Lawrence Berkeley National Laboratory
Recent Work

Title

Carotenoid Biosynthesis in Bacteria: {ital In Vitro} Studies of a {ital crt/bch} Transcription Factor from {ital Rhodobacter capsulatus} and Carotenoid Enzymes from {ital Erwinia herbicola}

Permalink

https://escholarship.org/uc/item/86z6p10d

Author

O'Brien, D.A.

Publication Date

1992-11-01

eScholarship.org Powered by the California Digital Library
University of California

LBL-33496
UC-408

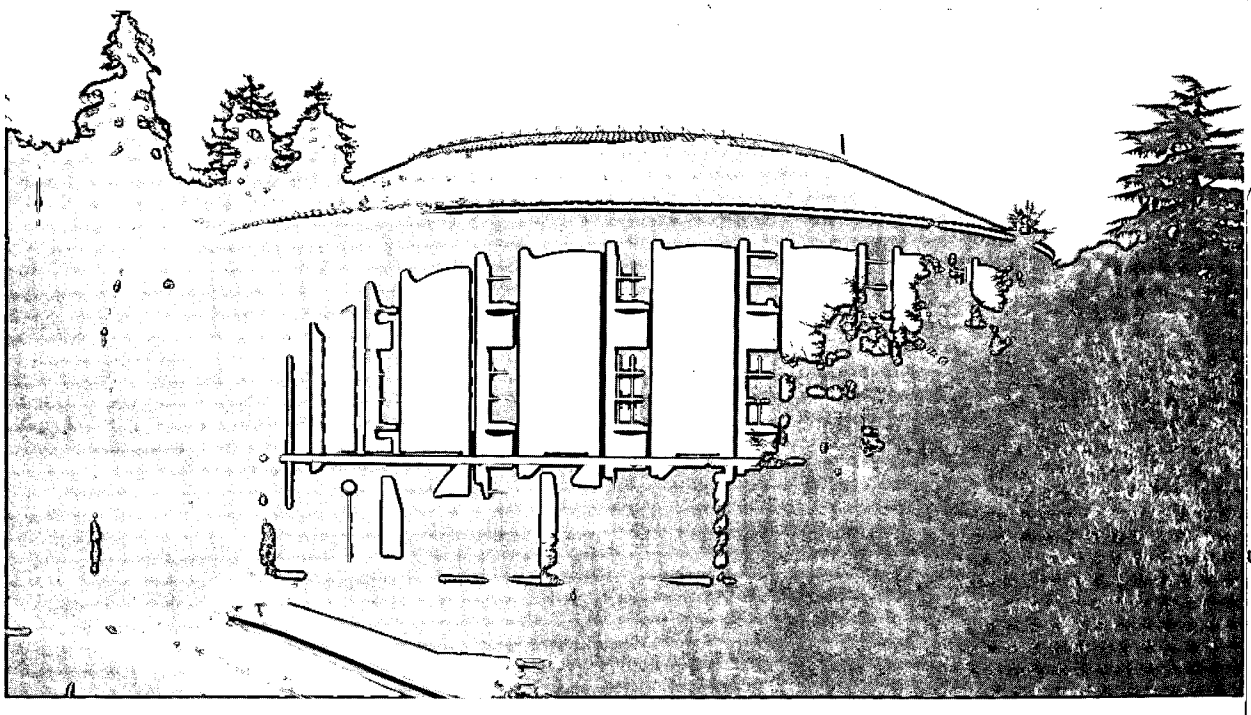
Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

CHEMICAL BIODYNAMICS DIVISION

Carotenoid Biosynthesis in Bacteria: *In Vitro* Studies of a *crt/bch* Transcription Factor from *Rhodobacter capsulatus* and Carotenoid Enzymes from *Erwinia herbicola*

D.A. O'Brien
(Ph.D. Thesis)

November 1992



| REFERENCE COPY |
| Does Not |
| Circulate |
LBL-33496
Copy 1
Bldg. 50 Library.

Prepared for the U.S. Department of Energy under Contract Number DE-AC03-76SF00098

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. Neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial products process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or The Regents of the University of California and shall not be used for advertising or product endorsement purposes.

Lawrence Berkeley Laboratory is an equal opportunity employer.

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

Carotenoid Biosynthesis in Bacteria: *In vitro* studies of a *crt/bch* transcription factor from *Rhodobacter capsulatus* and carotenoid enzymes from *Erwinia herbicola*

David Allen O'Brien
(Ph.D. Thesis)

Department of Chemistry
University of California

and

Chemical Biodynamics Division
Lawrence Berkeley Laboratory

November 1992

This work was supported by the Director, Office of Energy Research, Office of Basic Energy Sciences of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

Carotenoid Biosynthesis in Bacteria: In vitro studies of a crt/bch transcription factor from Rhodobacter capsulatus and carotenoid enzymes from Erwinia herbicola

by

David Allen O'Brien

Abstract

Carotenoid biosynthesis is ubiquitous in photosynthetic organisms, and occurs in some non-photosynthetic species as well. In studying carotenoid biosynthesis, the two bacteria, Rhodobacter capsulatus and Erwinia herbicola provide useful model systems for transcriptional regulation and enzyme biochemistry, respectively.

Erwinia herbicola, a non-photosynthetic plant pathogen, was selected for enzymology studies because of its similarity to and compatibility with Escherichia coli, and its lack of dependence on a specialized photosynthetic membrane or organelle. In this work, the first reported in vitro expression of individual post-phytoene carotenoid biosynthesis genes is presented. Erwinia lycopene cyclase was found to require an all-trans substrate, unlike some earlier reported plant chromoplast preparations in which inhibitors were used to focus on a particular reaction step. Erwinia β-carotene hydroxylase was found to require cofactors typical of other monooxygenases. Despite the hydrophobicity of both this enzyme and its substrate, higher in vitro activity was obtained in membrane-free preparations. The less common zeaxanthin glucosyl-transferase was discovered to be more robust in vitro than either of the other two enzymes. This enzyme is shown to be limited by zeaxanthin substrate availability and able to convert the mono-

glucoside intermediate much more efficiently to the diglucoside. This glucosyltransferase was found to be weakly associated with the cell membrane and subject to proteolysis when removed from the membrane. In addition, a UDP binding site is proposed based on amino acid sequence homology with other enzymes requiring UDP-activated substrates.

Rhodobacter capsulatus is a purple non-sulfur facultative phototroph with a 46 kb cluster of anaerobically induced photosynthetic genes, including all those required for carotenoid biosynthesis. As such, it provides an excellent vehicle for the study of transcriptional regulation, which is currently not well understood. Using a palindromic sequence found upstream of carotenoid and bacteriochlorophyll biosynthesis genes, a putative transcription factor has been identified. This protein binds the DNA site in a sequence-specific manner, and DNA-protein complex formation is cooperatively enhanced by the presence of an upstream AT-rich DNA sequence acting in cis. Both the binding constant and dissociation rate are estimated. The in vitro binding studies, together with parallel in vivo experiments, indicate that this factor acts as a repressor under aerobic growth conditions. Also, more limited evidence indicates a possible second role for this factor as an activator under anaerobic/ photosynthetic conditions. Finally, these results, when taken together with the distribution of putative regulatory sites throughout the gene cluster as well as the results of other researchers, suggest a possible mechanism of transcriptional regulation in which the expression of these pigment biosynthesis genes is coordinated with that of the structural proteins of the photosynthetic apparatus.

Table of contents

Acknowledgements

This thesis would truly not be complete without an acknowledgement of thanks for some of the other people who helped to make it possible. I am indebted to Greg Armstrong for much of the earlier work on carotenoids, and to both Greg and David Cook who introduced me to the "strange new world" of molecular biology. I wish to especially thank my collaborators, Bhupinder Hundle and Peter Beyer, with whom I worked on the overexpression and enzymology studies; and Zack Ma and Dave Cook for our collaborations on some of the transcriptional regulation studies. I also wish to acknowledge Don Burke, Ning Pon, Enrique Dalmasso, and Tariq Rana for their valuable help and suggestions, and Theresa Sullivan for the generally unsung task of keeping us supplied with solutions and glassware. We are all particularly indebted to Marie Alberti, not only for her energy and enthusiasm, but also for her massive sequencing effort, which has provided a major starting point and driving force for many of the projects in the lab, including my own.

I thank John Hearst, Ken Sauer, and Hatch Echols for their guidance on and critical reading of this thesis. I am also grateful to John for providing his input as my research director, while being patient enough to give me the independence to go out and make my own mistakes and learn from them, as well.

I also consider myself fortunate to have had the opportunity to have interacted with the other people at MCL, both scientifically and socially. Thanks to the friendship of Jeff Pelton and Joe Pease,

the first two years of classes and lab work were at least bearable. I thank the members of the Bartholomew Lab, particularly Jim, Hisao, and Catharina for my brief introduction into mammalian cell biology and for my more permanent appreciation of afternoon coffee. I am also grateful to Bill Scott and the members of the "woodshop" for their usually-welcome distractions. Also, I owe a special thanks to all the people I dragged to the beach on early weekend mornings to participate in sand castle contests.

Finally, I thank my parents and friends outside the lab (especially Lisa and Norman) for their moral support through both the good and bad times these last years.

Chapter 1:
Introduction to Carotenoids and their Biosynthesis

I. A general review of carotenoid structure, function, and distribution

Carotenoids are methyl-branched C_{40} compounds composed of poly-isoprene units, which have been further desaturated to produce a chromophore of conjugated double bonds. These colored compounds typically serve to protect plants as well as some bacteria and fungi from photo-oxidative damage. In addition, carotenoids are the pigments primarily responsible for the characteristic red and yellow colors of autumn leaves, ripened fruits, and flowers. Also, many birds, fish, insects, and crustaceans owe their colors to carotenoids found in their diets. Finally, dietary carotenoids such as β-carotene serve as metabolic precursors to vitamin A and provide the source of retinal and retinoic acid, compounds which play essential roles in animal vision and the regulation of cell differentiation. Carotenoids have commercial applications as well, primarily as vitamin supplements, pigments (notably as food coloring), animal food additives (e.g. in chicken feed to produce a desired egg yolk color), and the treatment of certain cancers (Mathews-Roth, 1987).

Although many of the over 500 carotenoids identified to date are basically linear, the most abundant ones, such as α- and β-carotenes and their derivatives, have undergone cyclization of their termini into six membered rings (Straub, 1987; Fig. 1-1A).

Carotenoids are divided into two classes by nomenclature, carotenes, which are strictly hydrocarbons, and their oxygen containing derivatives called xanthophylls. Naturally produced carotenoids amount to about 10^{11} kg each year, the vast majority being associated with green plants and other photosynthetic organisms.

Carotenoid pigments exhibit a characteristic type of absorbance spectrum with three peaks in the major band. Phytoene, the first C_{40} precursor to later carotenoids, absorbs in the UV range. The more common carotenoids absorb visible light and appear yellow, and then red as the chromophore is lengthened as more conjugated double bonds are introduced into the poly-isoprene chain (reviewed in Britton, 1983).

Carotenoids have been found to protect against photo-oxidative damage by three separate mechanisms: quenching of triplet photo-sensitizers to produce an unreactive triplet carotenoid species; reaction of carotenoids (with nine or more conjugated double bonds) with singlet oxygen to produce carotenoid triplets; and reaction with free radicals (such as lipid radicals) to form stable carotenoid radicals (reviewed in Krinsky, 1979; Cogdell and Frank, 1987; see Fig. 1-1B).

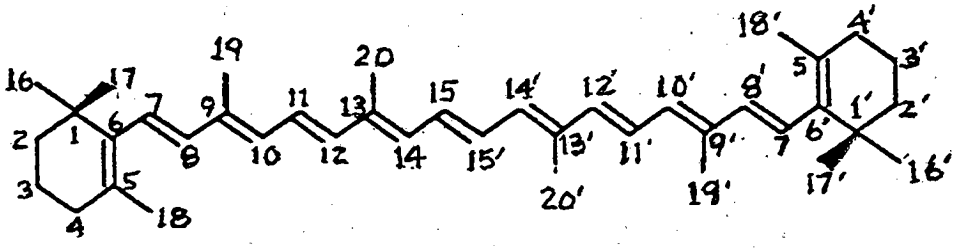
Particularly active and potentially damaging photosensitizing agents are the chlorophyll and bacteriochlorophyll molecules at the core of the pigment-protein complexes in the photosynthetic apparatus of all photosynthetic organisms. Although the architecture and location of these reaction center and light

Figure 1-1.

(A). Structure of β-carotene with systematic numbering of carbon atoms.

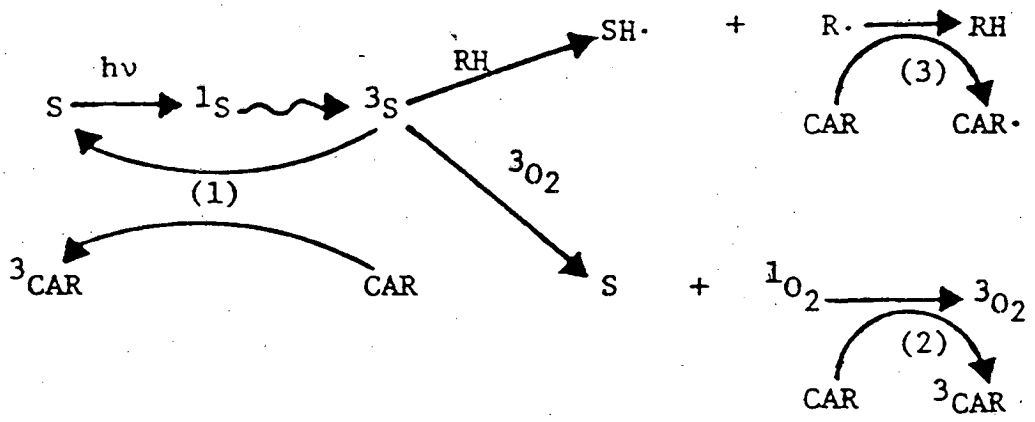
(B). Three mechanisms of carotenoid protection against photo-oxidative damage (from Krinsky, 1979): (1), reaction with triplet sensitizers; (2) the direct reaction with singlet oxygen; and (3), the reaction with free radicals to form less reactive carotenoid radicals

A



β-Carotene

B



harvesting complexes differ widely in plants, fungi, and bacteria, lipophilic carotenoids are an integral part of each (reviewed in Siefermann-Harms, 1984). Cogdell et al. (1981) calculated that the closely bound carotenoids in these pigment protein complexes are present at an effective concentration high enough to be approximately 90% efficient at quenching the excited chlorophyll. In addition, carotenoids have been shown to play a role in light harvesting in the wavelength range of 450 to 570, where chlorophylls do not absorb strongly, and to transfer this energy to chlorophyll species via a singlet-singlet mechanism (Cogdell and Frank, 1987)

In plants, the photosynthetic apparatus is located in leaf chloroplasts, where the carotenoid color is masked by the presence of chlorophylls. The plant reaction centers of Photosystems I and II contain carotenes, primarily β-carotene, as the bound carotenoids, while the light harvesting complexes utilize oxidized xanthophylls, primarily lutein, neoxanthin, and violaxanthin (reviewed in Siefermann-Harms, 1984). Carotenoids are also found association with the chloroplast envelope (Joyard et al. 1991). During leaf senescence, these xanthophylls often undergo esterification (Goodwin, 1958).

Flowers, pollen, fruit, and root tissue contain a huge variety of carotenoids, including common ones as well as the more exotic. The identity and distribution of these compounds are specific to each species and tissue. The more exotic compounds include oxidized keto-carotenoids and carboxylic acids, pentacyclic ring derivatives,

and C_{30} apocarotenoids (reviewed by Goodwin, 1980). More familiar examples of carotenoids occur in the tomato fruit, whose characteristic red color is due to lycopene, a linear carotene, and carrot roots, which contain β-carotene.

The reaction centers of Photosystems I and II in green, red, and brown algae, as well as cyanobacteria generally contain β-carotene, like their analogs in higher plants. However, some varieties of green algae utilize the similar α- or ε-carotenes instead. These "simpler" photosynthetic organisms contain vastly different antenna complexes, ranging from the carotenoid-less phycobilisomes of cyanobacteria to the internal light harvesting complexes of most algae and diatoms, containing violaxanthin and its keto- derivative, fucoxanthin (reviewed by Siefermann-Harms, 1984).

Green bacteria possess both a membrane-associated antenna as well as an external chlorosome, both of which contain carotenoids. The chlorobiaceae family utilizes chlorobactene (ϕ,ψ-carotene) in both the chlorosome and internal pigment protein complexes. However, the chloroflexaceae contain mostly β- and γ-carotenes, with smaller amounts of xanthophyll glucosides (Schmidt, 1980).

The purple bacteria possess a single photosystem, completely integrated into the thylakoid, a deeply invaginated intracytoplasmic membrane. These bacteria contain linear carotenoids which vary among different species, such as: spiroloxanthin in Rhodospirillum rubrum and Chromatium vinosum; rhodopin in Rhodopseudomonas acidophila, dihydroneurosporene in Rps. viridis; and sphaeroidene

in Rhodobacter sphaeroides and capsulatus (Van der Rest and Gingras, 1974; Drews et al, 1976).

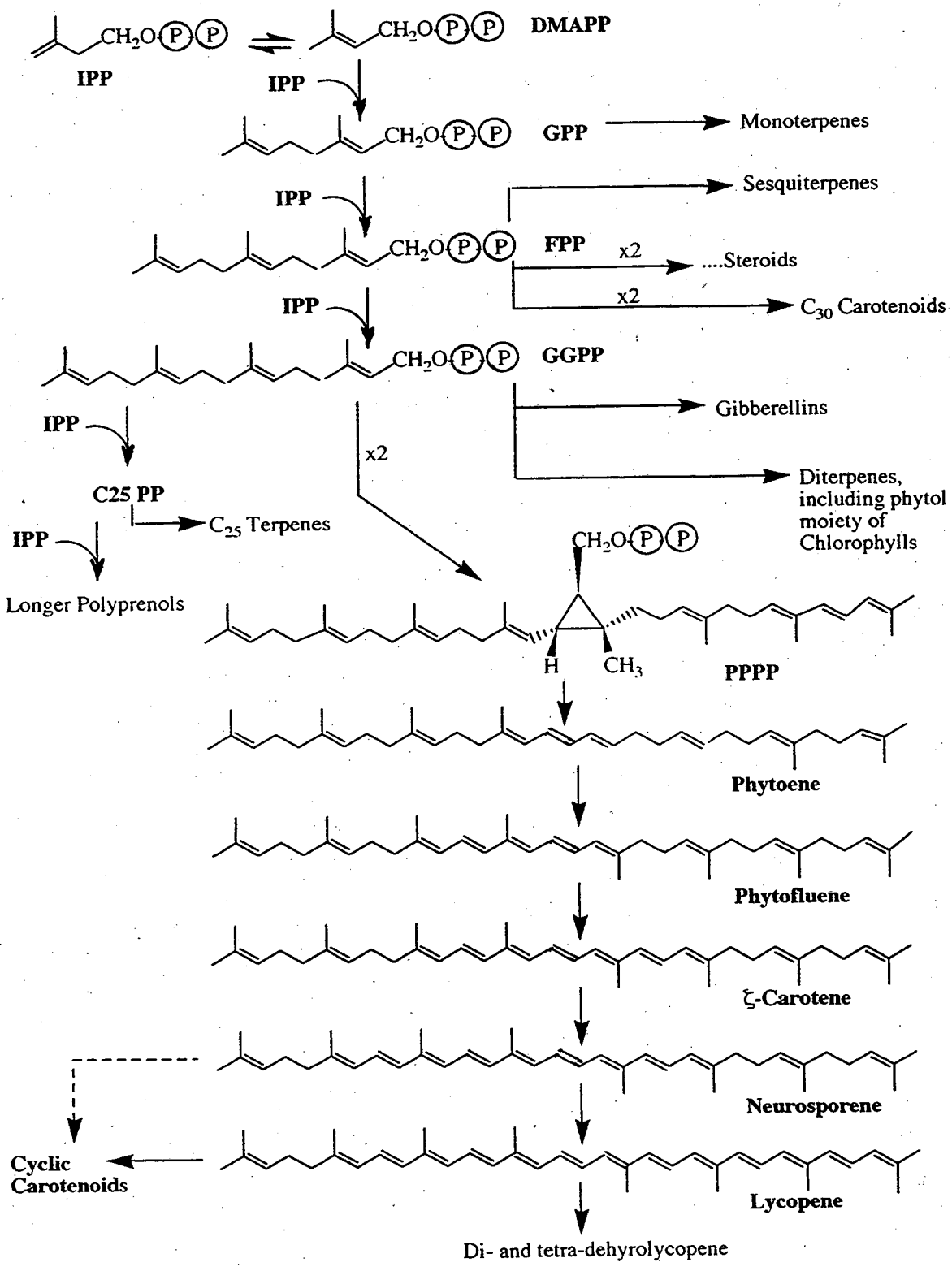
In addition to these photosynthetic organisms, carotenoids are also found in a number of non-photosynthetic fungi and bacteria, particularly ones inhabiting environments exposed to sunlight and oxygen.

II. Carotenoid biosynthesis

A conserved carotenoid biosynthetic pathway was first proposed by Porter and Lincoln (1950). The earlier steps of this pathway are common to all plant and bacterial systems, and consist of three condensations of isoprenyl pyrophosphate units to form the the twenty carbon geranylgeranyl pyrophosphate (GGPP), which is then dimerized to the first C_{40} carotenoid, phytoene. All of these reactions utilize soluble, phosphorylated substrates, and the enzyme activities for several of these reaction steps have been isolated in vitro from a variety of organisms (Poulter and Rilling, 1983; Qureshi and Porter, 1983; Kleinig, 1989). These steps, up to the formation of GGPP, are not unique to carotenoid formation, but are shared with other biological isoprenoid compounds, such as sterols, terpinol quinones, and gibberellins (Fig. 1-2).

The first biosynthetic step truly unique to carotenogenesis is the "tail-to-tail" condensation of GGPP to phytoene via the cyclopropyl intermediate, pre-phytoene pyrophosphate (PPPP). Either all-trans-phytoene or the 15-cis isomer results from the elimination of pyrophosphate, depending upon which of two protons is abstracted from the carbon skeleton. The 15-cis-phytoene is the predominant isomer produced in plants and other eukaryotes, while in the purple bacterium, Rhodobacter capsulatus, a 1:1 ratio of the isomers has been reported (Bramley and MacKenzie, 1988). A single 47.5 kDa phytoene synthase enzyme,

Figure 1-2. General pathway of carotenoid biosynthesis, indicating common intermediates with other poly-isoprenoid compounds (reviewed in Britton, 1983). The C_{40} compounds are shown in the all-trans configuration, although their isomeric forms differ widely among species. Phytoene, in particular is often found with its central 15-15' double bond in the cis configuration. Abbreviations used are: IPP, isopentenyl pyrophosphate; DMAPP, dimethylallyl pyrophosphate; GPP, geranyl pyrophosphate; FPP, farnesyl pyrophosphate; GGPP geranylgeranyl pyrophosphate; and PPPP, prephytoene pyrophosphate.


IPP
DMAPP
GPP
Monoterpenes
Sesquiterpenes
FPP
x2
....Steroids
x2
C_{30} Carotenoids
GGPP
Gibberellins
Diterpenes, including phytol moiety of Chlorophylls
C25 PP
C_{25} Terpenes
Longer Polyprenols
PPPP
Phytoene
Phytofluene
ζ-Carotene
Neurosporene
Lycopene
Cyclic Carotenoids
Di- and tetra-dehyrolycopene

catalyzing both the dimerization of GGPP and the subsequent dephosphorylation, has been purified from Capsicum annum (red pepper) chromoplasts by Dogbo et al. (1988). Two separate enzymes, designated crtB and crtE, have been reported in Rhodobacter capsulatus for the two reactions leading to phytoene (Armstrong et al., 1990a). However, the second enzyme, crtE, was found to bear considerable amino acid homology to the albino-3 gene product from Neurospora crassa, which codes for the last prenylation step forming GGPP (Nelson et al., 1989; Carattoli et al., 1991), as well as to the corresponding GGPP synthase in Erwinia herbicola (Math et al., 1992).

Subsequent enzymatic activities in the carotenoid biosynthesis pathway result in the wide range of end products particular to each species. These later enzymes utilize lipophilic substrates, and their activities are difficult to isolate in vitro. No homogeneous protein fractions catalyzing individual post-phytoene reactions have been isolated thus far. Most of the enzymes for carotenoid biosynthesis after phytoene appear to be membrane bound and have been postulated to be clustered into a multienzyme complex (Beyer et al., 1985). Due to the paucity of observable intermediates in the plant and fungal systems studied, it has been proposed that the carotenoid substrate may be channeled through this multienzyme complex (Beyer et al., 1985) and (Candau et al., 1991).

Four dehydrogenation steps are required to transform phytoene to lycopene (Fig. 1-2). Phytoene dehydrogenation is the target of a class of herbicides, and their effects as well as the effects

of mutations have been used to study carotenoid dehydrogenases. In some organisms, such as the fungus Phycomyces (Aragon et al., 1976) and the bacteria of the genus Erwinia, (Misawa et al., 1990) all four dehydrogenations appear to be catalyzed by a single enzyme. In other species, the analysis of mutant strains and the accumulation of intermediates have led researchers to suggest the presence of at least two separate enzymes, usually a phytoene dehydrogenase and a ζ-carotene dehydrogenase. Narcissus pseudonarcissus (daffodil) chromoplast carotenoids undergo trans removal of hydrogens in the two step conversion of phytoene to ζ-carotene, and cis dehydrogenations in both steps converting ζ-carotene to lycopene (Mayer et al., 1989). Also, a different sensitivity to herbicides was noted for the two processes in Narcissus as well as Aphanocapsa (Sandmann and Boger, 1989). Phytofluene is the accumulated intermediate in an apparently different two enzyme dehydrogenation process observed in spinach leaves (Kushwaha et al., 1969), tomato (Subbarayan et al.,1970), and Halobacterium cutirubin (Kushwaha et al., 1976), based on a different cofactor requirement for the two stages of the process. With only limited evidence, two competing mechanisms have been proposed for the dehydrogenations of carotenoids: simple abstraction of hydrogens; or oxidation followed by dehydration (Sandmann and Kowalczyk, 1989; Beyer et al., 1985).

Although most carotengenic organisms convert 15-cis-phytoene to all-trans-lycopene, no specific cis/trans isomerases have been identified by genetic lesions, or otherwise. Since isomerization occurs at different stages in the dehydrogenation

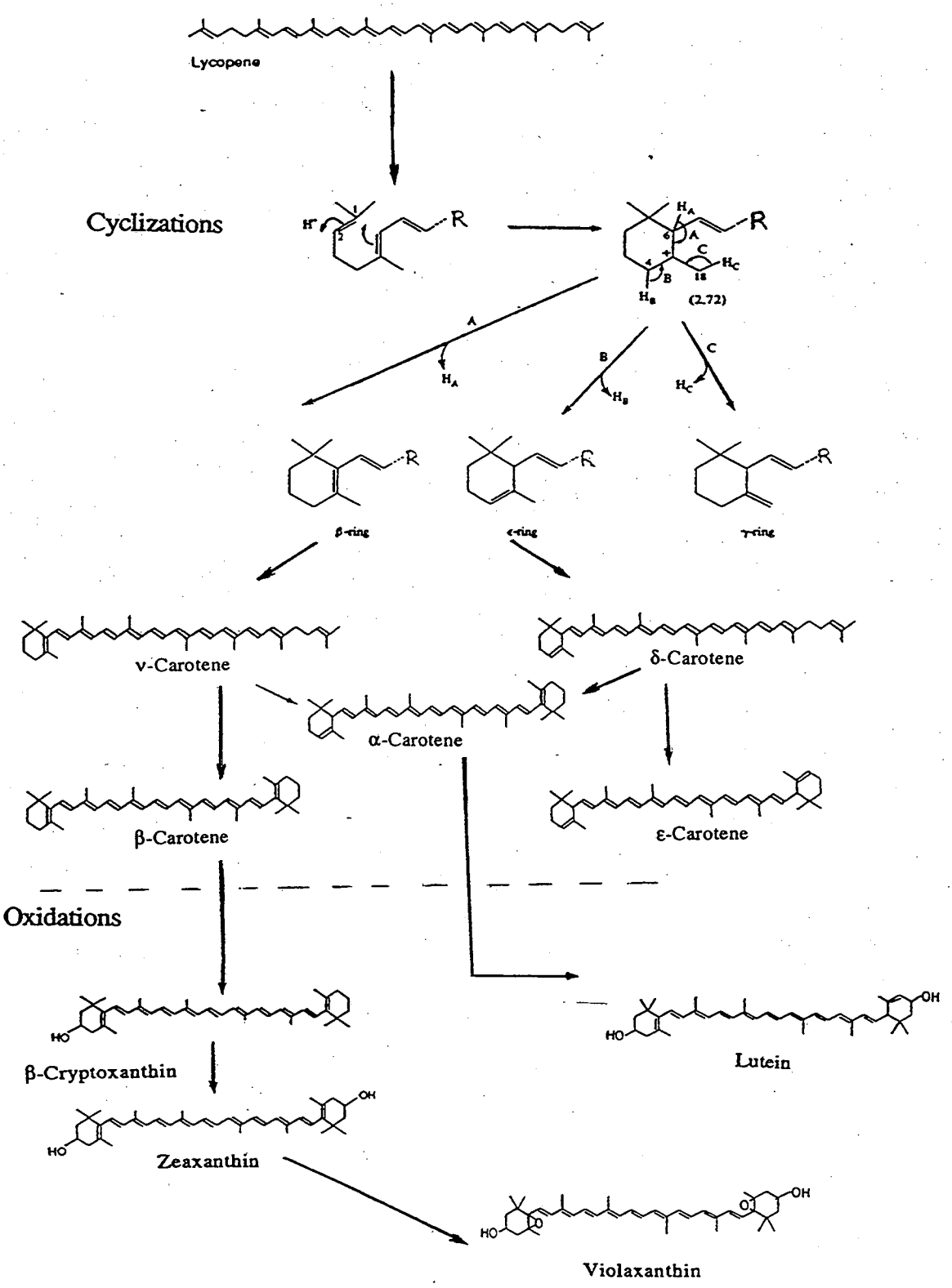
cascade with different organisms (Bramley and Sandmann, 1985; Beyer et al., 1985), it has been proposed that isomerization is accomplished upon binding of the carotenoid to a dehydrogenase (Goodwin, 1980).

Following the dehyrogenation reactions, lycopene can undergo cyclization of one or both of its termini, most commonly into β- or ε-rings. Cyclization can also occur with lycopene's immediate precursor, neurosporene, although it is less common. Cyclization also produces a shorter chromophore, resulting in a yellow carotenoid. The particular type of ring structure formed appears to depend on which proton is abstracted from a carbonium ion intermediate, and is dictated by the stereospecific enzyme (Britton, 1983; Fig. 1-3). Despite an inability to purify or over-express these membrane bound enzymes, certain individual reactions, such as cyclization, have been studied in relative isolation by the manipulation of cofactors (primarily O_2) to inhibit other enzymes in a crude isolate (Beyer et al., 1989). The combined carotene desaturation and cyclization activities have been demonstrated in vitro from the collection of membrane bound enzymes solubized from Narcissus pseudonarcissus (daffodil) chromoplasts and reconstituted into liposomes (Beyer et al., 1985). The preferred substrates for cyclization to β-carotene in the daffodil system were lycopene isomers containing cis double bonds at the 7 and 7' positions, and NADPH was found to be an essential cofactor (Beyer et al., 1989) and (Beyer et al., 1991). In the fungus, Phycomyces blakesleeanus, evidence of an enzyme aggregate containing two cyclases which convert lycopene to β-carotene via γ-carotene has been reported

(Candau et al., 1991).

A very wide variety of xanthophylls is formed via oxidation of both cyclic and non-cyclic carotenoids. The hydroxylation of β-carotene utilizes molecular oxygen, as demonstrated by $^{18}O_2$ labelling (Yamamoto et al., 1962) and (McDermott et al., 1974). In vitro experiments indicate that membranes of the cyanobacterium, Aphanocapsa, accomplish the hydroxylation of β-carotene using an O_2 dependent monooxygenase (mixed function oxidase), and that the reaction is stimulated by NADPH (Sandmann and Bramley, 1985). Two very common xanthophylls, zeaxanthin and lutein, both require oxidation of β-rings at the non-activated C-3 position, a somewhat unusual site preference for enzymatic oxidation (Walsh, 1976). However, monooxygenases have been reported to sometimes target other non-activated substrates, such as methane (Fox et al., 1989). The further oxidation of zeaxanthin to vioaxanthin is an olefinic epoxidation, which is also catalyzed by a monooxygenase requiring O_2 and NADPH (Siefermann and Yamamoto, 1975). Still other unrelated monooxygenases use O_2 to effect allylic oxidations of carotenoids to their ketone derivatives such as that of the non-cyclic sphaeroidene to sphaeroidenone in the purple bacteria of the genus Rhodobacter (Schneour, 1962). Finally, even non-carotenogenic organisms produce retinal by the oxidative scission of β-carotene, catalyzed by a dioxygenase. Other xanthophylls arise through the non-oxidative process of hydration of double bonds, examples being rhodopin and demethyl sphaeroidene (Britton, 1983).

Figure 1-3. Cyclization and oxidation pathways determining some of the more common end-product carotenoids (see text). The cyclization mechanisms shown are from the review by Britton (1983).


Lycopene
Cyclizations
(2.72)
β-ring
ε-ring
γ-ring
ν-Carotene
δ-Carotene
α-Carotene
β-Carotene
ε-Carotene
Oxidations
β-Cryptoxanthin
Lutein
Zeaxanthin
Violaxanthin

III. Rhodobacter capsulatus as a model system for the regulation of carotenoid biosynthesis

Rhodobacter capsulatus, formerly known as Rhodopseumonas capsulata, is a purple non-sulfur bacterium, generally found in ponds and other bodies of still water which contain suitable nutrients. As a facultative phototroph, this bacterium lives by aerobic respiration in the presence of oxygen, but with a drop in oxygen tension, the organism rapidly induces an intracytoplasmic membrane containing a complete photosynthetic apparatus (Drews and Oelze, 1981). The single photosystem found in this species and other purple bacteria is unable to evolve molecular oxygen from water. The R. capsulatus photosystem contains a reaction center and two light- harvesting complexes, LH-I (or B870) and LH-II (or B800-850), each of which include bacteriochlorophyll-a and the non-cyclic carotenoid, sphaeroidene (Clark et al., 1984; Klug et al., 1985). Energy is transferred in a downhill fashion from LH-II to LH-I to the reaction center.

However, even in the presence of oxygen, carotenoids are still produced at lower levels, giving R. capsulatus and its close relative, R. sphaeroides their characteristic red-purple color, primarily as the result of the formation of sphaeroidenone, the 2-oxo- derivative of sphaeroidene (Cohen-Bazire et al., 1957). In addition to aerobic and photosynthetic growth modes, R. capsulatus can also grow chemoheterotrophically in the absence of oxygen by using compounds such as dimethyl sulfoxide or an amine-oxide as electron acceptors (Yen and Marrs, 1977; Scolnik et al., 1980b).

Also, the organism can survive by anaerobic fermentation in the dark, but with a very long cell-doubling time (Schultz and Weaver, 1982).

All of the genes coding for the structural polypeptides in the photsynthetic reaction center and LH-I complex, as well as those coding for all of the enzymatic steps of bacteriochlorophyll and carotenoid biosynthesis, are clustered on the chromosome in a 46 kb region, which was discovered and mobilized on a large plasmid by B. Marrs (1981). This R-prime plasmid was designated pRPS404 (Marrs, 1981). This gene cluster provdes an excellent vehicle for the study of photosynthetic gene expression and regulation as a function of environmental changes. The cluster of genes was found to complement all known R. capsulatus photosynthetic mutations, but is not expressed in E. coli, or other common hosts, probably due to a lack of common transcription signals and the likely requirement of a specialized membrane for the assembly of the pigment-protein complexes. The multiple growth modes available to R. capsulatus make it possible to cultivate mutant strains which are not photosynthetically competent.

The genes in this photosynthetic cluster have been extensively mapped by mutational analysis (Yen and Marrs, 1976; Taylor et al., 1983; Zsebo and Hearst, 1984; Bauer et al., 1988), and subsequently the entire 46 kb cluster was sequenced by M. Alberti (Alberti and Hearst, 1991; EMBL Data Library accession number Z11165). Like other sequenced regions of R. capsulatus, this gene cluster is unusually GC-rich, having a GC content of approximately 67%. The organization of these genes is shown in figure 1-4, with

the highly expressed structural protein genes located in outwardly-transcribed operons at the ends of the cluster. A similar cluster of photosynthetic genes has also been found in R. sphaeroides (Lee and Kaplan, 1992).

The two structural polypeptides in the R. capsulatus LH-I antenna (α and β) and the L and M proteins of the reaction center are all encoded by the puf (photosynthetic unit, fixed) operon, while the third reaction center polypeptide, H, is encoded by the puh operon at the opposite end of the 46 kb cluster. The secondary antenna, LH-II, is induced by low light levels and contains its own α- and β- polypeptides encoded by an operon which is separate from the cluster and designated puc (Youvan and Ismail, 1985), as well as a third polypeptide, LH-II γ, arising from a gene of unknown location. An X-ray crystal structure of the R. sphaeroides reaction center indicates that the sphaeroidene is held in a hydrophobic pocket in a bent conformation, suggestive of an internal cis double bond in the carotenoid (Yeates et al., 1988). This is consistent with the 13-cis, twisted structure of the 1,2-dihydroneurosporene carotenoid in the high-resolution X-ray structure of the reaction center of Rhodopseudomonas viridis (Deisenhofer and Michel, 1988). In contrast, the carotenoids in light harvesting complexes are apparently held in an all-trans conformation (Cogdell and Frank, 1987).

The bacteriochlorophyll biosynthesis (bch) operons, as well as some others of unknown function, are located adjacent to the puf and puh operons, while the carotenoid biosynthesis (crt) operons are largely concentrated in the center of the photosynthetic cluster

(Fig. 1-4). The crt genes were mapped on this cluster using physical and genetic techniques with a wide variety of available Tn5.7 transposon, interposon, and point mutants (Yen and Marrs, 1976; Taylor et al., 1983; Scolnik et al., 1980; Zsebo and Hearst, 1984; Giuliano et al., 1988; Armstrong et al., 1990a; Fig. 1-5). These mutants were characterized by analysis of accumulated carotenoid intermediates, allowing enzymatic functions to be assigned to specific loci.

Transcription of the puf and puh operons has been studied extensively and found to be induced by a factor of at least 8 upon the reduction of oxygen tension (Cook et al., 1989; Wellington et al., 1991; Sganga and Bauer, 1992). However, the carotenoid biosynthesis genes are generally induced to a lesser extent, and the magnitude of this induction varies greatly among operons (Armstrong, 1989b). The sequencing of the 11 kb internal crt cluster (all the crt genes, except crtJ) has been reported by Armstrong et al. (1989), confirming the gene locations, organization and direction of transcription, as well as identifying the gene product amino acid sequences, potential ribosome binding sites, translation start sites, and rho-independent transcription termination signals (hairpin loop type). Also, Armstrong and co-workers identified several sequence motifs recurring upstream of the crt operons which could serve as promoters or other regulators of transcription. These include sequences bearing nucleotide homology to σ^{70}-like promoters (-35 and -10 regions), islands of non-homologous AT-rich sequence, and a specific palindrome or inverted repeat sequence with a consensus of GTGTAA N_8 TTACAC.

This palindrome bears some general homology to the consensus form found in a variety of prokaryotic transcription factors (Gicquel-Sanzey and Cossart, 1982; Buck et al., 1986; Armstrong et al., 1989), but is entirely different from the GC-rich palindrome found upstream of the puf structural operon, which deletion mutations have shown to be required for normal expression of puf (Adams et al., 1989).

The reason for the clustered arrangement of the photosynthetic genes is not completely understood, but several theories have been advanced. The notion that the entire cluster is somehow coordinately regulated by superhelicity changes throughout the region upon the removal of oxygen was effectively refuted by Cook et al. (1989).

However, several reseachers have demonstrated that overlapping transcriptional units exist at both ends of the 46 kb structure, and that read-through is necessary both for the proper expression of downstream operons and for the organism's rapid adaptation to photosynthetic growth (Young et al., 1989; Wellington and Beatty, 1989; Wellington et al., 1991; Bauer et al., 1991). Young et al. (1989) have demonstrated that transposon mutants which contain transcription terminators in the crtEF operon produced reduced levels of transcript from the downstream bchCXYZ operon (formerly bchCA; Burke et al., 1992; Alberti and Hearst, EMBL Data Library accession number Z11165). This suggested some type of read-through of transcription, which was confirmed by the S1 mapping of mRNA start sites of species extending into the downstream operon (Wellington and Beatty,1989). Also, the use of

S1 mapping and lacZ fusions has demonstrated that read-through from bchC into the puf structural operon is important both in providing significant amounts of total puf transcript (especially during aerobic growth), and in stimulating initiation from the puf promoter itself (Wellington et al., 1991). A similar organization of genes into a superoperonal arrangement was found at the opposite end of the 46 kb cluster in which transcription extends from the operon beginning at bchB into the puh operon (coding for the reaction center H protein), and the insertion of an interposon into bchB reduces the expression of puhA (Bauer et al., 1991). It appears that the basal level of transcription of structural proteins provided by read-through from pigment biosynthesis operons under aerobic conditions facilitates the rapid adaptation of R. capsulatus to photosynthetic growth when oxygen is subsequently removed. However, the mechanisms of transcriptional regulation of these essential upstream pigment biosynthesis genes themselves remain largely unknown.

Figure 1-4. Rhodobacter capsulatus photosynthetic gene cluster as sequenced by Alberti and Hearst (EMBL Data Library accession number Z11165, 1991) and the genes mapped to nucleotide resolution by Burke et al. (1992). The gene functions in this photosynthetic cluster have been extensively mapped by mutational analysis (Yen and Marrs, 1976; Taylor et al., 1983; Zsebo and Hearst, 1984; Bauer et al., 1988). The resolution of the former bchA locus into three separate genes is reported by Burke et al. (1992)

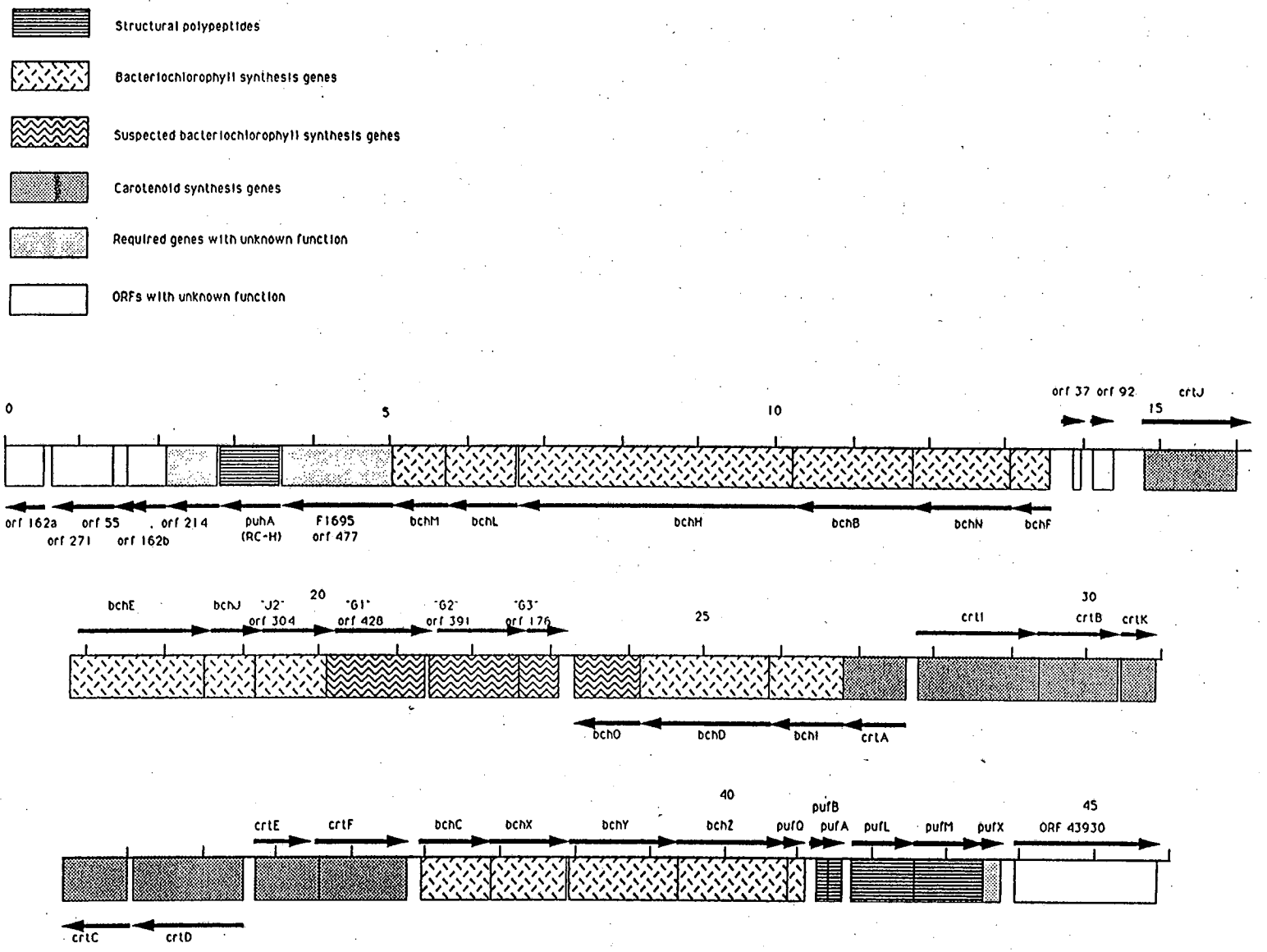
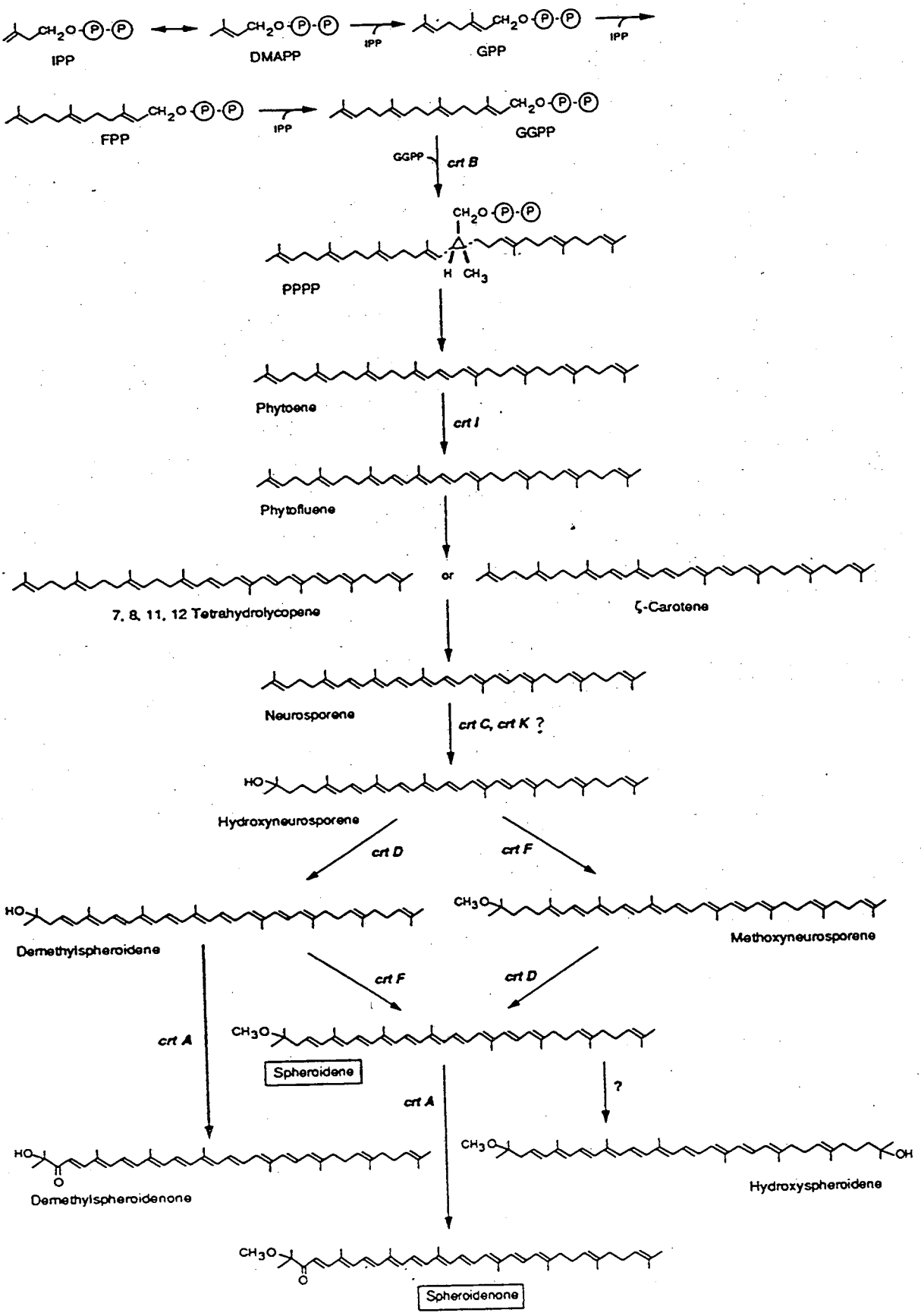

Structural polypeptides
Bacteriochlorophyll synthesis genes
Suspected bacteriochlorophyll synthesis genes
Carotenoid synthesis genes
Required genes with unknown function
ORFs with unknown function
0
5
10
15
orf 37
orf 92
crtJ
orf 162a
orf 271
orf 55
orf 162b
orf 214
puhA
(RC-H)
F1695
orf 477
bchM
bchL
bchH
bchB
bchN
bchF
20
25
30
bchE
bchJ
"J2"
orf 304
"G1"
orf 428
"G2"
orf 391
"G3"
orf 176
crtI
crtB
crtK
bchO
bchD
bchI
crtA
40
45
crtE
crtF
bchC
bchX
bchY
bchZ
pufQ
pufB
pufA
pufL
pufM
pufX
ORF 43930
crtC
crtD

Figure 1-5. Rhodobacter capsulatus carotenoid biosynthesis pathway as reported by Armstrong et al. (1989). Both 7,8,11,12-tetrahydrolycopene and ζ-carotene have been proposed as possible intermediates by Giuliano et al. (1986, 1988) Biochemical functions were proposed for crtA, crtC, crtD, and crtF by Scolnik et al. (1980a), crtI by Giuliano et al (1986), and crtB and crtE by Armstrong et al.(1990a). Abbreviations used are: IPP, isopentenyl pyrophosphate; DMAPP, dimethylallyl pyrophosphate; GPP, geranyl pyrophosphate; FPP, farnesyl pyrophosphate; GGPP geranylgeranyl pyrophosphate; and PPPP, prephytoene pyrophosphate.


IPP
DMAPP
IPP
GPP
IPP
FPP
IPP
GGPP
GGPP
crt B
PPPP
Phytoene
crt I
Phytofluene
or
7, 8, 11, 12 Tetrahydrolycopene
ζ-Carotene
Neurosporene
crt C, crt K ?
Hydroxyneurosporene
crt D
crt F
Demethylspheroidene
Methoxyneurosporene
crt F
crt D
crt A
Spheroidene
crt A
?
Demethylspheroidenone
Hydroxyspheroidene
Spheroidenone

IV. Overview of the thesis

Chapters 2 and 3 describe the characterization of a putative transcription factor in Rhodobacter capsulatus which binds upstream of the crt and bch pigment biosynthesis operons and appears to play a role in the adaptation of the organism from the aerobic to the anaerobic-photosynthetic growth mode. Chapter 2 describes the identification of this factor through an in vitro mobility shift assay, as well as the determination of its binding properties and sequence specificity. Chapter 3 focuses on the isolation of this factor.

Chapters 4 and 5 describe the overexpression and biochemistry of later carotenoid biosynthesis enzymes derived from the non-photosynthetic bacterium, Erwinia herbicola. Chapter 4 describes the separate overexpression and in vitro analysis of two enzymes involved in the main sequence of the carotenoid biosynthesis pathway, lycopene cyclase and β-carotene hydroxylase. Chapter 5 examines the overexpression and enzymology of functionally active zeaxanthin glucosyltransferase, an enzyme which carries out a more unusual transformation, converting a carotenoid into its more hydrophilic mono- and diglucoside derivatives. In addition, amino acid homology with other glucosyltransferases suggests a putative binding site for the UDP-activated glucose substrate.

Chapter 2:

Identification and *In vitro* Characterization of a Sequence- Specific DNA-Binding Protein, an Apparent Transcription Factor for the Pigment Biosynthesis Genes in *Rhodobacter capsulatus*

I. Introduction

The carotenoid biosynthesis genes in the purple non-sulfur photosynthetic bacterium *Rhodobacter capsulatus* have been extensively mapped and sequenced, and the levels of transcription measured for a variety of environmental growth conditions. However, the actual mechanisms of transcription initiation and regulation have remained largely obscure. Like the *puf*, *puh* , and *puc* operons which encode the structural polypeptides of the photosynthetic apparatus, and the various bacteriochlorophyll (*bch*) operons, many of the carotenoid biosynthesis (*crt*) operons appear to lack recognizable promoters and other transcriptional control signals (Armstrong et al., 1989).

Oxygen tension is a major factor controlling the coordinate expression of these genes. At least part of the oxygen-dependent regulation of pigment biosynthesis occurs at the transcriptional level (Clark et al., 1984; Zhu and Hearst, 1986; Sganga and Bauer, 1992). Although a great deal has been learned recently about the regulatory mechanism of structural polypeptide biosynthesis, much less is known about regulation of pigment biosynthesis.

Although the genes encoding the structural polypeptides and pigment biosynthesis enzymes are triggered by reduced oxygen

tension, several lines of evidence indicate that the regulatory mechanisms for these two classes of operons may be different. First, transcription of puf and puh is highly induced when cells are shifted from aerobic to anaerobic conditions (Sganga and Bauer, 1992; Cook et al., 1989; Wellington et al., 1991), whereas the induction of most bch and crt operons is less than 5-fold (Armstrong, 1989b; Wellington et al., 1991; Sganga et al., 1992; Fig. 2-1). Second, mutations in a newly discovered trans-regulatory factor, RegA, eliminate anaerobic induction of structural polypeptide biosynthesis but do not have any significant deleterious effect on pigment gene expression (Sganga et al., 1992).

Finally, Armstrong et al. (1989) have noted three conserved DNA sequence motifs that are found in the regions upstream of many pigment biosynthesis genes, but are not found upstream of the puf and puh operons. In addition to the presence of non-homologous AT-rich regions upstream of crt and bch operons, two recurring types of specific sequences were found as part of a homology search for possible promoters (Armstrong et al., 1989). One such sequence exhibits homology to a σ^{70}-like promoter site, with -35 and -10 regions. The other recurring sequence appeared to be new, and upon further inspection it was noticed (by D.A.O.) that the consensus nucleotides formed an inverted repeat or palindrome, and could serve as a protein binding site. This palindrome and its similarity to the conserved features of DNA binding sites for known prokaryotic transcription factors are noted in Armstrong et al. (1989) and shown in figure 2-2.

Recent studies have focused on the superoperonal organization

of pigment biosynthesis and structural genes and the functional importance of this organization for the adaptation from aerobic to anaerobic environmental conditions (reviewed in Wellington et al., 1992). Data from a number of researchers indicate that expression of the pigment biosynthesis genes enhances the expression of downstream, stringently controlled structural genes through both transcriptional read-through and an as yet poorly understood activation of the downstream operons themselves (Wellington et al., 1991; Young et al., 1989; Bauer et al., 1991). R. capsulatus strains in which the superoperon is inactivated by polar mutations are significantly impaired during the adaptation to photosynthetic growth (Wellington, 1991). Despite the apparent importance of transcription initiation of pigment biosynthesis genes, both for their own expression and for expression of downstream structural gene products, the DNA regulatory sequences controlling transcription have not yet been identified in detail for any pigment biosynthesis operon.

Perhaps the most intensively studied region from a genetic/ regulatory point of view is the superoperon that begins at the crtEF operon, and includes the bchCXYZ (formerly bchCA, Burke et al., 1992; Alberti and Hearst, EMBL Data Library accession number Z11165) and pufQBALMX operons (Wellington et al., 1991; Young et al., 1989). This crtEF superoperon is also of interest because it contains multiple examples of the recurring sequences: two copies of the palindrome sequence, a perfect match to the consensus sequence upstream of crtEF and one with a single mismatch upstream of the bch operon; a σ^{70}-like promoter sequence in front

of bchC; and a pronounced AT-rich region further upstream of bchC (Armstrong et al., 1989). Transcription initation at the crtEF promoter is increased approximately 6-fold by reduced oxygen tension (Armstrong, 1989b), while initiation at the bchCXYZ is increased by only a factor of 3 to 4 (Wellington and Beatty, 1989; Young et al., 1989; Armstrong, 1989b). Giuliano et al. (1988) identified the 5' ends of two closely-spaced mRNA transcripts upstream of the crtEF operon by S1 mapping. Young et al. (1989) mapped the promoter and regulatory signals for the bchCXYZ operon to a 134 bp fragment by deletion analysis. Wellington and Beatty (1989) identified the 5' end of an mRNA species from this same region, which they proposed resulted from transcription initiation from a site with similarity to a sequence upstream of the puf operon. However, Armstrong et al. (1989) suggested that the sequences in this region with similarity to an E. coli-type σ^{70} promoter and the putative prokaryotic transcription factor binding sites might instead function to regulate bchCXYZ expression.

The experiments in this chapter demonstrate that the recurring palindrome sequence does indeed serve as a specific protein binding site in R. capsulatus. A gel mobility-shift assay demonstrates that a DNA-protein complex forms at both the inverted repeat sequences found upstream of crtE and bchC, and that the stability of this protein is further enhanced in the case of the bchCXYZ operon through cooperative binding to the AT-rich region as well as the palindrome sequence. Additional data indicate that this complex may repress transcription under aerobic conditions. Based on these results, a general model for regulation of pigment biosynthesis genes in R. capsulatus is proposed.

Figure 2-1. Map of the 11 kb carotenoid gene cluster as determined by Armstrong et al.(1989) showing anaerobic induction of transcription (Armstrong, 1989b) as well as important features of the nucleotide sequence. Above the gene map is a graph showing the average A+T nucleotide content, as calculated by averaging over a 151 bp window using 10 bp increments. Only values above 40% are shown (the genomic mean AT content is 33%). Induction as a function of shifting from aerobic to anaerobic/ photosynthetic growth conditions was determined by mRNA dot blot analysis (Armstrong, 1989b) and summarized as shadings on the gene map as follows: 5-10 fold induction, black; 2-5 fold induction, gray; no induction, white. Immediately above the map, black circles indicate the position of the recurring palindrome motif and the σ's indicate the σ^{70}-like sequences, with putative transcriptional directions indicated by the small arrows. Note that these three sequence motifs are all segregated about the 5' ends of genes.

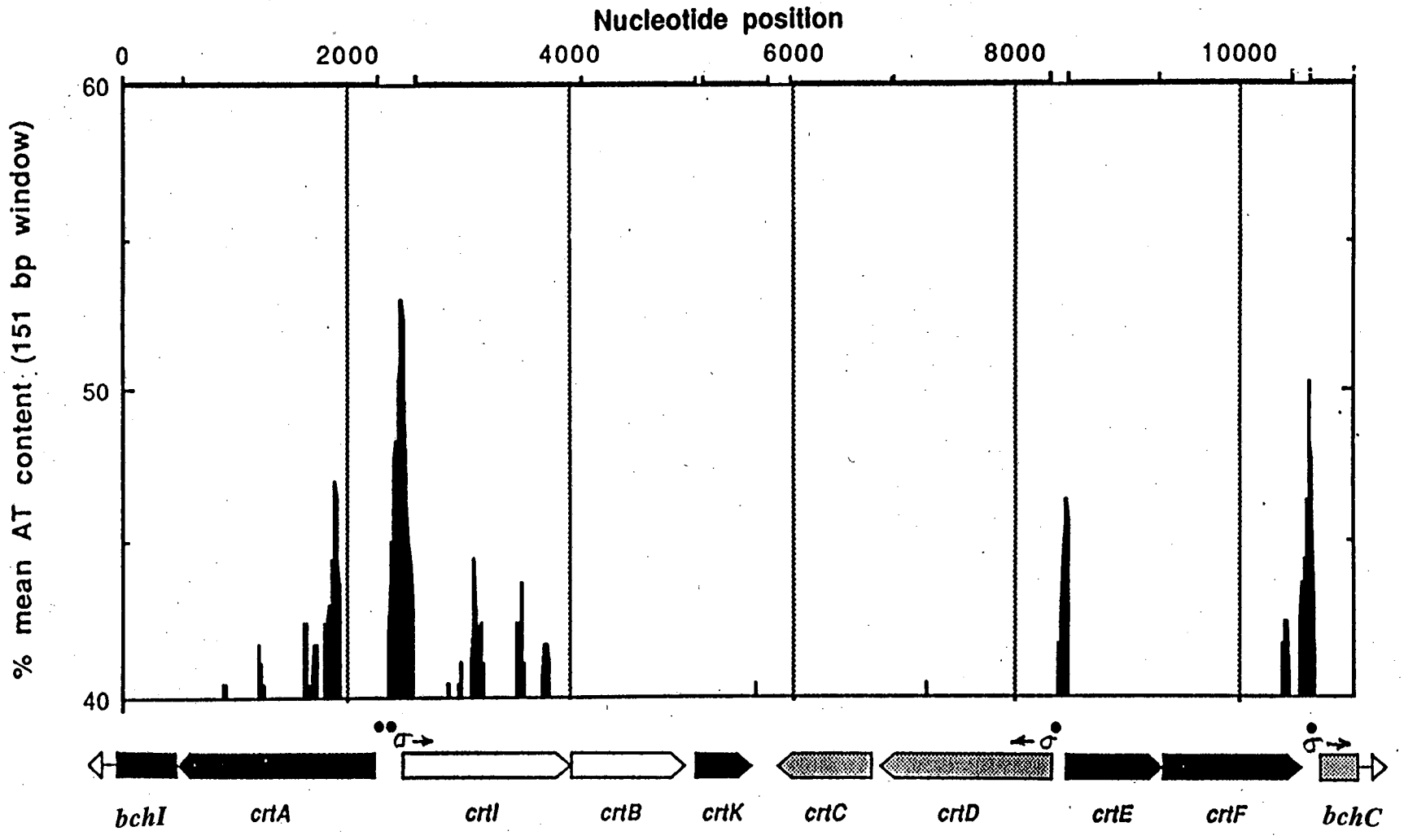
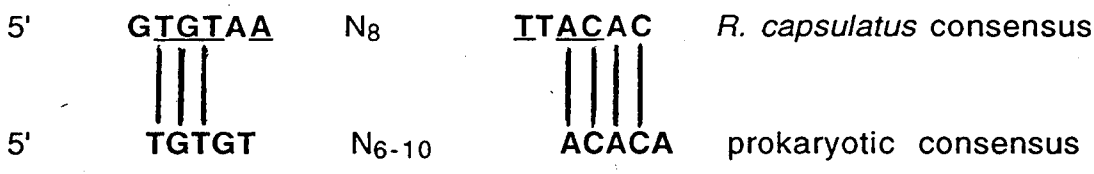

Nucleotide position
0
2000
4000
6000
8000
10000
60
50
40
% mean AT content (151 bp window)
bchI
crtA
crtI
crtB
crtK
crtC
crtD
crtE
crtF
bchC

Figure 2-2. The location of palindromes and σ^{70}-like sequences in the photosynthesis gene cluster of R. capsulatus (Armstrong et al., 1989; M. Alberti and J. E. Hearst, EMBL Data Library accession number Z11165). (**A**) Palindrome sequences found 5' to bch, crt and puc operons in R. capsulatus and the consensus derived from them. Conserved nucleotides matching the consensus are shown in bold type. Note that the crtD and crtE genes are transcribed in opposite directions (see Fig. 2-1) and share a single upstream palindrome. These sequences are compared to the consensus sequence of many procaryotic transcription regulators, including CAP, LacI, AraC, GalR, LexA, and NifA (Gicquel-Sanzey and Cossart; 1982; Armstrong et al., 1989); (**B**): Comparison of the σ^{70}-like sequences found 5' to bch and crt operons in R. capsulatus with the consensus σ^{70} promoter of E. coli N represents any nucleotide. The right half of the figure indicates the number of base pairs between these σ^{70}-like sequences and the start codon of the nearest downstream gene. Note that both the -35 and -10 sequences and also the spacing between them are conserved in all three cases.

A

5' GA	**GTGTAA**	GTTTTCAT	**TGACAC**	TTT	... 60 bp ... *bchF*
5' AC	A**TGT**C**A**	ACTGAGGT	**TTACAC**	CTA	... 19 bp ... *bchE*
5' AG	A**TGTAA**	ATATCCCG	**TTACAC**	ATC	... 17 bp ... *crtA*
5' AG	T**TGTAA**	ATCGGAAT	**TGAC**GA	CCT	..139 bp ... *crtI*
5' CT	**GTGTAA**	ACTGAAAC	**TTACAC**	CCA	... 66 bp ... *crtD*
5' GG	**GTGTAA**	GTTTCAGT	**TTACAC**	AGG	... 48 bp ... *crtE*
5' GC	**GTGTAA**	GTTCAATG	A**TACAC**	ACA	... 48 bp ... *bchCXYZ*
5' CA	**GTGTAA**	GCCCGACT	**TTACAC**	TTG	..140 bp ... *pucB*



B

5' TTGTTA	N_{16}	TATCAT	...35 bp... *crtI*
5' TTGGCA	N_{16}	TAAACT	...77 bp... *crtD*
5' TTGACA	N_{16}	AATGAT	...54 bp... *bchCXYZ*
5' TTGACA	N_{15-19}	TATAAT	*E. coli* consensus

II. Materials and Methods

Bacterial strains and growth conditions

The R. capsulatus strain used in these experiments was B100 (Marrs, 1981). R. capsulatus cells were grown under aerobic conditions at 30-32°C until a cell density of about 1.7 X 10^8 per ml was reached prior to harvesting or shifting to photosynthetic growth conditions. Photosynthetic cultures were then shifted to anaerobic conditions in the presence of 500 W/m^2 light provided by a bank of lumiline lamps for 60 minutes prior to harvesting. Dark, aerobic cultures were grown in covered glass vessels which were sparged with a mixture of $N_2:O_2:CO_2$ (80:20:2). Anaerobic cultures were uncovered and sparged with a mixture of $N_2:CO_2$ (80:2). Gas flow and composition were controlled using a Matheson Gas Products Multiple Dyna-blender, Model 8219. Growth rates were monitored on a Bausch and Lomb Spectronic 21 spectrophotometer.

Probe design and construction

Complementary synthetic DNA oligonucleotides were synthesized to reconstruct the completely conserved example of the consensus palindrome sequence found upstream of the crtEF operon (Armstrong et al., 1989). Approximately 1pmol of each strand was 5' end-labelled using T4 Kinase and γ-[^{32}P] ATP according to an established protocol (Sambrook et al., 1989). Following this reaction, the enzyme was heat-deactivated, the complementary strands

annealed, and the double stranded 31mer probe separated from the [^{32}P] ATP on a Sephadex G-50 spin column.

Double stranded probes of 80 to 210 base pairs in length were obtained by polymerase chain reaction (PCR) amplification of cloned plasmids containing the palindrome and surrounding intergenic DNA from either the region upstream of the crtEF or bchCXYZ operons. DNA upstream from bchCXYZ was cloned from a SmaI site at a position -85 relative to the 5' mRNA end mapped by Wellington and Beatty (1989) to an ApaI site at position +79. This pBR-derived construct, obtained courtesy of D.N. Cook and designated pDC410, contains the σ^{70}- like promoter, AT-rich region as well as the palindrome sequence. All site-specific mutations were constructed by PCR mutagenesis methodologies (Ho et al., 1989), utilizing primers containing the desired mutations for separate amplifications in each direction from the mutation site, followed by full length extension of the combined re-annealed products. Construction of the bchC promoter mutant probes was performed by D. Ma, and also used for a parallel in vivo promoter analysis (Ma et al., 1992). The nucleotide sequence of each mutation was verified by double-stranded sequencing. In a similiar manner, the entire intergenic region between the opposing genes, crtD and crtE, was amplified by PCR as a 200 bp fragment from pFL268, a plasmid obtained from F. Leach and M. Alberti. PCR amplifications were performed for 30 cycles, each consisting of a 1 min strand denaturation step 94°C, a 1.5 min annealing step at 40°C, and 2 min temperature ramp to a 1min, 72°C primer extension step. Following gel purification, these longer probes were end-labelled and isolated

on a Sephadex G-50 column as described above.

Mobility shift assays

R. capsulatus cells were grown as described above and harvested by first cooling the vessel in an ice-water bath for 15 min while continuously sparging with the gas mixture used for cell growth. This was followed by centrifugation for 15 min at 6000 x g at 4°C. Each culture was resuspended at 4°C in 1% of its original volume in 50 mM Tris-HCl, pH 7.5, 25 mM KCl, 5mM EDTA, 20% (v/v) glycerol, 0.1% Surfact-amp non-ionic surfactant (Pierce), 1 mM dithiothreitol, and 0.5 mM phenylmethylsulfonyl fluoride (PMSF) (Sigma). Cell lysates were obtained by either sonication or one passage through a French pressure cell at 13,000 psi. Large cell debris and unbroken cells were removed by centrifugation at 4500 x g for 10 min. Lysates were then divided into aliquots, frozen in liquid N_2, and stored at -70°C. Protein concentrations were determined using a Bradford assay (Bradford, 1976).

For mobility shift assays, 2 µl of [^{32}P]-labelled DNA probe (approximately 20 fmol) was added to a mixture of 8 µl non-specific competitor DNA and 10 µl cell lysate containing varying amounts of protein. Unless stated otherwise, the competitor DNA sample contained 500-fold excess (by weight) of poly (dI·dC). For the binding competition experiments, a 5 to 200-fold molar excess of specific unlabelled competitor DNA was also included. After incubation for 20 min at room temperature, samples were loaded on a native polyacrylamide gel, run for approximately 4 h at 8 V/cm, and then visualized by autoradiography or exposure on a

Molecular Dynamics Phosphor-Imager.

Phophatase treatment

Protein extracts were subjected to alkaline phosphatase treatment according to the procedure of Simmons et al. (1986). Alkaline phosphatase (type VII; Sigma) was added to a 20 mM Tris, 5 mM $MgCl_2$ solutions containing lysates or protein fractions and incubated at 30°C for 30 min. The pH of each solution was adjusted to 7.6, 8.0, or 8.5 by the addition of Tris base, and the free EDTA in the protein fractions was titrated out by additional $MgCl_2$. The enzyme was introduced, both directly in the salt-stabilized form and with prior dialysis, at levels ranging from 20 to 160 standard units. For all experiments, parallel incubations were carried out in the absence of enzyme, as a negative control.

III. Results

A. A protein binds the palindrome sequence found upstream of the crtEF and bchCXYZ operons in a sequence specific manner.

Gel mobility shift assays were performed with R. capsulatus extracts using either short synthetic oligonucleotides or longer plasmid-derived DNA's as probes (see Materials and Methods section). A stable DNA-protein complex was observed on a native polyacrylamide gel when cell lysates were incubated with a [^{32}P] labelled double stranded DNA 31mer containing a region found upstream of the crtEF operon. This particular DNA contains an example of the palindrome sequence which is an exact match with the consensus palindrome found upstream of various pigment biosynthesis operons (Armstrong et al., 1989). The shifted complex was resistant to competition by a large excess of non-specific DNA, but was successfully competed away by a lesser amount of the specific 31mer (Fig. 2-3). Protease treatment effectively eliminated the shifted band. The formation of the complex with this small probe which contains little DNA in addition to the palindrome sequence, and the resistance of the complex to excess sheared DNA as well as to non-specific single and double strand oligonucleotides, demonstrates that the complex is formed by a sequence-specific DNA binding protein.

The same type of complex formation was also observed when a 200mer containing the 85 base pair regulatory region upstream

of the bchCXYZ operon was used as a probe (Fig. 2-4). However, this probe, which contains the σ^{70}- like sequence and AT-rich region as well as the palindrome, appeared to bind protein much more efficiently than the crtEF 31mer. This tighter binding did not appear to be due simply to better electrostatic contacts to a longer piece of DNA, since the use of a 200mer containing the entire intergenic region between crtD and crtE did not result in a greater proportion of shifted counts. Both the crtEF 31mer and the longer bch probe effectively competed with each other in the gel shift assay, indicating that both probes bind to the same protein.

B. Effects of aerobic versus anaerobic cell growth conditions and other environmental factors on protein binding

In order to directly compare the binding efficiency of aerobic and anaerobic cell extracts, R. capsulatus was grown under vigorous aeration as described in the Materials and Methods section. After harvesting a portion of the culture for preparation of an aerobic extract, the remaining cells were shifted to anaerobic conditions in the light and grown for an additional 60 min. During this interval, cell growth is significantly slowed while the culture adapts to the newly required metabolic conditions and genes for photosynthesis are activated (Cook et al.,1989). Sampling the anaerobic cells during this adaptation period facilitates a direct comparison between aerobic and anaerobic states, since the same culture is used over a short time interval to make both extracts.

Under identical assay conditions with the bchCXYZ 200mer DNA probe, a higher yield of complex was obtained using the aerobic cell lysate than using the shifted, anaerobic lysate (Fig 2-5). This implied that the complex may function to repress transcription initiation at the bchCXYZ promoter under aerobic growth conditions. However, no significant difference was observed between aerobic and anaerobic lysates when the crtEF probes were used, despite the greater induction of transcription observed at this operon upon shifting to anaerobic growth conditions (Armstrong,1989b; Fig. 2-1).

To test the possibility that the DNA binding protein itself is oxidized or reduced to alter its binding characteristics, incubations were performed in reducing or oxidizing environments. Levels of dithiothreitol as high as 50 mM produced no discernable difference in binding patterns, although complex formation was essentially eliminated with 100 mM DTT. Binding was not effectively altered by the use of two common oxidizing agents, hydrogen peroxide and hydroxylamine. Also, no effects were observed upon the addition of the divalent cations, Mg^{2+} and Ca^{2+}, or ATP (with or without Mg^{2+}).

Phosphorylation is known often to play an important role in the regulation of transcription, both in eukaryotic systems (Baumann and Hand, 1982) as well as in the adaptive response of bacteria to changing environmental demands (reviewed in Stock et al., 1989). A classic example of a two-component system of regulation by a phosphorylating kinase plus a DNA binding protein is nitrogen fixation (Keener and Kustu, 1988). Phosphorylation has been proposed by Taremi and Marrs (1990) to regulate the switch between activation and repression of transcription at the puf

operon by binding in two different modes to a GC-rich palindrome (which bears no homology to the palindrome upstream of the pigment genes). Using the phosphatase assay of Simmons et al., 1986), no major change was observed in the fraction of probe bound or the pattern of binding to the bchCXYZ 200mer in this study.

C. Importance of the palindrome sequence and the AT-rich region in protein binding to the bchCXYZ control region - A mutational analysis.

To identify which sequences were involved in complex formation, binding experiments were performed using DNA probes with a variety of mutations in the promoter region (constructs shown in Fig. 2-6A). A gel mobility-shift assay using a DNA probe containing a point mutation in the -35 region of the σ^{70}-like promoter produced a DNA-protein complex of identical mobility to the one observed with the wild-type DNA sequence (Fig.2-6B, lanes 2 and 3). Therefore, the shifted band did not arise from the binding of RNA polymerase to the σ^{70}-like promoter. Futhermore, this complex did not appear to assemble with probes which contained substitutions within either or both halves of the palindrome sequence (lanes 14-12). In addition, mutations in the AT-rich region also had a deleterious effect on complex formation. Insertion of a single G at position -41 within the AT-rich region nearly abolished the shifted band, when the assay was performed at a total protein concentration of 1.4 mg/ml (1/4 x O_2 lysate, Fig. 2-7A, lanes 7-9). At higher protein concentrations, complex formation was

still observed, but at a lower yield than found with the wild-type sequence (Fig. 2-7B, compare lanes 7-9 to 10-12). Similar results were obtained with a truncated probe in which the entire AT-rich region was deleted. Finally, a probe containing the AT-rich region alone was not able to bind protein at either concentration (Figs. 2-7A and B, lanes 1-3). These in vitro results suggest that optimal complex formation results from a cooperative binding of protein(s) to both the palindrome and the AT-rich region, since far less complex was observed when either one of these elements was absent from the probe.

To further test the notion that a cooperative complex was formed, a binding competition experiment was performed (Fig. 2-8). Unlabeled wild-type DNA competed efficiently with wild-type labelled DNA such that, at a 50-fold molar excess, only very little complex was detected in a mobility-shift experiment (compare lanes 2-4, Fig 2-8). This control demonstrated that the specific DNA binding proteins in the cell extract were limiting at the 50-fold probe level in this experiment. As expected, an unlabelled DNA fragment with a mutation in the -35 region also effectively competed with the labeled probe (lane 5). However, DNA fragments that either lacked the AT-rich region (lane 7) or contained mutations in the palindrome sequence (lanes 9-11) did not effectively compete for binding, even at 50-fold molar excess over labelled probe. The competitions with a mutation in the only the left half of the palindrome decreased binding by about 50% (lane 9), indicating that the binding affinity of this altered sequence was decreased by a factor of at least 10 compared to the wild type

sequence. This was the best competing sequence of those with changes in either the palindrome or the AT-rich region. This result may mean that the right half of the palindrome is more critical to complex formation than the left half, despite the fact that the left half is an exact match with the consensus sequence, while the right half contains a mismatch. The AT-rich region alone (lane 8) or a DNA fragment which contained a G insertion in the AT-rich region (lane 6) also failed to inhibit specific complex formation. Finally, even when both the AT-rich region and the palindrome sequence were introduced on separate DNA fragments at 50-fold excess, the shifted wild-type band still formed (lane 12), demonstrating the importance of the two sequences acting in cis.

D. Determination of binding constants using a competition assay

Various amounts of unlabelled specific probes were added to the shift assay in order to determine minimum binding constants for both the crtEF 31mer and bchC wild type 200mer. This competition assay assumes that no dissociation occurs in the gel matrix itself, and therefore produces only minimum values for in vitro binding constants. If a single binding protein is assumed, the following equilibrium is used to determine a binary binding constant:

$$\text{DNA (D)} + \text{Protein (P)} \underset{}{\overset{K}{\rightleftharpoons}} \text{DNA-Protein Complex (C)}$$

in which K is the equilibrium constant, equal to [C] / [D][P].
Substituting the two mass balances: total DNA, D_O = [D] + [C] ; and total protein, P_O = [P] + [C] into the equilibrium, [C] = K [D] [P] :

$$[C] = K (D_O - [C]) (P_O - [C])$$

To avoid the error associated with determination of the efficiency of detecting the radioactive label, only the fraction of counts shifted, f = [C] / D_O, was calculated for each incubation and substituted into the above equation to yield:

$$f D_O = K D_O (1 - f) (P_O - f D_O) , \text{ or}$$

$$f / (1 - f) = K P_O - K (f D_O)$$

In a linear plot of f/(1-f) versus fD_O for various concentrations of probe DNA, the slope equals K, the binary equilibrium binding constant, and the x-intercept equals P_O, the concentration of binding protein in the lysate.

Using the consensus palindrome 31mer upstream of the crtEF operon as a probe, a minimum binding constant on the order of 10^7 M^{-1} (i.e. 8 X 10^6 M^{-1}) was obtained from such a plot. Using the DNA fragment containing the bchCXYZ operon control region as the probe, the binding of protein appeared to be approximately three orders of magnitude stronger with a binding constant of about 8 X 10^9 M^{-1} and a protein concentration of about 2-3 X 10^{-9} M (Fig. 2-9).

Alternatively, if two proteins are contained in the binding complex, the relationship becomes:

$$f / (1 - f) = - K (f D_O)^2 - K (A_O + B_O) (f D_O) + K A_O B_O$$

where A and B are the two proteins, and the best quadratic is fit to a plot of f/(1-f) versus fD_O. In this case, the tertiary binding constant for the bchCXYZ probe is estimated at 1.7 X 10^{18} M^{-1}, the limiting protein is still 2-3 X 10^9 M, and the other protein is approximately twice that concentration. Of course, the data are not

precise enough to discriminate between the one and two protein cases on the basis of this type of plot alone.

E. Kinetics of protein binding

The kinetics of DNA-protein complex formation and dissociation were examined using time-course mobility shift assays. Off rates for the complex were estimated using chase experiments in which a very large excess of unlabelled specific competitor DNA was added at different time points prior to loading the polyacrylamide gel (Fig. 2-10). The results indicated that the protein bound to the bchCXYZ fragment has a dissociation rate constant of about 0.02 min^{-1}. The complex association rate, although too fast to be measured directly, can be calculated from the equilibrium constant times the reverse rate constant as approximately 1-2 X 10^8 $M^{-1}min^{-1}$.

Also, it is interesting to note that an upper shift band also forms, but only upon longer incubation times, and that this band is more resistant to competition by specific competitor DNA (Fig. 2-10).

Figure 2-3. (**A**) Sequence of the "crtEF" 31mer double stranded DNA oligonucleotide probe. The conserved palindrome region of the sequence is highlighted in bold type. (**B**) 10% Native polyacylamide gel electrophoresis (PAGE) showing DNA-protein complex formation. Lane 1, [^{32}P] end-labelled crtEF 31mer palindrome probe only; Lane 2, γ-[^{32}P] ATP plus 10 µl protein extract from a R. capsulatus cell culture grown under photosynthetic conditions; Lane 3, 1 fmol labelled probe plus 10 µl protein extract; Lanes 4-8 each contain 10 fmol labelled probe and 10 µl protein extract; Lane 5, also includes 10x unlabelled crtEF 31mer probe; Lane 6, 100x unlabelled crtEF probe; Lane 7, 100x molar excess of a double strand 57mer oligonucleotide without the palindrome sequence; Lane 8, 100x molar excess of single strand non-palindromic 57mer DNA; Lanes 9-11 each contain 10 fmol probe and protein extract concentrated 3-fold by centricon; Lane 10 also includes 10x unlabelled crtEF probe; Lane 11, 100x unlabelled crtEF probe. All lanes contain 250x by weight poly(dIdC) non-specific competitor. Note: Apparent single strand bands arise from the short double stranded 31mer probe, even when the double stranded DNA is previously purified on a native gel.

A

5' GATCGG**GTGTAA**GTTTCAGT**TTACAC**AGATC
CTAGCC**CACATT**CAAAGTCA**AATGTG**TCTAG 5'

B

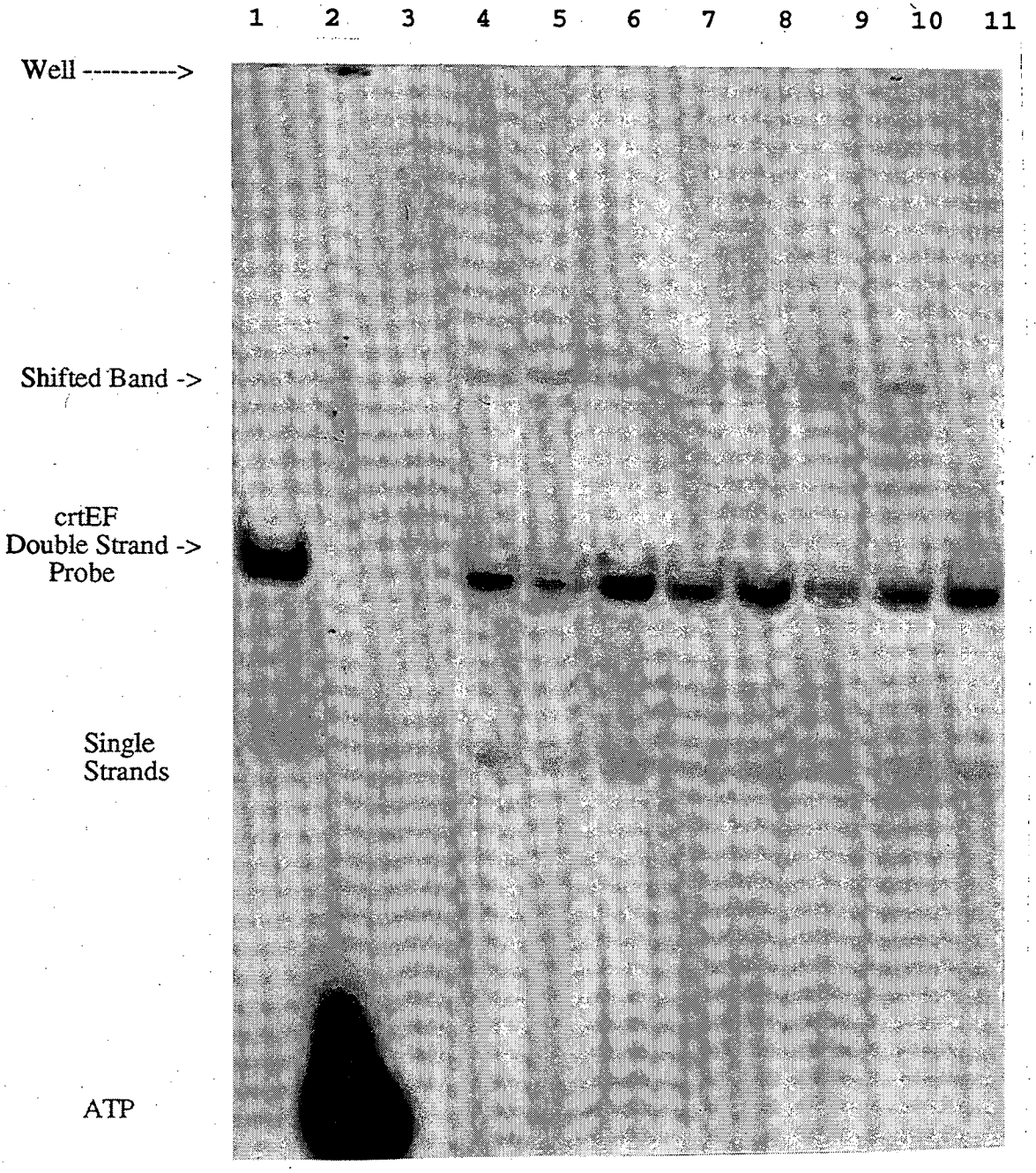
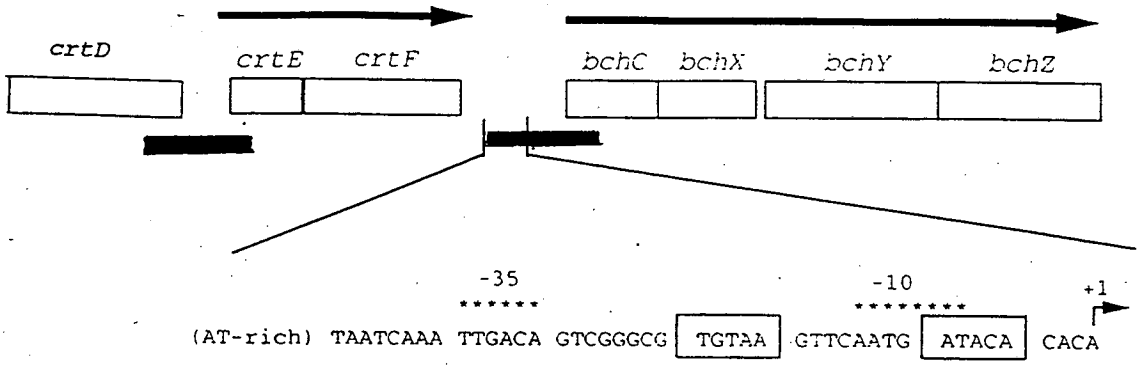


Figure 2-4. (**A**): Location of the 200mer probes upstream of the crtEF and bchCXYZ operons. (**B**): 4.5% native PAGE gel showing a comparison of mobility shift using crtEF and bchCXYZ 200mer probes: Lanes 1-3 contain 10 fmol [^{32}P] crtEF 200mer; Lane 1: 10 µl of aerobic lysate (about 3.9 mg/ml total protein); Lane 2, a negative control with no protein; Lane 3, 10 µl of anaerobic/ photosynthetic shifted lysate; Lanes 4-6 contain 10 fmol ^{32}P bchCXYZ 200mer and 10 µl anaerobic/ photosynthetic shifted lysate; Lane 4 also includes 100x unlabelled bchCXYZ 200mer as a competitor; Lane 6 includes 100x unlabelled crtEF 200mer as a competitor. All lanes contain 400x by weight poly(dIdC) as a non-specific DNA competitor.

A


crtD
crtE
crtF
bchC
bchX
bchY
bchZ
-35
-10
+1
(AT-rich) TAATCAAA TTGACA GTCGGGCG TGTAA GTTCAATG ATACA CACA

B

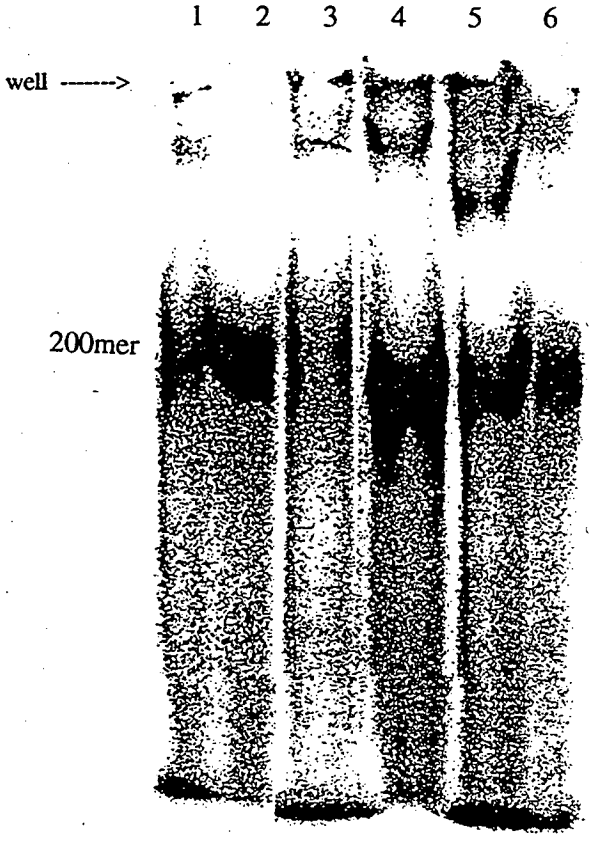

1 2 3 4 5 6
well ------>
200mer

Figure 2-5. In vitro mobility shift assay with cell lysates from aerobic and shifted, anaerobic cultures. Equal molar amounts of [^{32}P]-labelled wild type bchCXYZ 200mer DNA probe were titrated with increasing concentrations of cell lysate from a bacterial culture grown under aerobic conditions and then shifted to anaerobic conditions for 1 h. On the 4.5% native PAGE gel shown, 1x represents a total protein concentration of 5.6 mg/ml as measured in a Bradford assay. At the same total protein content, the aerobic cell lysate exhibited an approximately 2-fold higher binding affinity for the probe DNA. Lanes 1-5: Use of aerobic cell lysate at 1/16, 1/8, 1/4, 1/2, and 1x total protein concentration. Lanes 7-11: Use of anaerobic cell lysate at 1/16, 1/8, 1/4, 1/2, and 1x total protein concentration. Lanes 6 and 12: probe DNA with no cell lysate added. Note: At the highest protein concentrations, a considerable fraction of the probe was retained in the wells (e.g. lanes 4, 5, and 11).

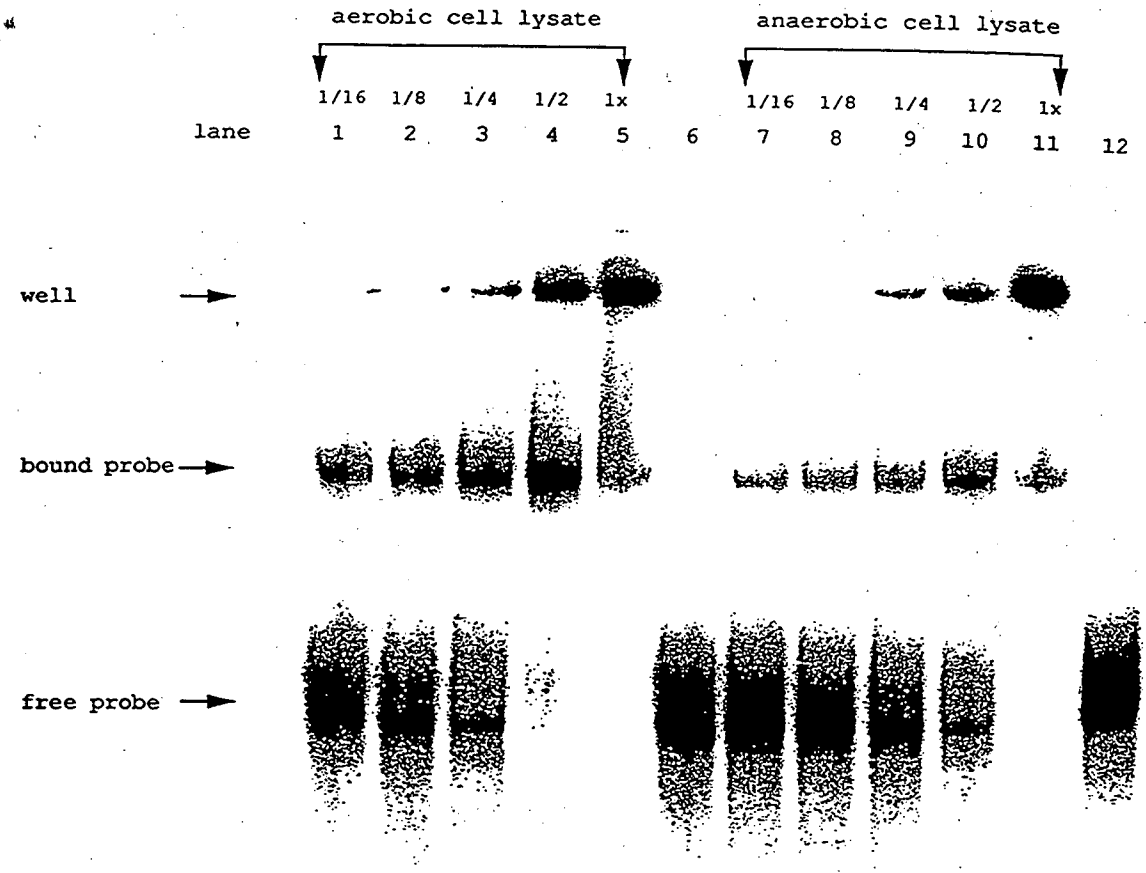
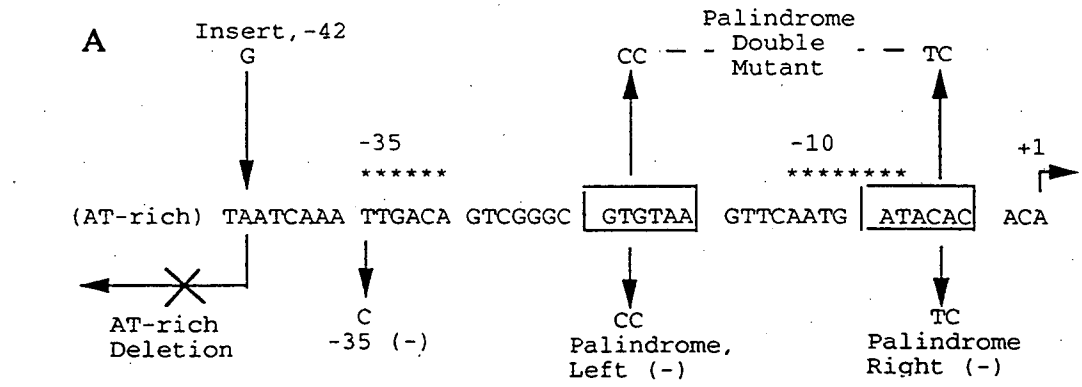

aerobic cell lysate
anaerobic cell lysate
1/16 1/8 1/4 1/2 1x
1/16 1/8 1/4 1/2 1x
lane
1 2 3 4 5 6 7 8 9 10 11 12
well
bound probe
free probe

Figure 2-6. (**A**): Schematic showing mutations in the bchCXYZ promoter region. (**B**): In vitro mobility shift assay using bchCXYZ DNA probes incorporating mutations in the σ^{70}-like -35 region or the palindrome region. Equal amounts of wild type or mutant [^{32}P]-labelled DNA probes were incubated with the same amount of cell lysate (1.4 mg total protein/ml), and loaded on a 4.5% native gel. Each probe was tested against an aerobic (+) or shifted anaerobic (-) lysate for complex formation. Lanes marked (0) contain only the probe DNA. Mutation of the -35 region of the promoter, did not affect complex formation.


A
Insert, -42
G
Palindrome
Double
Mutant
CC
TC
-35
******
-10
****
+1
(AT-rich) TAATCAAA TTGACA GTCGGGC GTGTAA GTTCAATG ATACAC ACA
AT-rich
Deletion
C
-35 (-)
CC
Palindrome,
Left (-)
TC
Palindrome
Right (-)

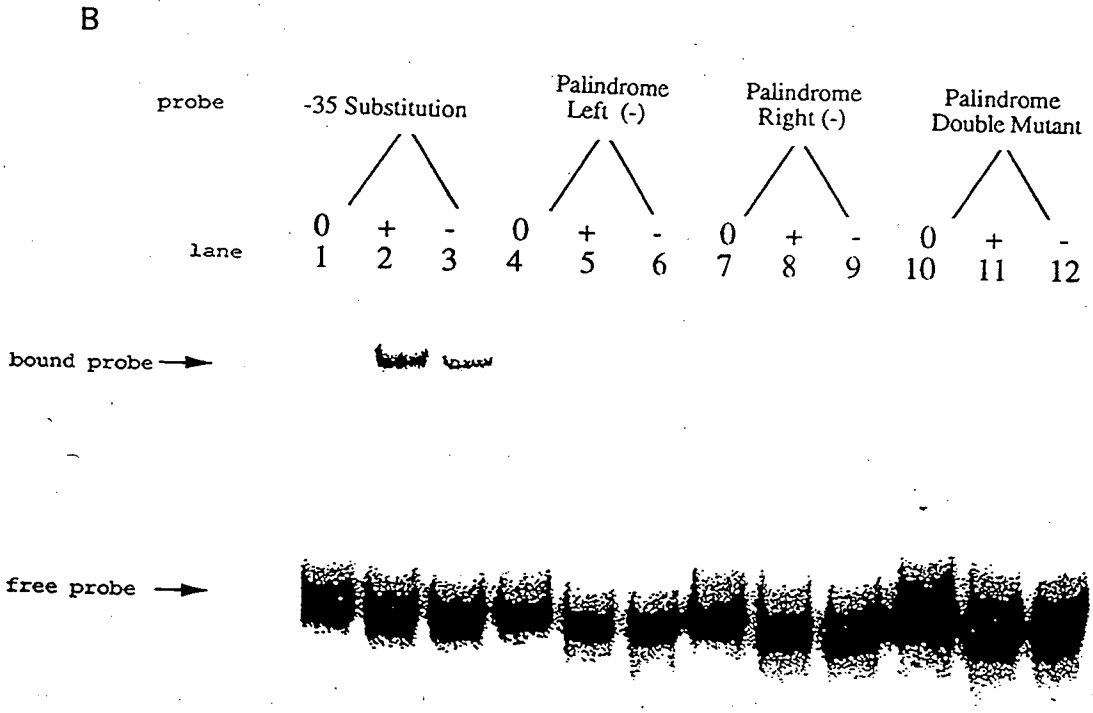

B
probe
-35 Substitution
Palindrome
Left (-)
Palindrome
Right (-)
Palindrome
Double Mutant
lane
0 + - 0 + - 0 + - 0 + -
1 2 3 4 5 6 7 8 9 10 11 12
bound probe
free probe

Figure 2-7. In vitro mobility shift assay using DNA probes derived from various deletion and insertion mutant bchCXYZ promoter fragments. Equal amounts of wild type or mutant [^{32}P]-labelled DNA probes were incubated with the same amount of cell lysate and assayed on a 4.5% native PAGE gel. Each probe was tested against an aerobic (+) or shifted anaerobic (-) lysate for complex formation. Lanes marked (0) contain only the probe DNA. The migration of free probe varied due to the length of the DNA. (**A**): 1.4 mg total protein/ml in the binding assay; (**B**): 5.6 mg total protein/ml in the binding assays.

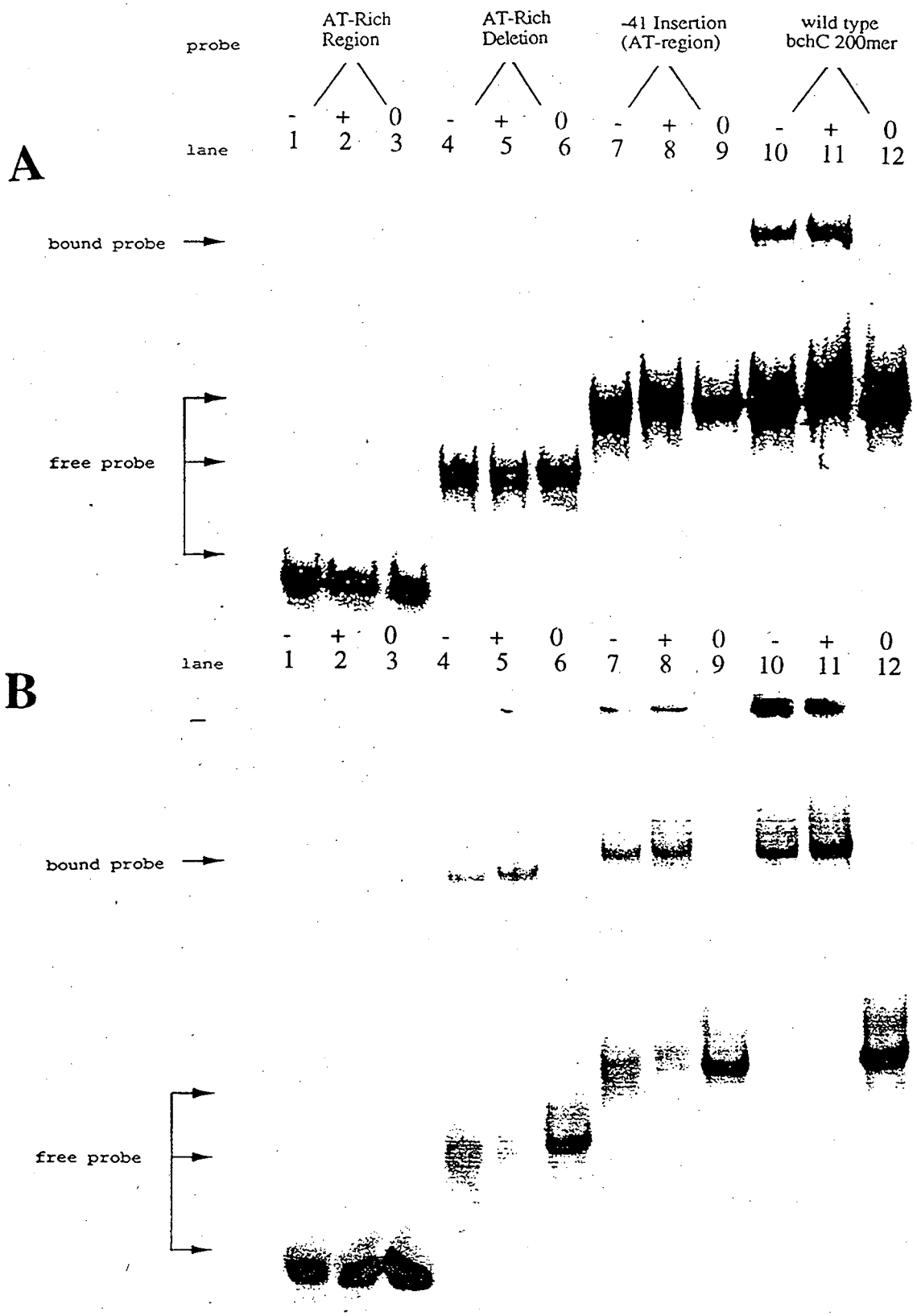

A
probe
AT-Rich Region
AT-Rich Deletion
-41 Insertion (AT-region)
wild type bchC 200mer
lane
- + 0 - + 0 - + 0 - + 0
1 2 3 4 5 6 7 8 9 10 11 12
bound probe
free probe
B
lane
- + 0 - + 0 - + 0 - + 0
1 2 3 4 5 6 7 8 9 10 11 12
bound probe
free probe

Figure 2-8. Competition binding assays between labelled wild type bchCXYZ probe and unlabelled competitor DNA from the various mutant constructs on a 4.5% native PAGE gel. Lane 1, free probe. Lane 2, complex formed in the absence of specific competitor DNA. A 50-fold molar excess of unlabelled DNA was added to an aerobic cell lysate to compete with labelled wild type DNA probe in the binding of protein. Only DNA fragments with the wild type sequence (lanes 3 and 4) or the -35 mutation (lane 5), both of which contained an intact AT-rich region and palindrome sequence, efficiently competed for binding. Insertion of a single G in the AT-rich region (lane 6) resulted in loss of ability to compete for complex formation. Neither the palindrome sequence (lane 7) nor the AT-rich region alone (lane 8) could successfully compete for binding. Site specific mutation of either the left or right half of the palindrome (lanes 9 or 10, respectively) or simultaneous disruption of both halves (lane 11) destroyed the ability of the fragment to compete for complex formation. Finally, even when the AT-rich region and the palindrome were present in the incubation mixture on separate fragments (lane 12), binding to the labelled probe was not inhibited.

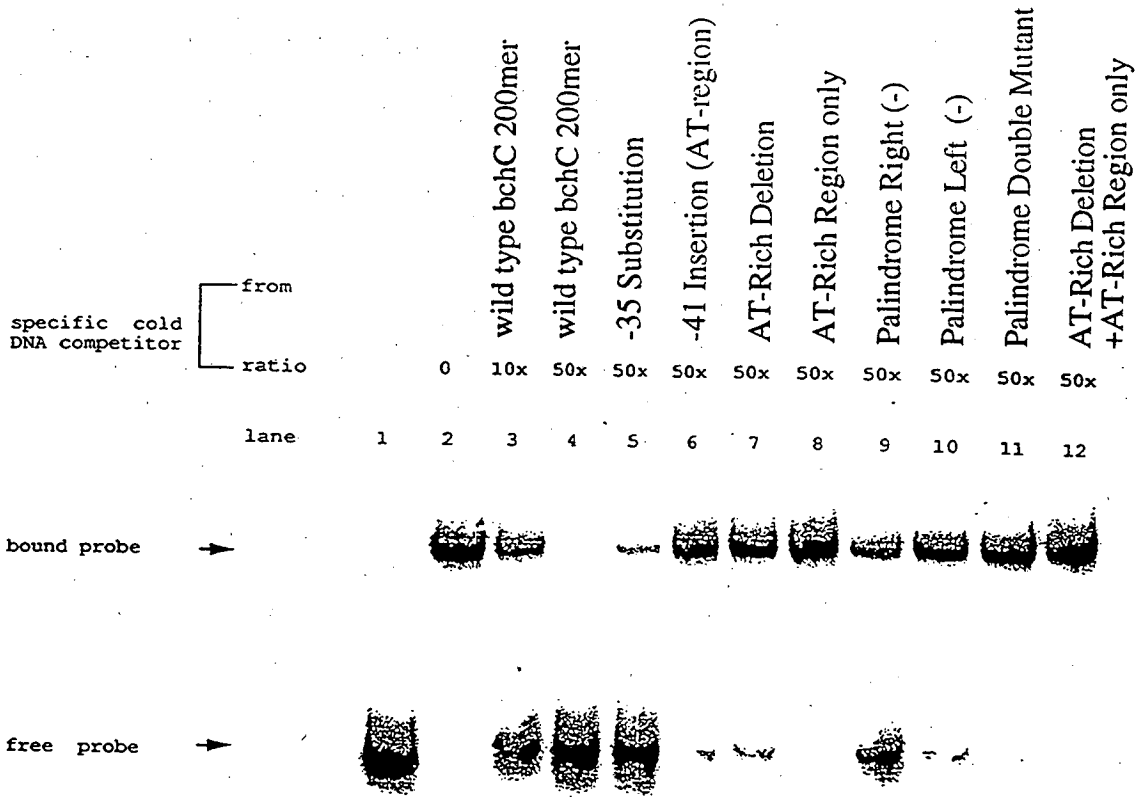
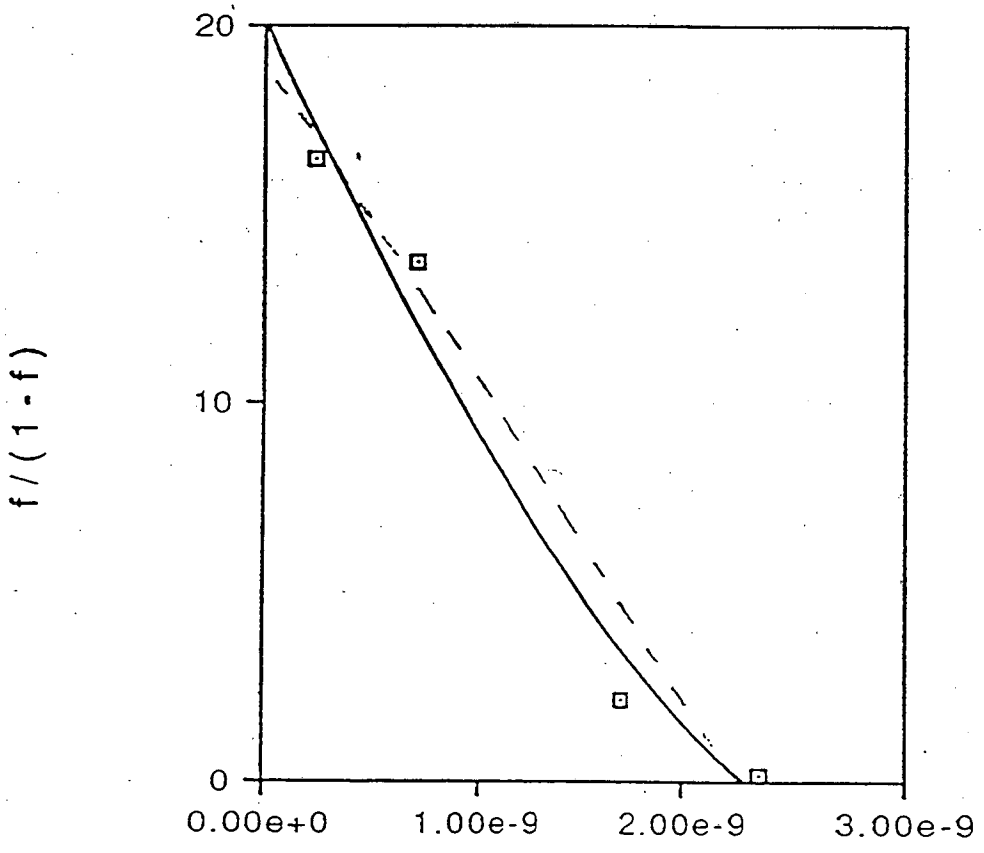
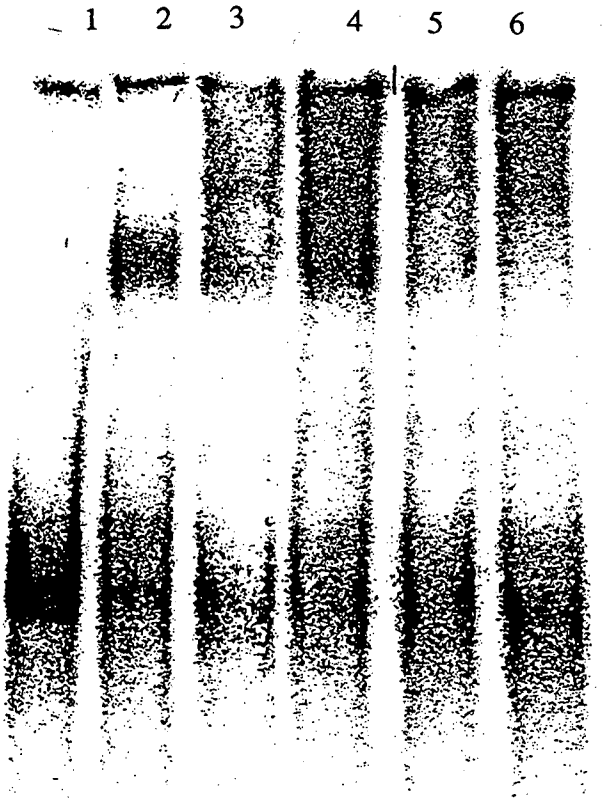

specific cold DNA competitor
from
ratio
wild type bchC 200mer
wild type bchC 200mer
-35 Substitution
-41 Insertion (AT-region)
AT-Rich Deletion
AT-Rich Region only
Palindrome Right (-)
Palindrome Left (-)
Palindrome Double Mutant
AT-Rich Deletion +AT-Rich Region only
0 10x 50x 50x 50x 50x 50x 50x 50x 50x 50x
lane 1 2 3 4 5 6 7 8 9 10 11 12
bound probe
free probe

Figure 2-9. Binding plot of f/(1-f) vs. fD_o for the bchCXYZ wild type 200mer probe and lysate derived from aerobically grown cells. "f" is the fraction of probe bound in the DNA complex (shifted band). "D_o" is the total concentration of DNA in the assay (bound plus free DNA). The slope of the linear (dashed line) fit to the data is equal to -K, the inverse of the binary binding constant. The solid line represents a second order polynomial fit, which is described by the equation below the graph. The tertiary (two proteins plus DNA probe) binding constant equals the inverse of the coefficient of the x^2 term.



$y = 20.137 - 1.2595e+10x + 1.6588e+18x^2 \quad R^2 = 0.974$

Figure 2-10. A pulse/chase gel assay measuring DNA-protein complex binding kinetics using the bchCXYZ wild type 200mer probe and a lysate derived from aerobically grown cells. On a 4.5% native PAGE gel: Lane 1, probe alone; Lane 2, a 10 second incubation prior to loading the gel; Lane 3, a 35 min incubation; Lanes 4-6, total incubation times of 95 min; Lane 5, 75 min plus a 20 min chase with 100x cold probe; Lane 6, 15 min plus a 80 min chase.


1
2
3
4
5
6

IV. Discussion

A. The palindrome-binding protein appears to act as a repressor of the bchCXYZ operon in R. capsulatus.

The binding of protein to the bchCXYZ DNA fragment in the mobility shift assay is approximately 2-3 times as strong when an aerobic cell lysate is used compared to the corresponding culture shifted to anaerobic/ photosynthetic growth conditions. Coupled with the result that such a shift to anaerobic growth conditions results in an increase in transcription of the bchCXYZ operon (Armstrong, 1989b), the data indicate that the palindrome sequence binds a repressor of aerobic transcription.

Second, the placement of the palindrome relative to the σ^{70}-like promoter is much more consistent with that of a repressor. Nearly all activator recognition sites associated with σ^{70} promoters in E. coli, with very few exceptions, are located upstream of the -35 region, and none has been reported in a position near the -10 region (reviewed in Colladovides et al., 1991). However, this palindrome sequence straddles both sides of the -10 region, an arrangement commonly occuring with prokaryotic repressors (Colladovides et al., 1991). In this way, binding of a transcription factor to the palindrome could presumably interfere with utilization of the -10 region by RNA polymerase.

Finally, the proposed role of the binding protein as a repressor under aerobic conditions is consistent with the results of an in vivo study (Ma et al., 1992) in which lacZ transcriptional

fusions were constructed in a plasmid compatible with R. capsulatus, and used to measure the effects of mutations in the bchCXYZ promoter region. Although mutations in the right, left, or both halves of the palindrome produced different basal levels of β-galactosidase expression, the differential effect of shifting to anaerobic growth was markedly reduced in each case in which the palindrome was disrupted. Since this study measured only transcription from the bchC promoter region, it did not include the effect of any read-through transcription from the upstream crtEF operon. Further mutational analysis demonstrated that basal transcription could be increased or entirely abolished by appropriate changes in the σ^{70}-like sequence, indicating that this sequence is the only true functional promoter in this region, although the palindrome does alter the efficiency of this promoter (Ma et al., 1992).

B. Evidence for possible multimer binding at the conserved palindrome

In many instances the formation of a second, higher band of lower mobility is observed in the gel-shift assay. This occurs with both the crtEF 31mer and the bchCXYZ restriction length fragment probes at approximately 60% of the mobility of the original shifted band (Figs. 2-3B, 2-4B, 2-10). Although this upper band is generally of lesser intensity than the primary shifted band with both of the above probes, the upper band is the only shifted band visible when the 200mer crtEF probe is used. Also, the formation of this upper band is kinetically slower than the lower shifted band (Fig. 2-10,

lanes 4, 5, and 6).

This upper protein-DNA band is usually present to a relatively greater extent in older lysates, those which have undergone a greater degree of manipulation (precipitation and dialysis, purification steps, multiple freeze/ thaw cycles, etc.), and when longer incubation times are used in the shift assay. This might indicate that some degree of protein degradation is required for the formation of this upper band, a somewhat unusual occurrence since degraded binding proteins nearly always result in lower shifted bands, more mobile than that of the complex formed with the intact protein. In this case, the presence of a second factor or a dimer of an altered form of the original protein would explain the upward shift. It is very unlikely that the upper band is just the result of a conformational change, since it occurs with the small 31mer as well as with longer probes.

In addition, a single lysate sample or partly purified fraction (Chapter 3) can produce larger amounts of the upper band by the use of higher concentrations of glycerol, increasing the amount of partly purified protein, or even adding larger amounts of specific probe. It is likely that the added probe is titrating the protein in the lower band, permitting a weaker interaction to produce the upper band.

Whether this apparent multiple protein binding mode is just the result of <u>in vitro</u> experiments, or if there is a parallel phenomenon <u>in vivo</u> with regulatory implications, is not known at present.

C. Toward a general model for transcription regulation of pigment biosynthesis in R. capsulatus.

The in vitro binding experiments in this chapter identify the palindrome found upstream of various pigment biosynthesis operons as a sequence-specific protein binding site. Furthermore, they suggest that the binding protein functions as a repressor of transcription under aerobic growth conditions, and demonstrate that the presence of an upstream AT-rich region greatly enhances binding to this palindrome. These results are confirmed by in vivo experiments with lacZ fusion constructs, which also identified a σ^{70}-like sequence as the funcional promoter for the bchCXYZ operon (Ma et al.,1982)

Recently, the sequencing of the entire 46 kb photosynthesis gene cluster of R. capsulatus has been completed by M. Alberti and J. Hearst (EMBL Data Library accession number Z11165). Analysis of these data demonstrates that σ^{70}-like sequences are located upstream of at least three other pigment biosynthesis operons (Fig 2-2B), that the palindrome sequence is also found upstream of many pigment biosynthesis operons (Fig 2-2A), and finally that AT-rich regions are located in all of the intergenic areas of the cluster (Armstrong et al., 1989). The anaerobic induction of the genes in the carotenoid portion of the cluster was studied using dot-blot hybridization to mRNA's by Armstrong (1989b), who classifed each operon as strongly induced (5-10 fold), induced (2-5 fold), or not induced by the removal of oxygen (summarized in Fig. 2-1).

It appears that the σ^{70}-like sequences function as promoters

and confer constitutive expression on these genes. The relatively high basal expression from these promoters could facilitate adaptation to anaerobic environments through the superoperonal organization of the photosynthesis cluster. By themselves, the σ^{70}-like sequences seem to lack the information to confer regulation of transcription (Ma et al., 1992). This view is supported by two additional observations in relation to crt gene expression. Two of the putative σ^{70}-like promoters are located upstream of the crtD, and crtI genes (Fig. 2-2B). Like the bchC region, a conserved palindrome sequence is also found overlapping the -10 region of crtD (Fig. 2-2B), a gene whose transcription is also induced by a shift to anaerobic conditions (Fig. 2-1). In contrast, only a very poor match to the palindrome sequence could be found near the σ^{70}-like promoter upstream of crtI (Fig.2-2B) and it is closer to the -35 region than to the -10. Transcription of crtI is constitutive (Fig. 2-1). Thus, when the σ^{70}-like sequence functions alone, it appears that it produces only constitutive gene expression.

In at least two species of purple photosynthetic bacteria, R. capsulatus and R. sphaeroides, (Armstrong et al., 1989; Lee and Kaplan, 1992), the same palindrome has also been reported upstream of the puc operon, which encodes the structural polypeptides of the LHII complex . This raises the possibility that the palindrome binding protein coordinates the transcriptional regulation between LHII, the major pigment-protein complex of the photosynthetic apparatus, and the pigment biosynthesis enzymes. In contrast, this palindrome does not occur in front of the puf or puh operons, a finding which is consistent with the idea that

pigment biosynthesis is regulated by a different transcriptional mechanism from the one for these structural polypeptide operons. The AT-rich region works together with the palindrome to regulate expression of bchCXYZ. Since AT-rich regions are consistently observed in the intergenic stretches of the photosynthesic gene cluster, they may play a general role in regulating transcription, perhaps via local melting to accommodate the binding of additional factors, facilitating the bending of the DNA, or by presenting a minor groove of distinct width (Nelson et al., 1987).

Finally, the third class of pigment biosynthesis operons is the highly induced group which includes crtEF and crtA. No σ^{70}-like sequences were evident in the promoter regions of these operons, but both contain AT-rich areas, as well as palindrome sequences which are matches with the consensus sequence. In the binding experiments using crtEF DNA probes, complex formation was much weaker than with bchCXYZ, despite the superior match with the palindrome consensus. The observation that the crtEF 200mer only forms the upper shifted band may mean that dimeric binding of protein could be part of a method of transcriptional activation. Also, while the in vivo results of Ma et al. (1992) showed that a mutation in the left half of the palindrome produced higher (aerobic-like) levels of transcription consistent with a repressor site, a mutation in the right half resulted in reduced transcription versus the wild type, indicating a possible activator role for this site. Alternatively, proximity to the -10 region has been postulated to account for this reduction (Ma et al., 1992; Fig. 2-6A). Also, the left "repressor" half of the palindrome is better conserved than the right in the other

copies of palindrome found throughout the gene cluster. It is possible that the binding of the dimeric form of the protein to the perfect match of the palindrome is weaker because it is transitory and leads to other events, such as activation of RNA polymerase, either directly or through a cascade of other factors, forming structures which are not stable in the gel assay. On the other hand, repressor binding should be strong in order to compete effectively with the polymerase. Oxygen tension could directly or indirectly affect the mode of complex formation.

The presence of imperfect (e.g. bchC) and even apparently vestigial (e.g. crtI) versions of the palindrome near the putative σ^{70}-like promoters might mean that the palindrome-binding activator/ repressor mechanism was the original promoter for these genes. In that case, constitutive expression from a σ^{70}-like promoter would be a more recent adaptation, providing an advantage in the shift to aerobic growth conditions through the superoperonal organization of the photosynthetic genes.

Chapter 3:

Isolation of a Sequence-Specific DNA-Binding Factor which affects Transcription of Pigment Biosynthesis Genes in R. capsulatus

I. Introduction

An R. capsulatus DNA binding factor was identified and characterized in Chapter 2. This protein was shown to bind in a sequence specific manner to a DNA palindrome found upstream of carotenoid and bacteriochlorophyll biosynthesis operons, using competition assays, and mutations in the binding site. The presence of an AT-rich region further upstream of the palindrome sequence was also demonstrated to enhance binding in a cooperative manner. Finally, the disruption of this binding site was independently shown to affect transcription in vivo (Ma et al., 1992).

Unlike many of the more extensively studied transcription factors, no overproducing strains have been identified for this pigment biosynthesis regulatory protein. The experiments in this chapter describe the isolation of this protein of apparent molecular mass of 95 kDa. N-terminal sequence data are also presented.

II. Materials and Methods

Bacterial strains and growth conditions

The R. capsulatus strain used in these experiments was B100 (Marrs, 1981). These R. capsulatus cells were grown under aerobic conditions until a cell density of about 1.7 X 10^8 per ml was reached prior to harvesting or shifting to photosynthetic growth conditions. Photosynthetic cultures were then shifted to anaerobic conditions in the presence of 500 W/m^2 light provided by a bank of lumiline lamps for 60 minutes prior to harvesting. On the protein preparative scale, dark, aerobic cultures were grown in 1 liter covered glass vessels which were sparged with a mixture of $N_2:O_2:CO_2$ (80:20:2). To shift to anaerobic/ photosynthetic conditions, cultures were uncovered and sparged with a mixture of $N_2:CO_2$ (80:2). Gas flow and composition were controlled using a Matheson Gas Products Multiple Dyna-blender, Model 8219. Growth rates were monitored on a Bausch and Lomb Spectronic 21 spectrophotometer.

E. coli DH5α cells were grown aerobically at 37°C on plates or on a rotary shaker according to standard protocols (Sambrook et al., 1989)

Recovery of protein

R. capsulatus cells were harvested by first cooling the vessel in an ice-water bath for 15 min with continuous sparging with the gas mixture used for cell growth, followed by centrifugation for 15 min at 6000 x g at 4°C. All further manipulations were performed at

4°C. Each culture was resuspended in 1% of its original volume in a lysis buffer consisting of 50 mM Tris-HCl, pH 7.5, 25 mM KCl, 5mM EDTA, 20% (v/v) glycerol, 1 mM dithiothreitol, and 0.5 mM phenylmethylsulfonyl fluoride (PMSF) (Sigma). The 0.1% Surfact-amp non-ionic surfactant (Pierce) in the original version of the buffer was removed after it was discovered that this anionic surfactant (alkylphenol ethoxylate) contributed to a loss of DNA-binding activity as the protein is purified. Cell lysates were obtained by one passage through a French pressure cell at 13,000 psi. Large cell debris and unbroken cells were removed by centrifugation at 4500 x g for 10 min.

Initially, proteins were precipitated by adding finely ground ammonium sulfate (Bio-Rad, enzyme grade) to 65% of saturation, and the pellets spun down at 15,000 x g for 15 min and stored at 4°C. For further use, the pellets were resuspended in the original Tris lysis buffer and dialyzed over a period of 5h against 4 changes of a buffer consisting of 10 mM Tris-HCl, pH 7.5, 50 mM KCl, 2.5 mM EDTA, 5-10% (v/v) glycerol, 1 mM dithiothreitol, 0.5 mM PMSF, and 0.1% Surfact-amp (Method A). Alternatively (Method B), the original lysates themselves were just divided into aliquots, frozen in liquid N_2, and stored at -70°C for further use. In either case, protein concentrations were determined using a Bradford assay (Bradford, 1976).

Heparin-Agarose column

DNA-binding proteins in the cell lysates were fractionated using a gravity column of heparin crosslinked to agarose beads

(Sigma, Type II). The column bed volume used was roughly equivalent to the volume of crude lysate. The heparin column was equilibrated with 8-10 column volumes of a buffer consisting of 10 mM Tris-HCl, pH 7.5, 50 mM KCl, 2.5 mM EDTA, 5% (v/v) glycerol, 1 mM dithiothreitol, 0.5 mM PMSF, and 0.1% Surfact-amp (column buffer). The dialyzed ammonium sulfate fractions were spun at 6000 x g for 15 min to remove any undissolved material and loaded directly onto the column. Alternatively, the crude lysates, which had been frozen directly (Method B), were thawed on ice, diluted by a factor of four to match the composition of the column buffer, and spun twice for 15 min at 12,000 x g to remove insoluble material prior to application to the heparin column. In either case, the column was then washed with 4-5 volumes of the above column buffer, and subsequently eluted stepwise with the same column buffer containing 225, 425, 700, and 1000 mM KCl, respectively. Individual column fractions were assayed for sequence specific binding activity using a gel mobility shift, as described in Chapter 2.

DNA-Affinity columns

Sequence specific DNA-binding columns were prepared according to the method of Kadonaga and Tjian (1986). Concatemers of the crtEF palindrome DNA (perfect consensus sequence) were prepared by kinasing and ligating complementary synthetic 27mers which contained compatible protruding ends (specifically a BamHI site). The yield and size of the concatemers were determined on a 2% agarose gel, prior to coupling 800 µg of the DNA to 5.0 g of cyanogen bromide activated Sepharose CL-2B beads (Pharmacia)

(Kadonaga and Tjian, 1986). In a similar manner, concatemers were prepared from a 200mer containing the whole transcription control region of the downstream bchCXYZ operon (including the palindrome sequence). The approximately 150 µg of DNA used was obtained from a 2.4 L plasmid preparation (Sambrook et al., 1989) of a pUC-18 construct containing blunt-end ligated bchCXYZ 200mer probe from Chapter 2. The construct (Fig. 3-1) was cloned in E. coli DH5α and screened on plates containing X-gal (Sambrook et al., 1989). The insert DNA was recovered from a 4.5% native polyacrylamide gel following restriction with AvaI and BstEII (sites engineered into the original PCR primers). These back-to-back concatemers were coupled to 1.5 g of Sepharose as described above.

The DNA-affinity columns were equilibrated with a buffer consisting of 25 mM Tris-HCl, pH 7.5, 75 mM KCl, 1 mM EDTA, 20% (v/v) glycerol, 1 mM dithiothreitol, 0.1 mM PMSF, and 0.1% Surfact-amp. Active binding fractions from the heparin-agarose column were pooled and diluted by a factor of 3 to match the composition of the buffer used with this column. Before application to the column, non-specific competitor DNA, poly(dIdC) was added at 25µg/ml of sample. The column was washed with 8 volumes of the buffer, and eluted stepwise with buffer containing 200, 400, and 1000 mM KCl.

Mobility shift assay and prep

For analytical scale mobility shift assays, 2 µl of [^{32}P]-labelled DNA probe (approximately 10 fmol) was added to a mixture of 8 µl of non-specific competitor DNA and 10 µl of cell lysate or protein

fraction. Unless stated otherwise, the competitor DNA sample contained 400-fold excess (by weight) of poly(dI·dC). The probes used in the assay were the end-labelled bchCXYZ 200mer and crtEF 31mer described in Chapter 2. For competition assays, as well as the preparative scale mobility shift, a 5 to 200-fold molar excess of unlabelled specific competitor DNA was also included to increase the yield of protein bound in the shifted band (although the fraction of bound DNA is reduced). In the case of the preparative scale band shift, the source of the bchCXYZ competitor was the insert from the pUC plasmid prep described above used to construct the DNA-affinity column. After incubation for 20 min at room temperature, samples were loaded on a native polyacrylamide gel, run for approximately 4 h at 8 V/cm, and then visualized by autoradiography or exposure on a Molecular Dynamics Phosphor-Imager.

SDS gel electrophoresis

Proteins were analyzed by an SDS- 10% PAGE gel using the discontinous buffer system of Laemmli (1970). Intact E. coli cells were centrifuged and resuspended in 400 µl 1X sample buffer per ml. Liquid samples were boiled for 2 min and quenched on ice. Proteins were separated on a 14 cm gel apparatus, and visualized by the silver staining protocol of Morrissey (1981).

Gel slices containing the shifted bands were excised from preparative scale native mobility shift gels and incubated in the denaturing sample buffer for SDS gels for 30 min at room temperature according to the protocol of Cleveland et al. (1977) and

stored at -70°C. Upon thawing, these gel slices were inserted into the wells of a Laemmli 10% SDS-PAGE protein modified to include a longer (4 cm) stacking gel (Cleveland et al., 1977), overlaid with the sample buffer, and run in the same manner as described for liquid samples.

To obtain protein for N-terminal sequencing, an unstained gel was electroblotted overnight onto an Pro-Blott™ membrane, according to the manufacturer's protocol (Applied Biosystems, Foster City, CA). The membrane was stained with Coomassie Blue, and the band of interest was excised from the membrane and sequenced.

III. Results

A. Heparin-Agarose column fractionation, gel shift assay

The binding activity of the protein appeared to be only partly recovered after ammonium sulfate precipitation and dialysis. The original level of activity was restored upon fractionation of the dialyzed lysate with the heparin column. Using either the short crtEF oligonucleotide or longer bchC region probe, the binding activity was recovered in column fractions eluted with 200-250 mM KCl in the buffer. In cases when an upper shifted band was also observed in the original lysate (Chapter 2), it could be eluted from the heparin-agarose column with 400 mM KCl. However, this upper band can also arise later from the 200 mM fraction upon extensive handling or long incubation times. It was further noted that the ammonium sulfate precipitation and dialysis steps could be eliminated prior to use of the heparin column without any deleterious effects on fractionation or yield of specific binding activity, as long as the cell debris was carefully removed by spinning at 12,000 x g before loading the crude material on the column (see Materials and Methods section).

Following fractionation on the heparin column, a rapid loss of bchC binding activity was observed in the active fractions. This loss was less pronounced with heparin fractions derived from photosynthetic lysates when compared with those derived from highly oxygenated cultures. A study of variables indicated that this loss of binding activity was due to the presence of the surfactant (a "Triton-type" nonionic, octylphenol-polyethoxylate) in the

purification process. This loss of activity in column fractions was only stopped when the surfactant was removed from the lysis buffer as well as the column wash and elution buffers. After heparin column conditions were determined on the analytical scale, a preparative scale column was run using 8 ml of lysate (from a 1 L aerobic culture) (Fig. 3-2).

B. DNA-Affinity columns

Use of the poly(crtEF palindrome) DNA-sepharose column only enriched the binding activity by a small amount. Furthermore, a high level of glycerol (45-50%) was required to effect protein binding to the column. As a result, binding activity was largely shifted to the upper band described in Chapter 2. Unfortunately, multiple passes through this column did not result in isolation of a specific protein (Fig. 3-3, lane 4). The poly(bchC 200mer) column produced a loss of activity, not merely a failure to bind. This did not appear to be a case of the activity remaining on the column, since elution with salt levels as high as 2.5 M did not result in recovery of binding activity.

C. Preparative band-shift

An alternate second fractionation step is a preparative version of the band shift assay, itself. Active fractions from the heparin-agarose column were run on the mobility shift assay with labelled bchC DNA probe as well as specific and non-specific competitor DNA (see Materials and Methods section). Following analysis of the native shift gel by phosphorimaging, the shifted

band containing DNA-protein complex was excised from the gel. To identify the shifted protein, a negative control lane was also run containing the same active heparin column eluate and polydIdC non-specific competitor, but no bchC probe. An unlabelled gel slice having the same mobility as the DNA-protein complex was also excised from this lane. Comparison of the excised gel slices with and without the DNA probe on a subsequent SDS-PAGE gel resulted in the identification of 95 kDa protein, present only when the probe was used, as the likely DNA-binding factor (Fig. 3-3).

Following this small scale experiment, active heparin-agarose column fractions (from aerobic growth of cells) were pooled, concentrated from 12 to 6 ml on a Centricon-10 spin column at 4°C, and run on a 4.5% native polyacrylamide gel with the bchC 201mer probe (with 25x specific competitor and 400x polydIdC).

D. Identification of a binding protein

Following the preparative gel mobility shift, the excised gel strip was denatured, and transferred to an SDS-PAGE protein gel which was subsequently electroblotted onto a membrane (see Materials and Methods section). The 95 kDa protein band was removed and the N-terminal amino acid sequence determined as: -DIIDPT(D)PS(G)-(T)-(K), in which dashes represent undetermined residues, and parentheses indicate residues determined with some degree of ambiguity. This sequence is too short for any direct homology comparisons with other factors (although a search of protein data bases was performed). However, the sequence should be long enough to permit the construction of a DNA oligonucleotide

probe for use in locating the corresponding gene.

Figure 3-1. pUC-derived plasmid used to overproduce the bchCXYZ 200mer probe

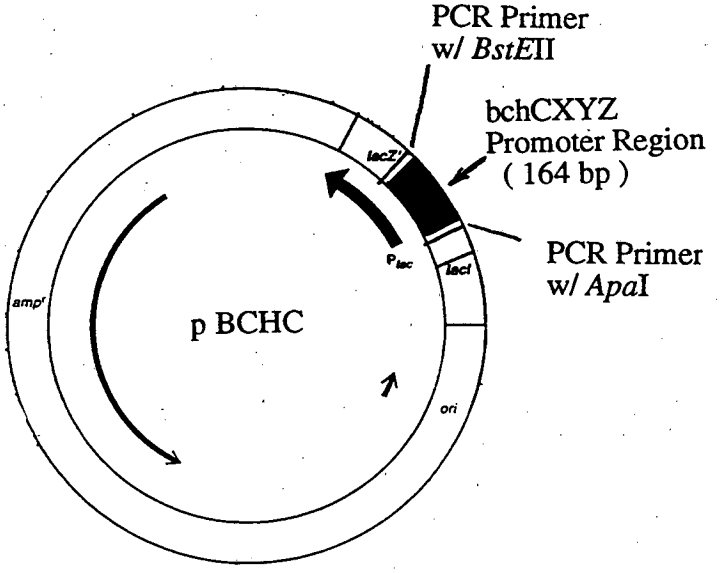

PCR Primer
w/ BstEII
bchCXYZ
Promoter Region
(164 bp)
PCR Primer
w/ ApaI
lacZ'
P lac
lacI
amp'
p BCHC
ori

Figure 3-2: Mobility shift assay of Heparin-Agarose column fractionation, preparative scale. Note: more narrow fractions (smaller volumes) were taken around expected active eluates. Numbers above the lane numbers indicate the volume of the assayed fraction (in ml).

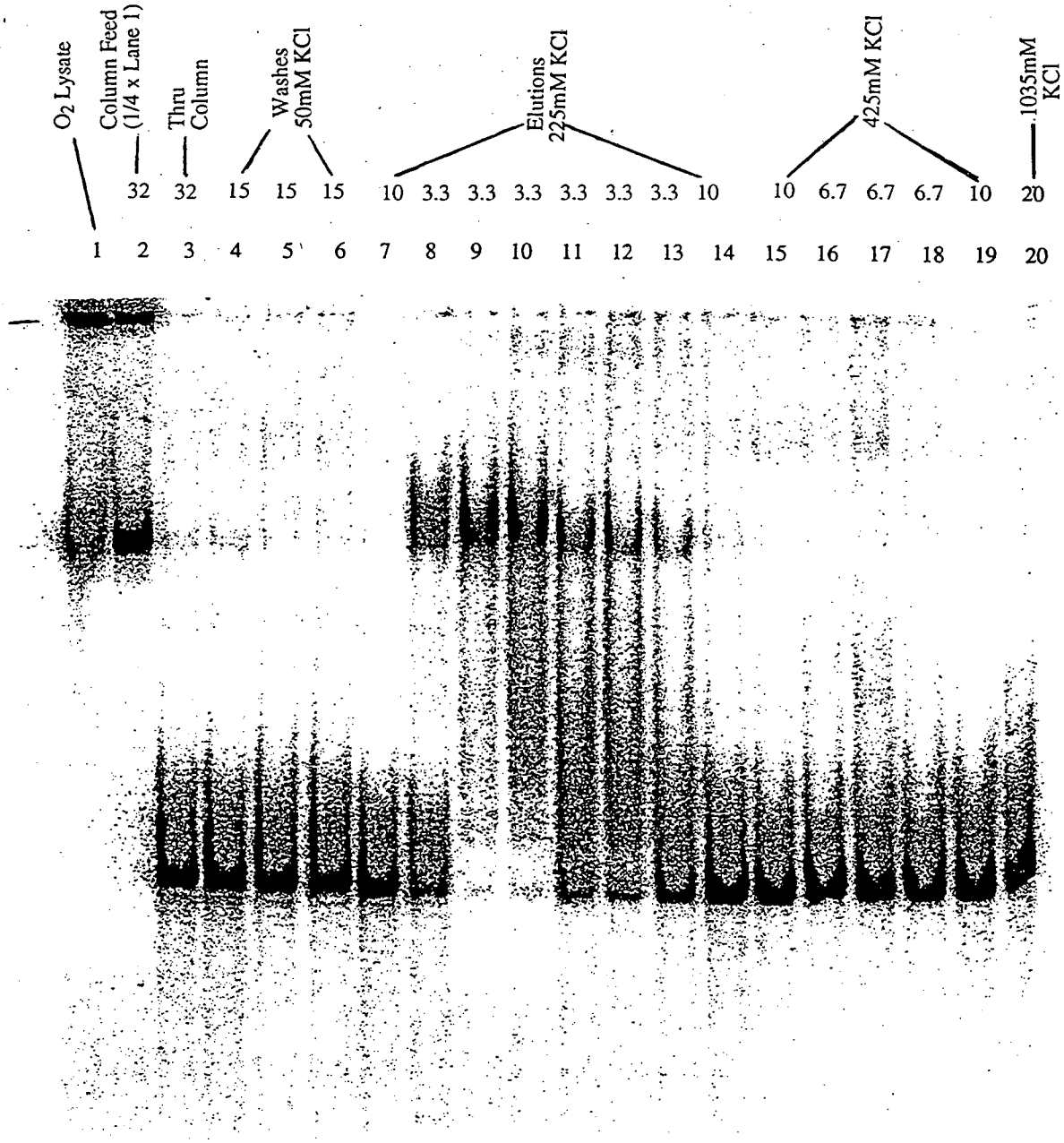
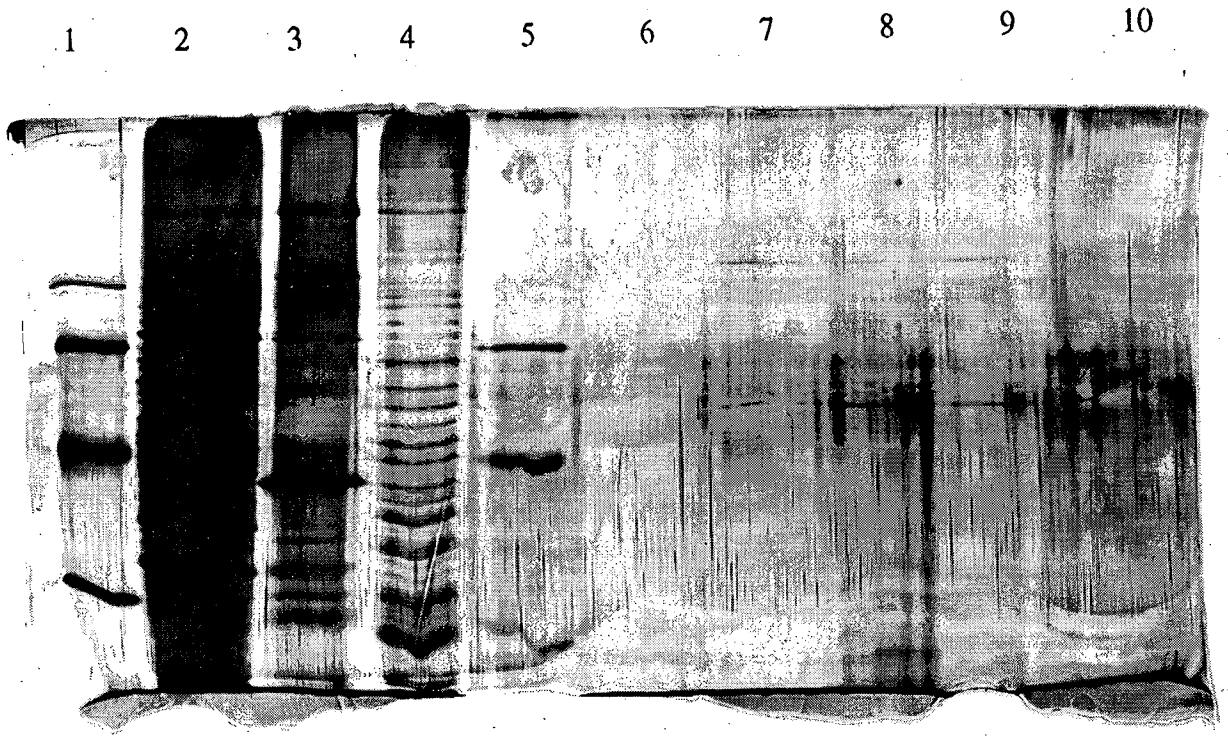

O2 Lysate
Column Feed (1/4 x Lane 1)
Thru Column
Washes 50mM KCl
Elutions 225mM KCl
425mM KCl
1035mM KCl
32 32 15 15 15 10 3.3 3.3 3.3 3.3 3.3 3.3 10 10 6.7 6.7 6.7 10 20
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

Figure 3-3: SDS-10%-PAGE, Lane 1, 10µl molecular mass standards: 20, 43, 67, 90 kDa.; Lane 2, 0.5 µl aerobic lysate; Lane 3, 5 µl active heparin column eluate; Lane 4, 5 µl active eluate from poly(crtEF) DNA-affinity column; Lane 5, 2 µl molecular mass standards; Lane 6, 25 fmol bchCXYZ 200mer DNA alone; Lanes 7-10, bands excised from mobility shift gel of active heparin-agarose eluate; Lane 7. shifted band using 5x probe in the heparin gel; Lanes 8 and 9, combined shifted bands, 2-10x probe in heparin gel; Lane 10, negative control, no probe in the heparin gel ,band from same Rf as the shifted band.


1
2
3
4
5
6
7
8
9
10

IV. Discussion

The yield of DNA-binding protein from this preparation was only 5 - 10 picomoles, which is near the limit of resolution for N-terminal sequencing, but consistent with the concentration of binding protein calculated in Chapter 2. Unfortunately, the size of this prep is also near the maximum limit for resolution by the shift assay, and of course, no overproducing strains for this protein are available. The strategy of using a preparative band shift gel to identify and isolate a protein present at very low concentrations was only viable because of the use of the heparin column as a first step. This column generally segregates positively charged proteins, specifically ones which bind poly-anions like DNA. Therefore, these proteins as a class are less likely than anionic proteins to complicate resolution in a subsequent native shift gel by migrating toward the anode based on their own mobility. The other DNA binding proteins in the active heparin column fraction are subject to competition from the large excess of non-specific DNA, while sequence-specific binding factors are left to migrate with the specific DNA probe.

Obtaining useful quantities of this protein in an active form is our ultimate goal. The sensitivity of the protein to even small levels of nonionic surfactant in a partly purified state appears to indicate that this binding factor is much more easily denatured as the concentration of total protein is reduced.

Ideally, the use of a specific sequence DNA column should effect isolation of the binding factor. In the case of the consensus palindrome sequence alone, the <u>in vitro</u> binding constant is much

lower than those of the eukaryotic proteins for which the procedure was developed, and appears to be too low for an effective separation. The larger binding site (bchC DNA) exhibits a much tighter affinity for the protein, however its use in a column resulted in complete loss of activity (see Results). This may be due to the fact that protein recognition elements are spread out over a region of about 75 base pairs, and crosslinks to the support are frequent enough to disrupt or disturb most of the sites on the column. The possibility of two different proteins or the presence of some cofactor being required for binding was also considered. Under this scenario, each factor could be separately released at different salt concentrations. The use of a single high salt elution to release all bound proteins still did not produce an active eluate, indicating that the situation was not quite that simple. It is still possible that the column separates the binding protein from a required cofactor, but perhaps this occurs at the initial binding stage, and may be due to the conformational constraints placed on the DNA by being crosslinked to the column support.

For future work, an alternate type of poly(bchC-200mer) DNA column should be constructed in which the active DNA concatemer is ligated in situ to compatible DNA previously crosslinked to the support. This would provide unhindered sites available for protein binding. Also, if denaturation of the factor at low total protein concentrations continues to be a problem, the addition of excess carrier protein (BSA, etc.) or DNA to all column wash solutions should be considered as ways of stabilizing the binding factor.

Chapter 4 :

In vitro Expression of Lycopene Cyclase and beta-Carotene Hydroxylase from *Erwinia herbicola*

I. Introduction

Erwinia herbicola is a nonphotosynthetic gram-negative bacterium, related to the common enterobacteria such as *E. coli*, but found in a wide variety of environments. Many strains of this bacterium are found as plant pathogens, some of which are capable of providing nucleation sites for ice crystal formation on plant tissues, thereby promoting frost damage to crops and other plants (Orser et al., 1985). Some *Erwinia* strains living on exposed plant surfaces produce polar, yellow colored pigments which were proposed to be glycosylated carotenoids by Starr (1981). All the genes required to produce these pigments have been cloned from the *Er. herbicola* strain Eho10 as a single 12.4 kb chromosomal fragment in a cosmid library and actively expressed on a plasmid in *E. coli* by Perry et al. (1986). The resulting yellow *E. coli* were afforded protection against near-UV light and photosensitizing agents, consistent with the pigments being carotenoids (Tuveson et al., 1988). These pigments were later definitively characterized in both *Er. herbicola* (Hundle et al., 1991) and the related *Er. uredovora* (Misawa et al., 1990) as mono- and diglucosides of zeaxanthin and cryptoxanthin, natural carotenoids of almost unique polarity. It was also confirmed that

the 12.4 kb cluster from Er. herbicola contains all the genes required for carotenoid production from the last isoprenoid condensation onward (Hundle et al., 1991), and a logical pigment biosynthesis pathway was proposed (Fig. 4-1).

The earlier steps of the carotenoid biosynthetic pathway leading to the first C_{40} carotenoid, phytoene are conserved across all carotenogenic organisms. These enzymes utilize soluble, phosphorylated substrates, and the enzyme activities for several of these reaction steps have been isolated in vitro from a variety of organisms . However, the subsequent enzymatic activities in the carotenoid pathway utilize lipophilic substrates, and their activities have proved difficult to isolate in vitro. Most of the enzymes for carotenoid biosynthesis after phytoene appear to be membrane bound, often in specialized photosynthetic membranes, and have been postulated to be clustered into a multienzyme complex (Beyer et al., 1985). Due to the paucity of observable intermediates in the plant and fungal systems studied, it has been proposed that the carotenoid substrate may be channeled through this multienzyme complex (Beyer et al., 1985; Candau et al., 1991).

No individual enzymes catalyzing the conversion of phytoene to the more abundant later carotenoids have been isolated and reported to date. However, certain individual reactions, such as cyclization, have been studied in relative isolation by the manipulation of cofactors (primarily O_2) to inhibit other enzymes in a crude isolate (Beyer et al., 1989). The combined carotene desaturation and cyclization activities have been demonstrated in vitro from the collection of membrane bound enzymes solubized

from Narcissus pseudonarcissus (daffodil) chromoplasts and reconstituted into liposomes (Beyer et al., 1985). These enzymes are responsible for the multiple dehydrogenations, isomerizations, and cyclizations which convert phytoene to the end product, β-carotene. The preferred substrates for cyclization to β-carotene in the daffodil system were lycopene isomers containing cis double bonds at the 7 and 7' positions, and NADPH was found to be an essential cofactor (Beyer et al., 1989) and (Beyer et al., 1991). In the fungus, Phycomyces blakesleeanus, evidence of an enzyme aggregate containing two cyclases which convert lycopene to β-carotene via γ-carotene has been reported (Candau et al., 1991).

The hydroxylation of β-carotene utilizes molecular oxygen, as demonstrated by $^{18}O_2$ labelling (Yamamoto et al., 1962) and (McDermott et al., 1974). In vitro experiments indicate that membranes of the cyanobacterium, Aphanocapsa, accomplish the hydroxylation of β-carotene using an O_2 dependent monooxygenase, and that the reaction is stimulated by NADPH (Sandmann and Bramley, 1985).

Due to their compatibility with E. coli, and ability to be actively expressed in vivo without the presence of a specialized membrane, the Erwinia gene products hold the promise of being a viable non-photosynthetic alternative for the study of carotenoid enzyme biochemistry. The crtE , crtB, and crtI genes have been mapped to sequenced open reading frames in the Erwinia cluster based on amino acid homology with the analogous genes in R. capsulatus (Armstrong et al., 1990b), despite an estimated 1.1 billion years of evolutionary separation (based on 16 S rRNA

comparisons, Wilson et al., 1987). The various later enzymatic steps, including cyclization and hydroxylation, have been assigned to specific loci within the Erwinia carotenoid gene cluster via mutagenesis, and this cluster has been sequenced in both Er. herbicola (Alberti and Hearst, EMBL submission M87280) and Er. uredovora (Misawa et al., 1990).

Using the sequenced DNA from Erwinia herbicola as an amplification template, very specific plasmids have been constructed to permit expression of the enzymes, lycopene cyclase and β-carotene hydroxylase, in E. coli under the control of an inducible T7 RNA polymerase promoter. The experiments in this chapter demonstrate the separate in vitro activities of these two enzymes to cyclize the termini of lycopene to form β-carotene, and to oxidize β-carotene to zeaxanthin.

Figure 4-1. Biosynthesis of carotenoids in Erwinia herbicola, general pathway from Hundle et al. (1991). The assignment of the crtB and crtI loci was based on amino acid homology with the corresponding gene products in R. capsulatus (Armstrong et al., 1990b; Alberti and Hearst, EMBL submission M87280). Characterization of the crtX, crtY, and crtZ loci was determined by identification of products accumulated when each locus was deactivated by mutation (Hundle et al., 1991).

Carotenoid Biosynthesis Pathway in *Erwinia herbicola*

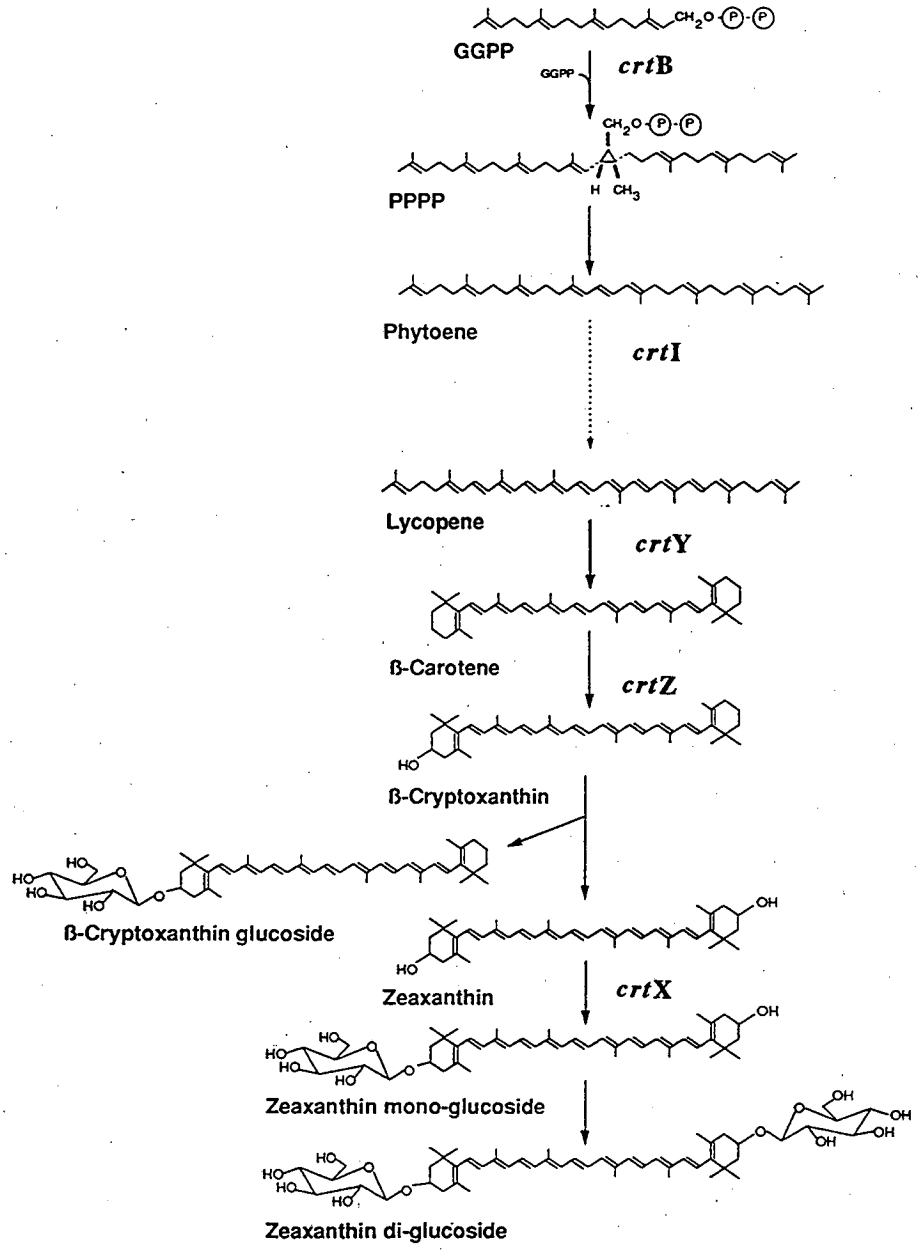
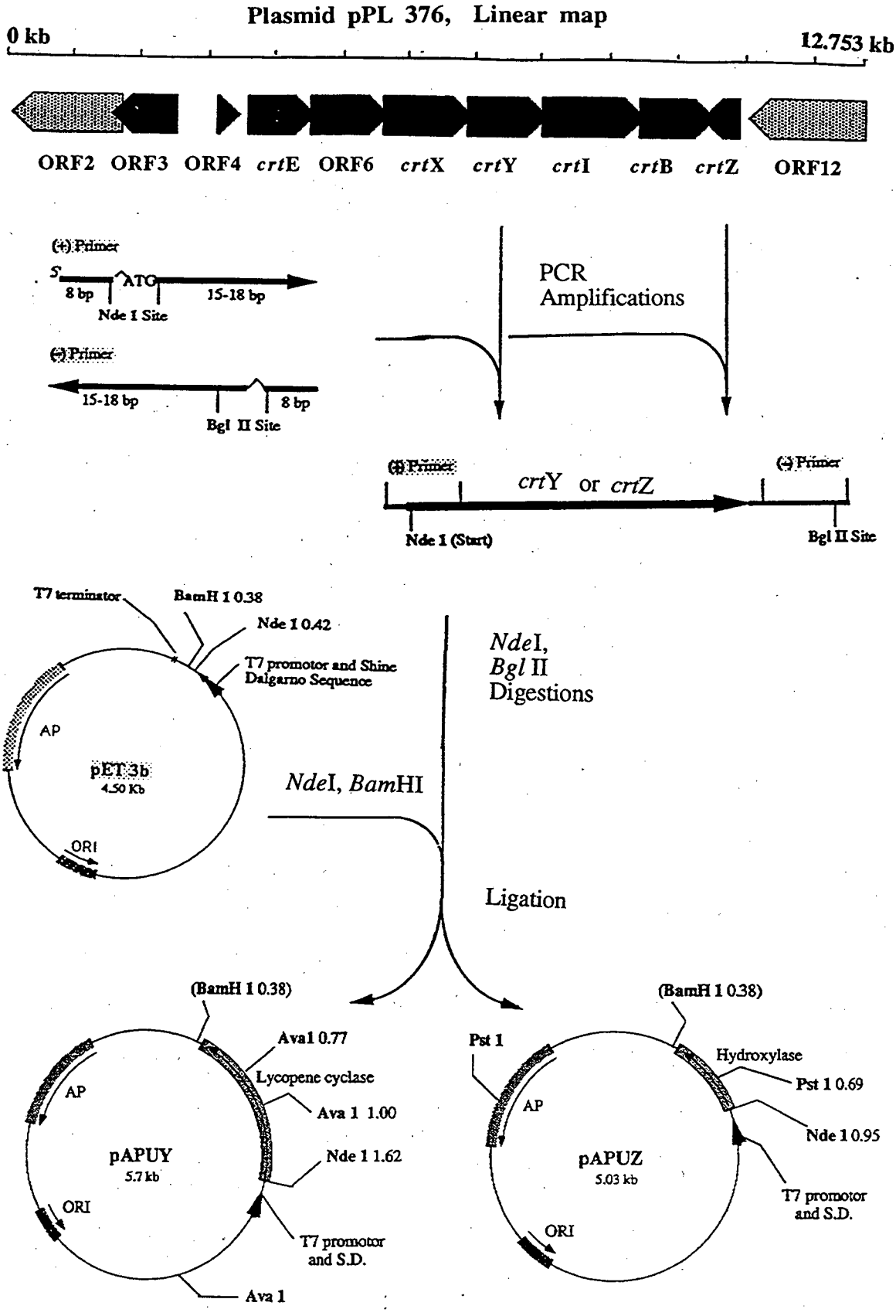


Figure 4-2. Construction of plasmids pAPUY and pAPUZ. Ap represents the ampicillin resistence gene. crtY is the Erwinia herbicola lycopene cyclase gene. crtZ is the E. herbicola β-carotene hydroxylase gene. Upstream of the inserted gene at an appropriate distance is a T7 promoter and Shine Dalgarno sequence. Downstream of the insert is a strong T7 terminator.


Plasmid pPL 376, Linear map
0 kb
12.753 kb
ORF2 ORF3 ORF4 crtE ORF6 crtX crtY crtI crtB crtZ ORF12
(+) Primer
5'
ATG
8 bp
15-18 bp
Nde I Site
(-) Primer
15-18 bp
8 bp
Bgl II Site
PCR
Amplifications
(+) Primer
crtY or crtZ
(-) Primer
Nde I (Start)
Bgl II Site
T7 terminator
BamH I 0.38
Nde I 0.42
T7 promotor and Shine
Dalgarno Sequence
AP
pET 3b
4.50 Kb
ORI
NdeI,
Bgl II
Digestions
NdeI, BamHI
Ligation
(BamH I 0.38)
Ava1 0.77
Lycopene cyclase
Ava 1 1.00
Nde 1 1.62
T7 promotor
and S.D.
Ava 1
AP
pAPUY
5.7 kb
ORI
(BamH I 0.38)
Pst 1
Hydroxylase
Pst 1 0.69
Nde 1 0.95
T7 promotor
and S.D.
AP
pAPUZ
5.03 kb
ORI

II. Materials and Methods

Organisms and growth conditions

The E. coli strains carrying the Erwinia herbicola genes contained on plasmids pAPU211, pAPUY (Fig. 2A), and pAPUZ (Fig. 2B) were grown at 37°C in LC medium, by selection for ampicillin resistance using 100 μg/ml ampicillin (Sigma Chemical Co., St. Louis, MO).

PCR amplification of the crtY and crtZ genes

The Erwinia herbicola lycopene cyclase (crtY) and β-carotene hydroxylase (crtZ) genes were amplified separately via polymerase chain reaction (PCR) using Thermus aquaticus Amplitaq™ DNA polymerase (Perkin Elmer Cetus, Emeryville, CA) in 100 μl of the standard buffer (Saiki et al., 1988) and (Loh et al., 1989). The N-terminal primers (crtY-n= 5'-GAGAGCGTCATATGAGGGATCTGATTTT-A, and crtZ-n= 5'-CGCGCACCCATATGCTAGTAAATAGTTTAATC, 1μM) each contained an NdeI restriction site to permit the ligation of the amplified gene directly into the translation start site of the pET-3b expression vector (Studier et al., 1990). The two C-terminal primers (crtY-c= 5'-GGCCACCAAGATCTGCGCCAATCACAAC, and crtZ-c= 5'-TGCCATGCAGATCTCGGCACCGGGGCAGG, 1μM) were designed to hybridize to a region immediately downstream of the gene, and to each contain a BglII site in order to permit insertion into the BamHI site of pET-3b. Linearized pAPU211 plasmid containing the E. herbicola carotenoid gene cluster was used as the template (0.84 ng, or 1.0 pM concentration) for both amplifications. Because insertion

of the restriction sites reduced the homology of the primers to the original template, an annealing temperature of 45°C was used for the first five cycles of amplification, with 55°C being employed in subsequent cycles. The amplification cycle was as follows: denaturation at 94°C for 1 min; annealing for 1.5 min as described above; extension at 72°C for 2.5 min. Amplifications were performed for 30 cycles, and the products were digested with NdeI and Bgl II and extracted from an agarose gel slice and precipitated with ethanol.

Plasmids, cloning techniques

Plasmid pAPU211, was derived from pPL376 (Perry et al., 1986; Hundle et al., 1992a) and contains the essential E. herbicola genes for carotenoid production. Plasmids pAPUY and pAPUZ (Fig. 4-2) were constructed by ligating PCR amplified crtY and crtZ, respectively, into the NdeI and BamHI restriction sites of the pET-3b vector (Studier et al., 1990). The plasmids were each cloned and maintained in E. coli DH5α cells. Clones containing pAPUY and pAPUZ were screened by the mini-prep method of Riggs and McLachlan (1986). Nucleic acid and enzymatic manipulations were performed according to standard published procedures (Sambrook et al., 1989) or manufacturers' protocols, and with the technical assistance of B. Hundle. For the overexpression of lycopene cyclase and β-carotene hydroxylase, each plasmid was transformed into E. coli BL21(DE3).

SDS gel electrophoresis

Induced proteins were analyzed by an SDS-10% PAGE gel using the discontinous buffer system of Laemmli (1970). Intact E. coli cells were centrifuged and resuspended in 400 µl 1X sample buffer per ml. Samples were boiled for 4 min and quenched on ice. Proteins were separated in a Bio-Rad minigel apparatus, and visualized by staining the gel with Coomassie Blue R250 and destaining in a solution of 40% methanol, 10% acetic acid, and water.

To obtain proteins for N-terminal sequencing, an unstained gel was electroblotted onto an Immobilon^{TM} membrane, according to the manufacturer's protocol (Immobillon Tech Protocol TP006, Millipore Corp., Bedford, MA). The membrane was stained with Coomassie Blue, and bands of interest were excised from the membrane and sequenced.

Substrate Materials

All-trans-lycopene, 7,9,9',7'tetra-cis-lycopene, 5,5'di-cis-lycopene,as well as δ-carotene were kindly provided by Hoffman La-Roche (Basel, Switzerland). All-trans- β-carotene was obtained from Sigma Chemical Co.(St. Louis, MO). γ-Carotene was isolated from Chloroflexus aurantiacus.

Preparation of enzyme extract:

E. coli BL 21(DE3) cells containing the plasmid pAPUY or pAPUZ were grown to a cell density of about 3.5 X 10^8 per ml and induced with 0.4 mM isopropyl β-D-thio-galactopyranoside (IPTG) for 30 min. at 37°C. To account for difficulties in the reproduction

of enzyme activities (associated with misfolding of the over produced proteins, see below), inductions were also performed at 15°C and for shorter (minimal 10 min) and longer (maximal 3.5 hours) periods of time. For the preparation of enzyme extracts (Method A), cells were pelleted, frozen in liquid nitrogen, and thawed on ice. All further steps were carried out at 4°C unless otherwise stated. The pellets obtained from 200 ml cultures were washed and suspended in 3 ml Buffer A (0.2M Hepes, pH 7.0, 1mM EDTA, 1mM PMSF), aliquoted (1ml), and allowed to stand for 30 min after the addition of 0.100 ml of a lysozyme solution (1mg/ml). A low-pressure French Press passage (4000 psi) completed disintegration. This lysate was used directly for incubations with carotene substrates or submitted to fractionation. A 15 min centrifugation at 13,000x g was used to remove large debris as well as inclusion bodies, and a subsequent centrifugation of the supernatant at 120,000x g for 1.5 hours yielded a soluble fraction and a membrane pellet. The membrane pellet was resuspended in the original volume of buffer, and both fractions were incubated separately.

Alternatively (Method B), the induced cells were washed in Buffer B (100 mM Tris HCl, pH 7.4) and disintegrated by use of a French Pressure Cell at 15,000 psi. Fractionation was performed as outlined above, but in Buffer B.

Enzyme assays

Samples of the individual carotene substrates (5 nmol) dissolved in 20 µl acetone were added to 1 ml enzyme assays. The

cofactors examined in the hydroxylase assays were NADPH, NADH, FAD, ascorbate, 2-oxoglutarate (1 mM each), and Fe^{2+} (20μM). An enzymatic oxygen trap, according to the method of Lam and Malkin (1982), was used to stimulate the reaction, based on results involving the cyclase reaction in Narcissus pseudonarcissus chromoplasts (Beyer et al.,1989). The reactions were stopped by extracting with chloroform/ methanol (2:1 v/v) after an incubation at 28°C, performed for 8 hours, if not stated otherwise. The extracts were dried under a stream of nitrogen and analyzed by HPLC employing an ET 300/8/4 Nucleosil $5C_{18}$ Column (Macherey-Nagel). The column was developed isocratically with 5% tetrahydrofuran in acetonitrile at a flow rate of 1 ml/min. The enzyme assays were performed in collaboration with P. Beyer.

III. Results

A. Cloning of genes and over-production of proteins

The plasmid constructs, pAPUY and pAPUZ, are outlined in figure 4-2. PCR-amplified Erwinia herbicola DNA containing crtY or crtZ was inserted immediately beyond the T7 promoter and Shine Dalgarno sequence of the pET-3b expression vector. A Φ10 T7 termination signal is located downstream of the inserted DNA. Following an initial screening of colonies by plasmid size, restriction patterns were used to verify the identity and orientation of the inserts (Figs. 4-3A and 4-3B).

These pAPUY and pAPUZ constructs were each maintained in E. coli DH5α cells, and the crtY and crtZ genes expressed separately in E. coli BL21(DE3) via induction with IPTG. The accumulation of a new 43 kD protein was observed on a 10% SDS-polyacrylamide gel (Fig. 4A) upon induction of pAPUY with IPTG. The calculated molecular weight based on the derived amino acid sequence of the protein is also 43 kD. The identity of this band as the crtY gene product was confirmed by N-terminal amino acid sequencing. Likewise, the accumulation of a newly formed 22 kD protein was observed on a 10% SDS-polyacrylamide gel (Fig. 4B) upon induction of pAPUZ with IPTG, although the amount of expressed protein was considerably less than that of the cyclase under the same conditions. The calculated molecular weight based on the derived amino acid sequence of the hydroxylase enzyme is also 22 kD, and the identity of this band as the crtZ gene product was confirmed by N-terminal amino acid sequencing. While the expressed proteins

are present as major bands in the induced cells, uninduced cells also showed traces of these proteins due to the basal activity of the T7 RNA polymerase (Figs.4-4A and 4-4B).

B. In vitro enzymatic activity of proteins

Extracts from E. coli strains containing the overexpressed proteins were assayed in vitro for their individual activities. Hydroxylase activity could be obtained when cells were grown at 15°C and induced for 10 h at the same temperature. Figure 4-5 shows an incubation of the lysate (Method A), performed in the absence of any externally added cofactors. The formation of zeaxanthin, as well as the formation of some intermediate cryptoxanthin was observed. Structural identity was demonstrated by HPLC, both by coelution with the authentic reference, and with the aid of the congruent spectra taken with a photodiode array detector. In fractionation experiments, the enzyme activity was found in the supernatant of the 120,000x g centrifugation, wheras the membrane fraction was inactive. Moreover, the membranes were actually found to be inhibitory, since remixing of the supernatant with increasing amounts of membranes back to the original content abolished the original activity completely. This is qualitatively shown in figure 4-6. The β-carotene hydroxylase reaction proceeded in the absence of externally added cofactors; however, electron donors, such as NAD, NADPH, and also ascorbate were able to stimulate the reaction.

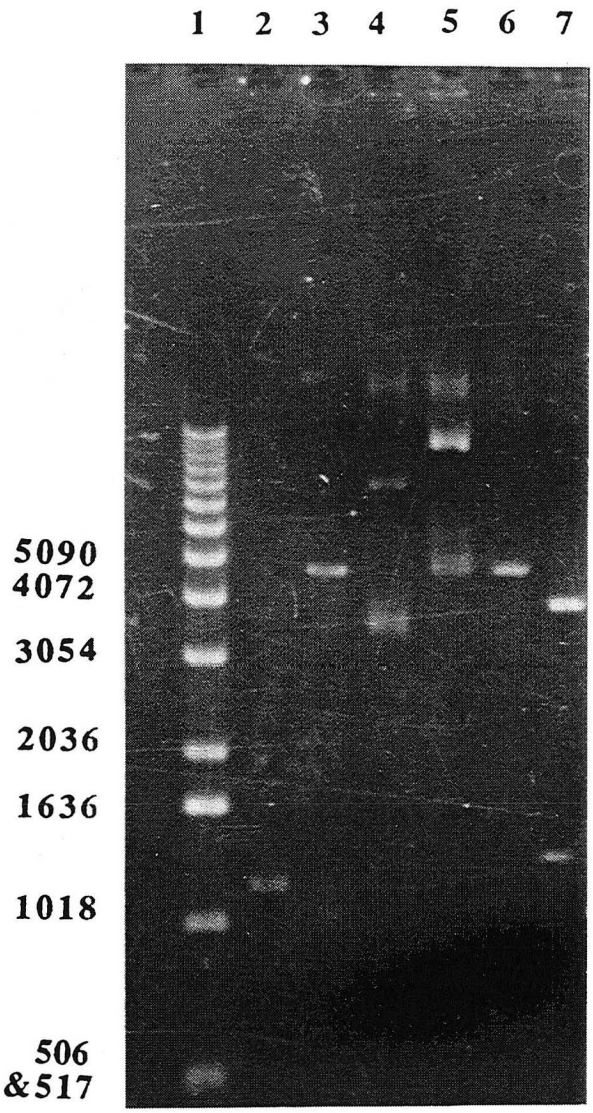
Although massively overproduced (see Fig. 4-4A), and more

hydrophilic based on its amino acid composition (Fig. 4-8A), the lycopene cyclase exhibited less enzymatic activity than the β-carotene hydroxylase. In addition, the cyclase activity was less reproducible. This could not be improved upon, either by numerous variations in the induction regime, or by fractionation of the cells. Anaerobic conditions administered by the use of an enzymatic oxygen trap, while essential for the function of the Narcissus pseudonarcissus cyclase in vitro (Beyer et al., 1989), had no stimulative effect on this Erwinia system. When extracts from E. coli strains containing pAPUY were assayed for in vitro lycopene cyclase activity by incubation with various lycopene isomers, all-trans-lycopene was found to be the only acceptable substrate for this Erwinia cyclase. An example of the conversion of all-trans-lycopene into all-trans-β-carotene in a cell lysate (Method B), obtained after 20 min induction at 37°C, is given in figure 4-7.

Figure 4-3. 0.7% Agarose gels showing DNA used in cloning enzymes from E. herbicola: (**A**), lycopene cyclase (CrtY): Lane 1, DNA size standards; Lane 2, PCR amplification product, digested with NdeI and BglII, used as insert; Lane 3, pET3b vector digested with NdeI and BamHI; Lane 4, uncut pET3b (nicked and supercoiled forms visible); Lane 5, uncut pAPUY construct (nicked and supercoiled); Lane 6, AvaI digest of pET3b; Lane 7, AvaI digest of pAPUY; and (**B**) β-carotene hydroxylase (CrtZ): Lane 1, DNA size standards; Lane 2, PCR amplification product, digested with NdeI and BglII; Lane 3, pET3b vector digested with NdeI and BamHI; Lane 4, uncut pET3b (nicked and supercoiled); Lane 5, uncut pAPUZ (nicked and supercoiled); Lane 6, PstI digest of pET3b; Lane 7, PstI digest of pAPUZ.

(A)

Lycopene Cyclase



Ⓑ β-Carotene Hydroxylase

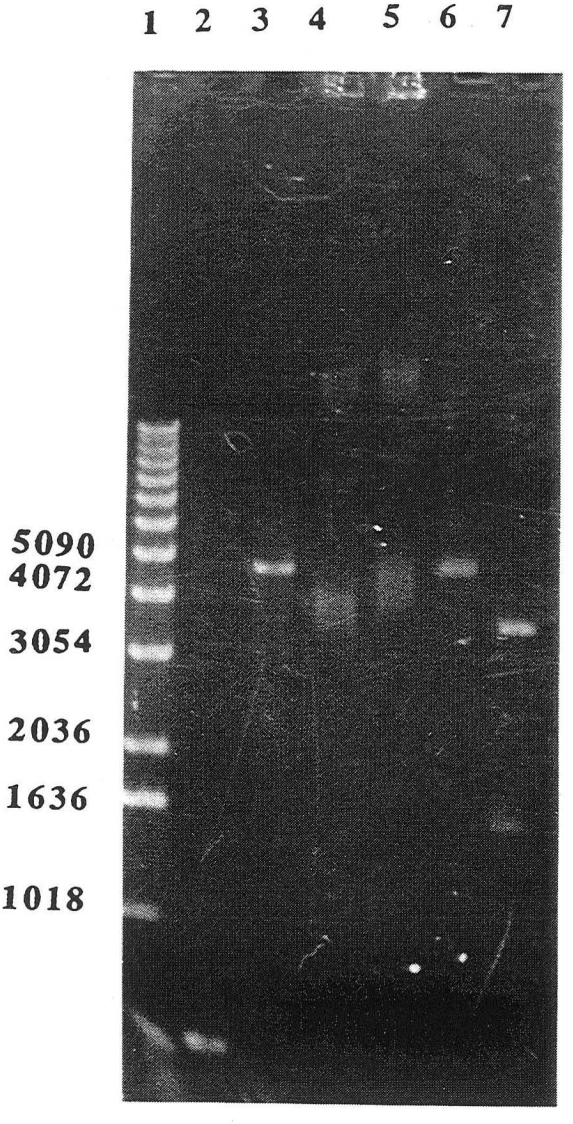
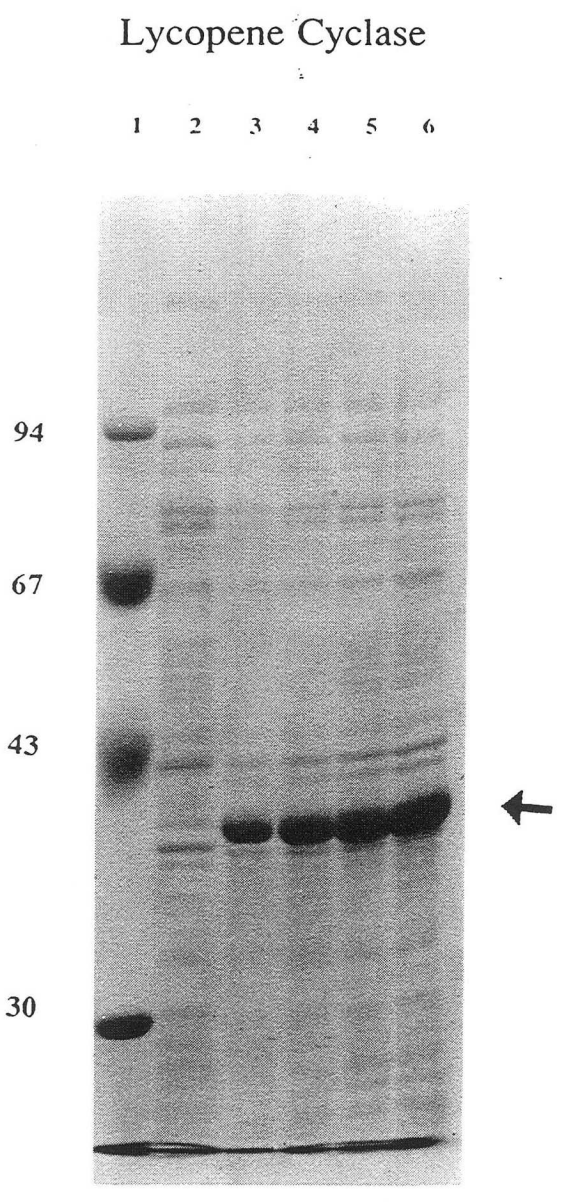


Figure 4-4. 10% SDS-PAGE gels showing expression in E. coli of: (**A**), lycopene cyclase; and (**B**), β-carotene hydroxylase. In each gel: Lane 1, protein standard markers of molecular mass 94, 67, 30, 20.1, and 14.4 kDa; Lane 2, total protein from uninduced whole E. coli cells containing the plasmid bearing the gene; Lanes 3 through 6, increasing amounts of total protein from induced cells.

(A)



Ⓑ β-Carotene Hydroxylase

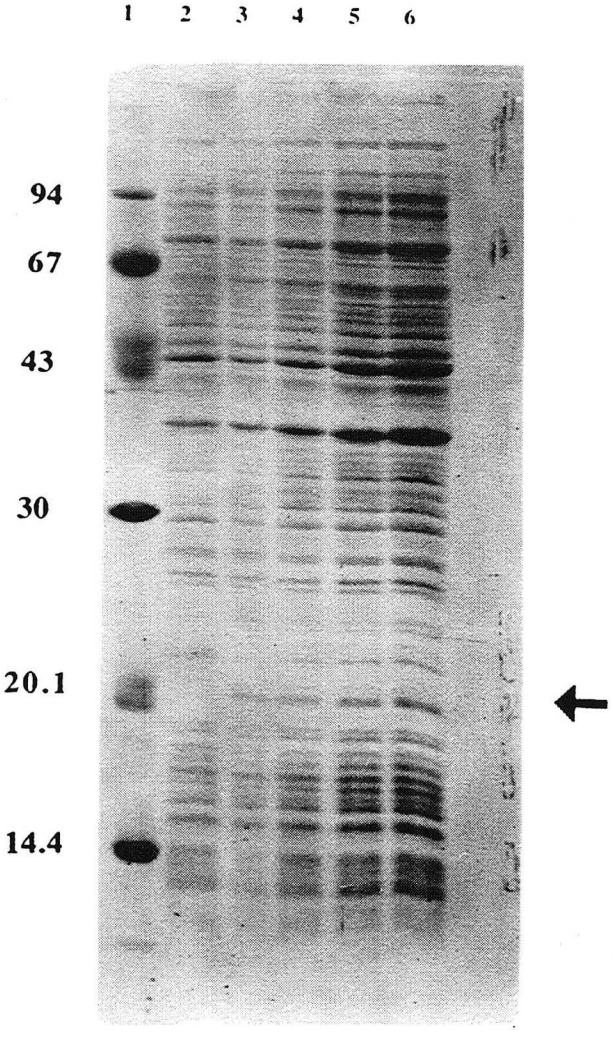


Figure 4-5. HPLC analysis of an incubation of an E. coli cell lysate (Method A) containing the overexpressed β-carotene hydroxylase with β-carotene as the substrate. Refer to materials and methods for the incubation conditions. 1, β-carotene; 2, β-cryptoxanthin; 3, zeaxanthin. (**A**), substrate (β-carotene) co-incubated in the presence of denaturing amounts (2 vol.) of $CHCl_3$/MeOH (2/1,v/v) as a control. (**B**), analysis of the enzymatic conversion.

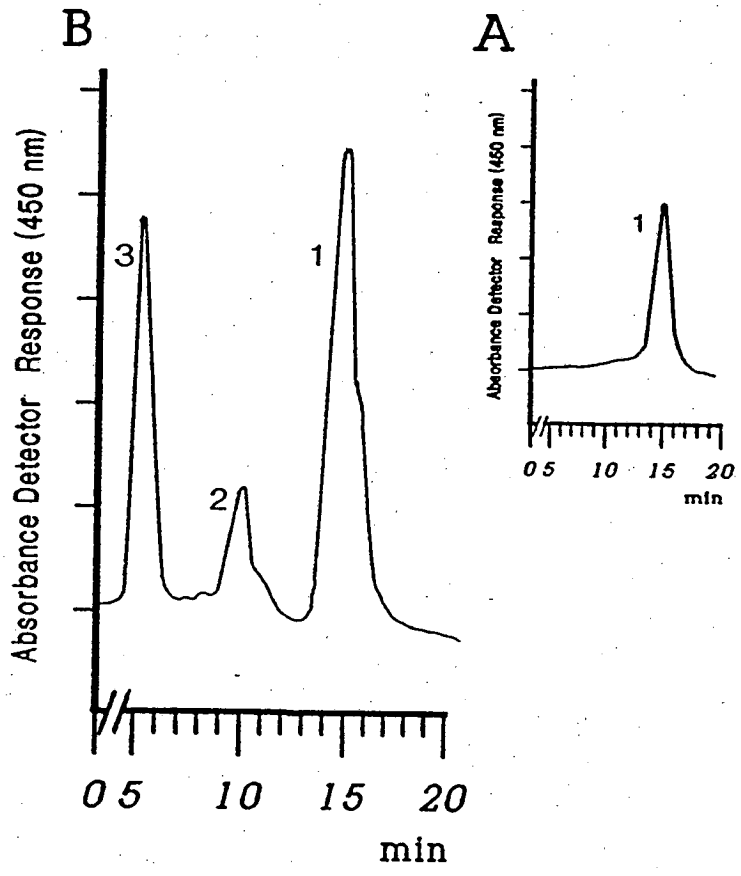

B
A
Absorbance Detector Response (450 nm)
3
2
1
0 5 10 15 20
min
Absorbance Detector Response (450 nm)
1
0 5 10 15 20
min

Figure 4-6. HPLC analysis of an incubation of an E. coli supernatant (**A**), derived from a cell lysate (Method B) with β-carotene; (**B**), conversion after readdition of 25% of the original membrane content; and (**C**), after readdition of 50% of the original membrane content. 1, β-carotene; 2, zeaxanthin. Cofactors are NADPH, NADH, FAD, ADP, Pi, 1mM each.

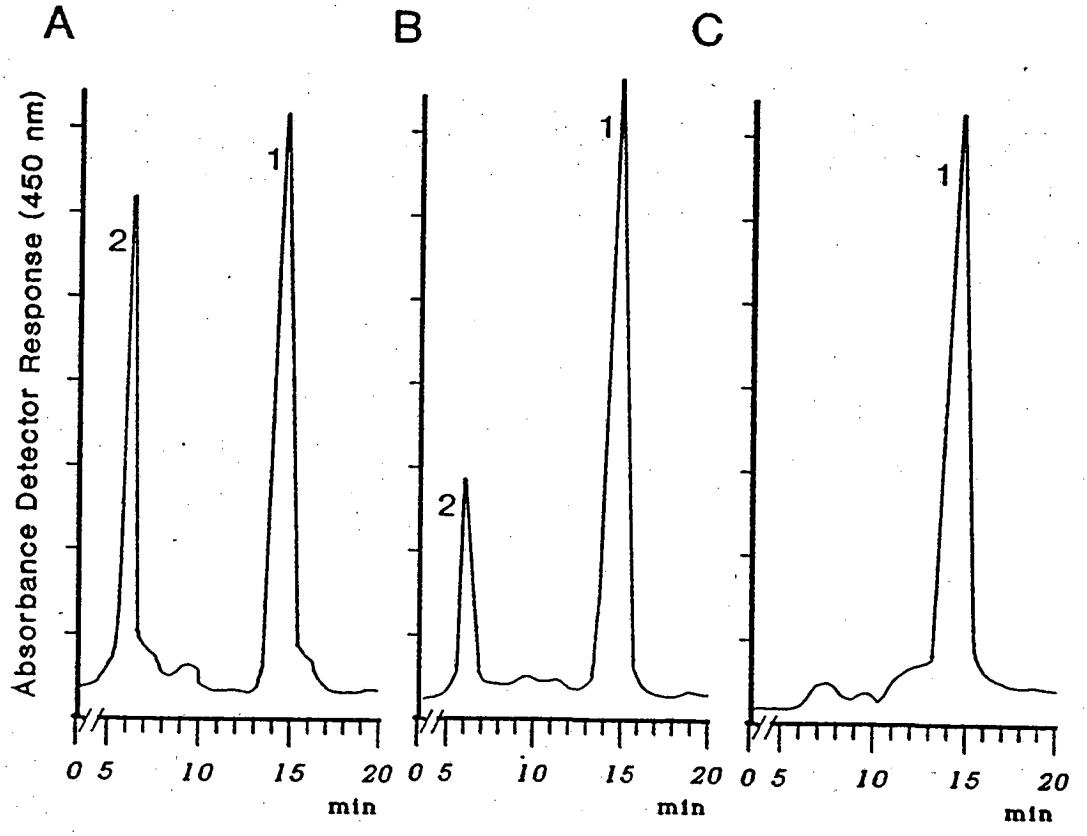

A
B
C
Absorbance Detector Response (450 nm)
1
2
0 5 10 15 20
min

Figure 4-7. HPLC analysis of an incubation of an E. coli cell lysate (Method B) containing the overexpressed lycopene cyclase. (**A**), substrate (all-trans-lycopene) co-incubated in the presence of denaturing amounts (2 vol.) of $CHCl_3$/MeOH (2/1,v/v) as a control. (**B**), analysis of the enzymatic conversion: 1, all-trans-lycopene; 2, β-carotene.

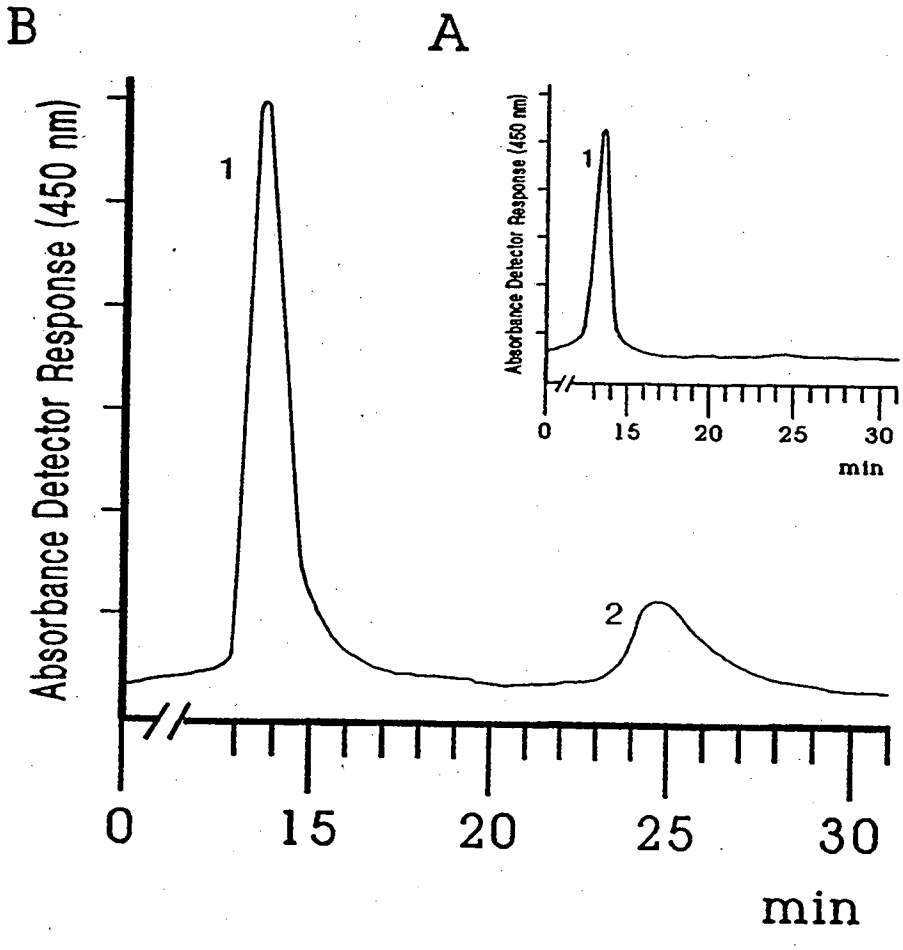
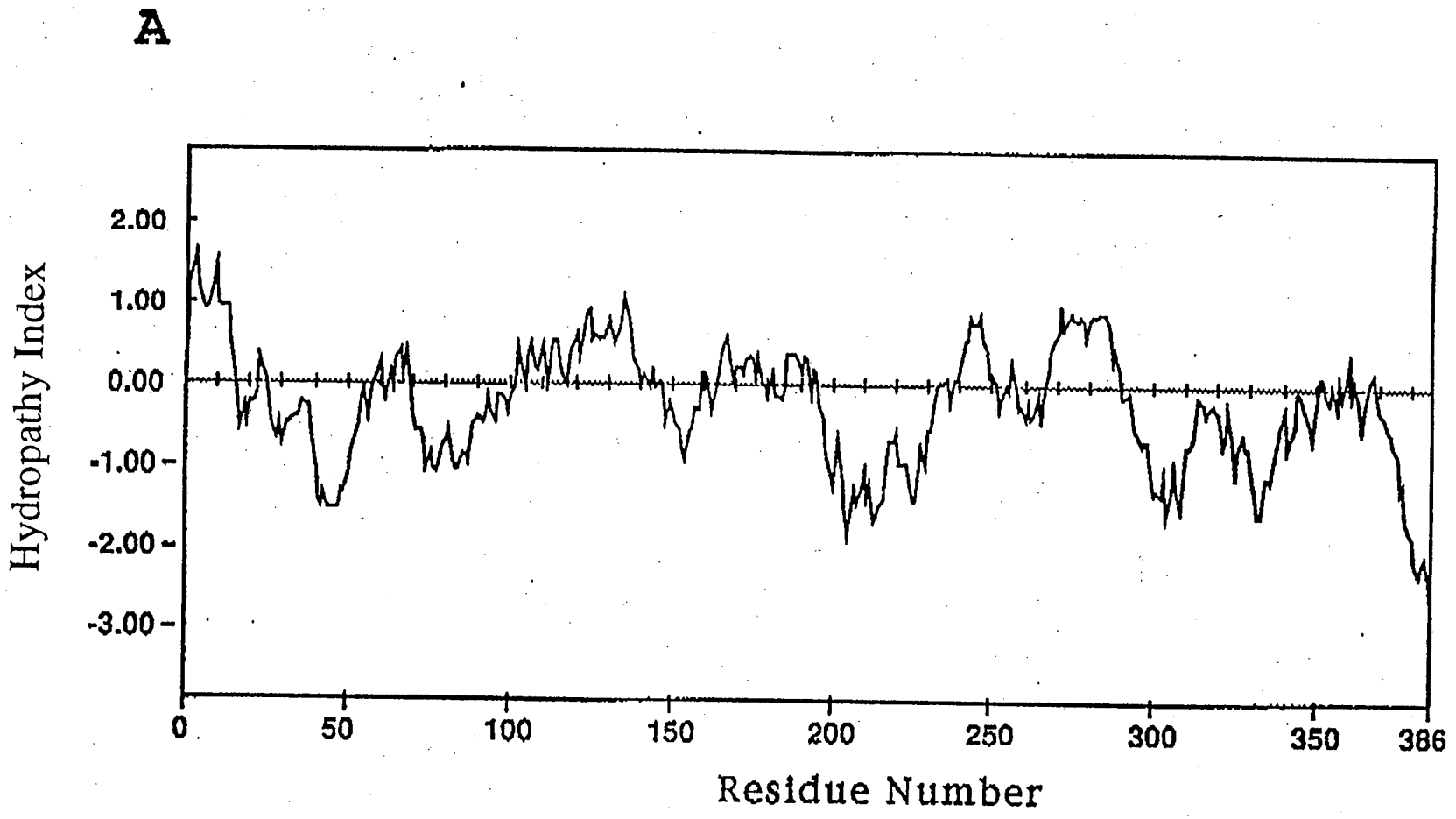
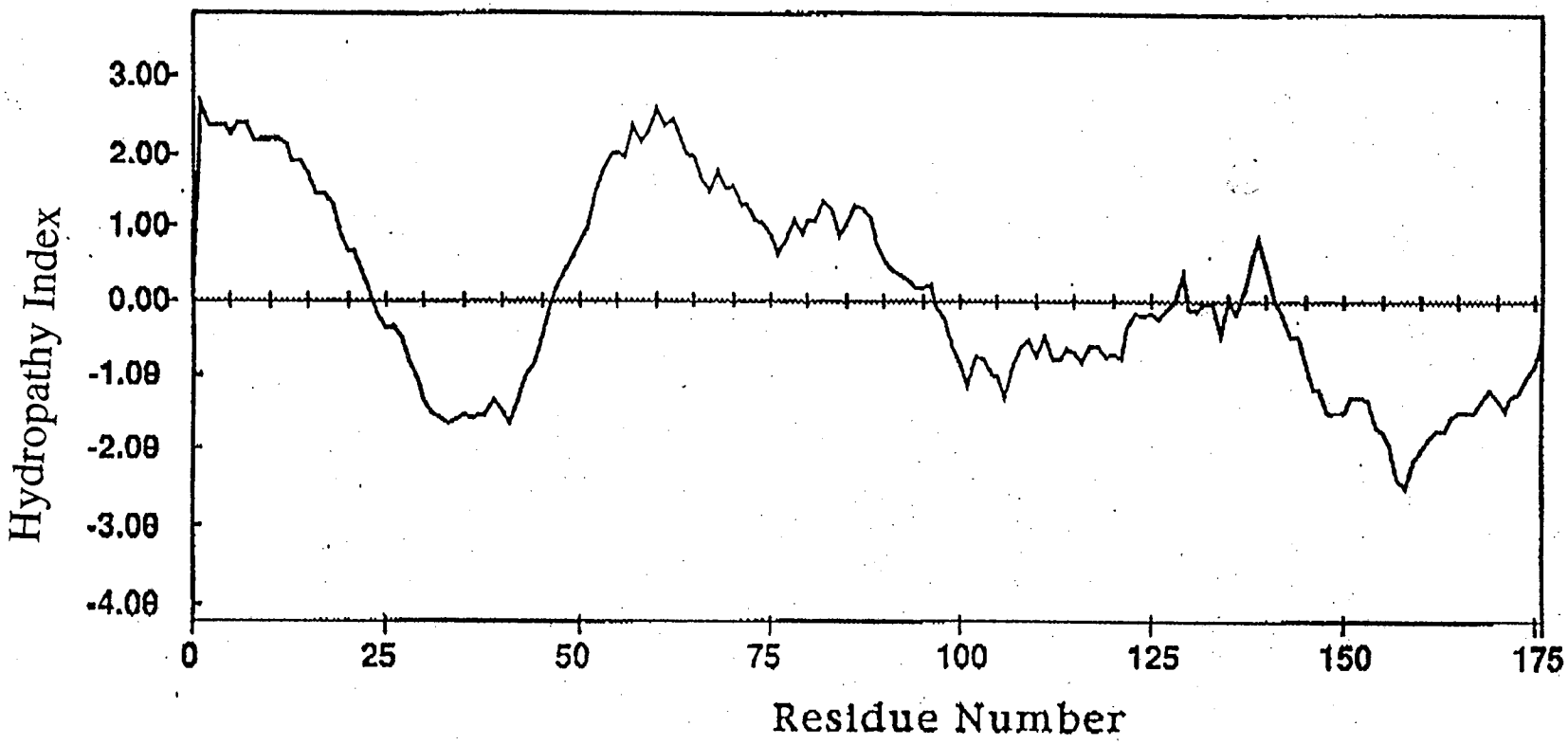

B
A
Absorbance Detector Response (450 nm)
1
2
0
15
20
25
30
min
Absorbance Detector Response (450 nm)
1
0
15
20
25
30
min

Figure 4-8. Hydropathy profile for the deduced amino acid sequences of: (**A**) CrtY (lycopene cyclase); and (**B**) CrtZ (β-carotene hydroxylase). The curve above the midline represents hydrophobicity and below represents hydrophilicity. The computer generated profile was determined using a moving window of 19 amino acid residues as recommended by Kyte and Doolittle (1982). A hydropathy index of +1.6 or greater for a segment containing at least 19 amino acids is indicative of a membrane spanning region.


A
Hydropathy Index
2.00
1.00
0.00
-1.00
-2.00
-3.00
0
50
100
150
200
250
300
350
386
Residue Number

B


Hydropathy Index
3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00
-4.00
0
25
50
75
100
125
150
175
Residue Number

IV. Discussion

Both the lycopene cyclase and the β-carotene hydroxylase from Er. herbicola were able to be overexpressed in E. coli in this study, although much of the cyclase protein ended up in inclusion bodies. Both overexpressed enzymes exhibited some degree of in vitro activity in cell-free lysates, unlike the previous attempts to achieve in vitro expression using post-phytoene biosynthesis enzymes derived from other organisms. In our results, each of the two enzymes could act separately and process exogenous substrates, indicating that neither the clustering of the carotenoid biosynthesis enzymes nor the channelling of intermediates was required, as had been proposed (Beyer et al., 1985; Candau et al., 1991). Upon fractionation, in vitro β-carotene hydroxylase activity was found to be localized in the high speed supernatant, while lycopene cyclase activity did not appear to segregate (see Results).

The levels of activity of both enzymes were found to be fairly low, particularly that of the lycopene cyclase. This result was not unexpected, considering the fact that both enzymes require extremely hydrophobic substrates which are introduced into a basically aqueous environment. In addition, prolonged incubations also led to significant bleaching of substrates and products, as expected. In an attempt to optimize enzyme activity, a variety of conditions were employed in both the induction period and the in vitro incubation step for each enzyme. Variables studied in the induction regime include time of induction, growth temperature,

and cell density at the onset of induction. In an extensive manipulation of these variables, the enzymatic activities were not always found to be reproducible, probably due to the misfolding of the overproduced proteins.

The enhancement of β-carotene hydroxylase activity upon the addition of electron donors (see Results) is consistent with a mixed-function oxygenase mechanism for carotene hydroxylation in which O_2 is first reduced to a more reactive species (e.g. a hydroperoxide). However, the presence of the hydroxylase activity in the high-speed supernatant instead of the membrane fraction was not expected, in the light of earlier studies in which the late steps of carotenoid biosynthesis were localized in the membrane (Beyer et al., 1985; Sandmann and Bramley, 1985; Beyer et al., 1991). Also, a hydropathy plot of the amino acid sequence of this enzyme indicates the presence of several regions of considerable hydrophobicity (Fig. 4-8B), consistent with an enzyme which is likely to be membrane bound. The inhibitory effect of the membrane itself may indicate that these complicated redox reactions could be severely disturbed upon lysis when localized in an inappropriate topological relation to another dominating redox phenomenon, such as the respiratory chain.

All-trans-lycopene was found to be the only acceptable in vitro substrate for this Erwinia lycopene cyclase, in contrast to the findings with Narcissus pseudonarcissus chromoplasts (Beyer et al., 1991) which only converted prolycopene or other cis-isomers of lycopene to β-carotene. The Erwinia cyclase protein itself is not particularly hydrophobic (Fig. 4-8A), and might have been expected

to have been more active in vitro than the hydroxylase. However cyclase overproduction in E. coli results in an enormous amount of protein being removed in an inactive form in inclusion bodies, which might explain the loss in activity. In fact, activity could be obtained only with the use of short induction times to minimize the production of excess protein and the resulting partitioning of protein into inclusion bodies.

The E. coli system, although very well suited for the overproduction of protein, may not be the ideal vehicle for these in vitro biochemical investigations of the cyclase and hydroxylase reactions. Therefore, the development of more sophisticated strategies as well as additional incubation systems will be pursued in the future, to better elucidate these enzyme mechanisms. Taking advantage of the similarity between the Erwinia strains and E. coli in their regulation of transcription, one strategy for studying these enzymes would involve introducing the plasmid-borne genes back into Er. herbicola wild type and/or mutant strains and using an Erwinia-based in vitro system.

Chapter 5 :

Expression and Characterization of Functionally Active Zeaxanthin Glucosyltransferase from Erwinia herbicola

I. Introduction

Numerous hydrophobic compounds are converted to more water soluble products via condensation with activated glucuronic acid or an activated sugar. The enzymes catalyzing such condensation reactions are known as glucuronosyl- or glycosyl-transferases and have been studied most extensively in xenobiotic contexts. The lipophilic substrate for such an enzyme must have an appropriate substituent, typically a hydroxyl or carboxyl group, which can be covalently modified via glycosylation. In mammals, UDP-glucuronosyltransferase is primarily a liver microsomal enzyme which catalyzes the transfer of glucuronic acid from UDP-glucuronic acid to many endogenous substances such as steroids, bilirubin, thyroid hormones or amines. This condensation reaction is of critical importance in the removal of many exogenous compounds, such as drugs, food additives, pesticides, and other ingested compounds, which may be toxic or carcinogenic (Dutton, 1980).

In plants, glycosylation of flavonoids is common. It is most widely studied in citrus fruit. The bitter taste of grapefruit is due to the presence of flavanone-neohesperidosides, whereas the isomeric flavanone-rutinosides predominant in oranges and lemon are

tasteless (Bar-Peled et al., 1991). One of the bronz loci (Bz-McC allele) in maize codes for UDP-glucose flavonoid glucosyltransferase, which catalyzes one of the last steps in anthocyanin biosynthesis. This enzyme has been cloned and sequenced (Ralston et al., 1988).

Ecdysteroids are insect molting hormones which are essential for normal development in lepidopteran species. A baculovirus, Autographa californica blocks insect molting by interfering with ecdysteroid biosynthesis. The virus achieves this by inserting a gene which codes for ecdysteroid glucosyltransferase into the host genome (O'Reilly and Miller, 1989). The product of this enzyme is a glucosyl- ecdysteroid, which is not recognized by the developing larvae, and as a result, the molting process is arrested.

In bacteria, an example of glycosylation of a lipophilic substrate occurs in carotenoid biosyntheis. Many common carotenoids, such as α- and β-carotenes and their xanthophyll derivatives, have undergone cyclization of their termini into six membered rings (Straub, 1987). Two examples of cyclic carotenoids with glycosylated ring hydroxyl groups have been identified: rhamnosylated zeaxanthin in Corynebacteria (Nybraaten and Liaaen-Jensen, 1974) and glucosylated zeaxanthin in Erwinia herbicola (Hundle et al., 1991) and Erwinia uredovora (Misawa et al., 1990). The diglucosylated zeaxanthin found in Erwinia is among the most polar of the natural carotenoids, with a water solubility of 800 ppm, compared to 100 ppm for zeaxanthin monoglucoside (Pfander and Hodler, 1974) and 12.4 ppm for zeaxanthin itself (Hundle et al., 1992b). The genes coding for zeaxanthin diglucoside

production from farnesylpyrophosphate (FPP) are clustered in the two Erwinia species. This gene cluster from Erwinia herbicola has been cloned and expressed in Escherichia coli, resulting in yellow colored E. coli (Perry et al., 1986). The various enzymatic steps have been assigned to specific loci within the Erwinia carotenoid gene clusters via mutagenesis, and the corresponding genes have been sequenced in both Er. herbicola (Alberti and Hearst, 1991, EMBL submission 87280) and Er. uredovora (Misawa et al., 1990). The experiments in this chapter demonstrate the expression of the E. herbicola enzyme, zeaxanthin glucosyltransferase, in E. coli under the control of an inducible T7 RNA polymerase promoter, as well as the enzyme's in vitro activity. A sequence comparison with other UDP-glycosyl transferases results in the identification of a putative UDP binding domain.

II. Materials and Methods

Organisms and growth conditions.

The E. coli strains carrying the Erwinia herbicola genes contained on plasmids pAPU211 (Hundle et al. 1992a) and pAPUX (Fig. 5-1) were grown at 37°C in LC medium, by selection for ampicillin resistance using 100 μg/ml ampicillin (Sigma Chemical Co., St. Louis, MO).

PCR amplification of the crtX gene

The Er. herbicola zeaxanthin glucosyltransferase gene, crtX, was amplified via polymerase chain reaction (PCR) using Amplitaq^{TM} DNA polymerase (Perkin Elmer Cetus, Emeryville, CA) in 100 ml of the standard buffer (Saiki et al., 1988) and (Loh et al., 1989). The N-terminal primer (5'-GGGATACCATATGAGCCATT-TTGCCATTG, 1μM) contained an NdeI restriction site to permit the ligation of the amplified gene directly into the translation start site of the pET-3b expression vector (Studier et al., 1990). The C-terminal primer (5'-AAATCAGATCTCTCACGATACGCTCTCACT, 1μM) was designed to hybridize to a region immediately downstream of the gene, and to contain a BglII site in order to permit insertion into the BamHI site of pET-3b. Linearized pAPU211 plasmid containing the E. herbicola carotenoid gene cluster was used as the template (0.42 μg in 100 μl, or 500 pM, a high concentration required because of competitive reactions). The amplification cycle was as follows: denaturation at 94°C for 1 min.; annealing at 45°C for 1.5

min.; a gradual temperature increase to 72°C over a 2 min. period; and a further primer extension at 72°C for 1 min. Amplification was performed for 30 cycles.

Plasmids and cloning techniques

Plasmid pAPU211 was derived from pPL376 (Perry et al.,1986) and contains the essential E. herbicola genes for carotenoid production. Plasmid pAPUX (Fig. 5-1) was constructed by digesting PCR-amplified crtX DNA with NdeI and Bgl II, and ligating it into the NdeI and BamHI restriction sites of the pET-3b vector according to the method of Studier et al. (1990). The plasmid was cloned and maintained in E. coli DH5α cells. Clones containing pAPUX were screened by the mini-prep method of Riggs and McLachlan (1986). The nucleic acid and enzymatic manipulations were performed according to standard published procedures (Sambrook et al., 1989) or manufacturers' protocols, and with the technical assistance of B. Hundle. For the overexpression of glucosyltransferase, plasmid pAPUX was transformed into E. coli BL21(DE3).

SDS gel electrophoresis

Intact E. coli cells were centrifuged and resuspended in 400 µl 1X sample buffer (Laemmli, 1970) per ml of culture and boiled for 4 min. Cell lysate fractions were also treated similarly, and proteins were analyzed by a 10% SDS-PAGE gel using the discontinous buffer system of Laemmli (1970). Proteins were stained with Coomassie Blue R250 and destained in 40% methanol, 10% acetic acid and water.

To obtain proteins for N-terminal sequencing, an unstained gel was electroblotted onto an Immobilon^{TM} membrane, according to Immobillon Tech Protocol TP006 (Millipore Corp., Bedford, MA). The membrane was stained with Coomassie Blue, and bands of interest were cut from the membrane and sequenced.

Preparation of enzyme extract

E. coli BL 21(DE3) cells containing the plasmid pAPUX were grown to a cell density of about 3.5 X 10^8 per ml and induced with 0.4 mM IPTG for 45 min. All further steps were carried out at 4°C unless otherwise stated. Cells were harvested by centrifugation at 4200x g for 10 min. The resulting pellet from a 100 ml cell suspension was resuspended in 1.0 to 4.0 ml of 0.05 M Tris-HCl buffer at pH 7.5, containing 1 mM βME, and 0.1 mM PMSF. Crude cell lysate was obtained by one passage through a French pressure cell at 13,000 psi. Large cell debris and unbroken cells were removed by centrifugation at 3100x g for 5 min.

Fractionation of the crude cell lysate was performed by centrifuging the sample at 13,000x g for 15 min. The pellet obtained in this manner is indicative of segregation of excess overexpressed protein in inclusion bodies. The low-speed spin supernatant was ultra-centrifuged at 100,000x g for 90 min. The pelleted membrane fraction was resuspended in 240 µl of the Tris buffer described above. A 120 µl fraction was saved for analysis, and the remaining 120 µl was washed with 4 ml of Tris buffer and again ultra-centrifuged for 90 min. This final pellet was resuspended in 120 µl of Tris buffer and is referred to as the

washed membrane fraction. Protein concentrations were determined by Bradford Assay (Bradford, 1976).

Enzyme assay

Glucosyltransferase activity was assayed by measuring the incorporation of [^{14}C]glucose into zeaxanthin. The assay mixture was prepared as follows: 20 to 80 µl enzyme extract; 13.5 nmol zeaxanthin in 5 µl acetone; and 100 nmol uridine diphosphate glucose (UDP-glucose) at a specific activity of 2.95 mCi/mmol ^{14}C, in 0.05 M Tris-HCl pH 7.5 containing 1mM βME to a final volume of 100 µl. UDP-glucose, radiolabelled at all six carbon atoms, was purchased from Dupont-NEN (Wilmington, DE) at a specific activity of 295.1 mCi/mmole and diluted by unlabelled UDP-glucose (Sigma). Zeaxanthin was provided as a gift by Hoffmann-LaRoche, Basel, Switzerland. Zeaxanthin was quantified based on $E^{1\%}_{1cm}$=2340 at 452 nm in acetone (Davies, 1976). Incubation was performed at 37°C for 5 h in the dark, and the reaction was stopped by the addition of 200 µl chloroform:methanol (2:1). Reaction products were identified by thin layer chromatography (TLC). To separate the products, 40 µl of the chloroform extract from each assay mixture was applied on a silica gel TLC plate (Whatman, Hillsboro, OR). Two solvent sytems were used. The TLC was first developed in Solvent System A {petroleum ether/ether/acetone (1:1:1)} to separate unreacted zeaxanthin from the products. After air drying the TLC plate for 30 min., a second solvent system, B {petroleum ether/ether/methanol (1:1:1)}, was used to separate zeaxanthin monoglucoside from diglucoside. The TLC plates were scanned for

radioactivity using a PhosphorImager™ 400 Series, and spots were quantified using ImageQuant™ v3.15 software supplied by Molecular Dynamics, San Diego, CA

Protein sequence comparisons

The deduced amino acid sequence of CrtX from E. herbicola was compared with the protein sequence data bases. Database searches were made using the FASTDB program (Bruttlag et al., 1990) of the Intelligenetics Suite of sequencing software. Databanks seached were PIR 28 (March 31, 1991) and SWISS-PROT 18 (May 1991). To identify common motifs or boxes of similarity, short overlapping sequences of 48 amino acid residues were used to probe the data bases. Sequences were aligned using the Intelligenetics Gene Alignment Program.

III. Results

A. Cloning of crtX and over-production of the CrtX protein

The plasmid construct, pAPUX, is outlined in figure 5-1. PCR-amplified Erwinia herbicola DNA containing crtX was inserted immediately beyond the T7 promoter and Shine Dalgarno sequence of the pET-3b expression vector. A Φ10 T7 termination signal is located immediately downstream of the crtX insert.

This pAPUX construct was maintained in E. coli DH5α cells, and the crtX gene successfully expressed in E. coli BL21(DE3) via induction with IPTG. The accumulation of a newly formed 45 kD protein was observed on a 10% SDS-polyacrylamide gel (Fig. 5-2). The calculated molecular weight based on the derived amino acid sequence of the protein is also 45 kD. The identity of this band as the crtX gene product was confirmed by N-terminal amino acid sequencing. An additional band of slightly lower molecular mass was observed only as a minor component in the crude extract, but as a major component in the supernatant fractions.

B. In vitro enzymatic activity of the enzyme, zeaxanthin glucosyltransferase

Extracts from plasmid- containing E. coli strains were assayed for in vitro glucosyltransferase activity by incubation with zeaxanthin and UDP-[^{14}C]glucose. When an extract from induced E. coli BL21(DE3) cells containing pAPUX was used, significant amounts of the yellow colored radiolabeled products, zeaxanthin

mono- and diglucosides, were observed. As a negative control, an extract from E. coli BL21(DE3) cells containing the pET-3b vector without crtX was incubated with zeaxanthin and UDP-[^{14}C]glucose in the in vitro assay, and no product was observed (data not shown). Thin layer chromatographic analysis of the products of this assay is shown in figure 5-3 (Lane 1). Solvent System A was used to separate unreacted zeaxanthin (Rf=0.8) from the products. Subsequent use of Solvent System B resulted in the resolution of the polar zeaxanthin mono- and diglucosides (Fig. 5-3, lane 1). An authentic sample of zeaxanthin as well as an E. herbicola lysate containing the zeaxanthin mono- and diglucosides were used as standards for the identification of products based on mobility (Fig. 5-3, lanes 3 and 4). The second substrate, UDP-[^{14}C]glucose, did not move from the origin and, being colorless, was detected only as a radio-labelled spot (lane 2). Phosphor-imaging of the TLC plate showed the two colored spots at Rf=1.0 and 0.3 in Solvent System B to be radioactive, thus indicating that [^{14}C]glucose was incorporated into both products.

The in vitro activity of the enzyme was assayed at 37°C over a pH range from 6.5 to 9.0, and maximum activity was found at pH 7.0 to 7.5. Similiarly, activity at pH 7.4 was assayed over a temperature range of 23° to 42°C, and found to be optimal at 32° to 37°C. The addition of larger amounts of enzyme extract did not increase the yield of glucosylated products, indicating that enzyme was not the limiting component. Also, higher concentrations of UDP-glucose beyond 1 mM did not increase the product yield. However, the addition of increasing amounts of zeaxanthin over the range of

3 to 50 nmoles/ 0.1 ml to the assay did result in a roughly proportional increase in products (Fig. 5-4A). This zeaxanthin dependence was observed despite the very limited solubility of zeaxanthin, which was determined by absorbance measurements to be approximately 12.6 ppm, or only 2.3 nmol per 0.1 ml reaction volume.

To determine an optimal time at which cells should be harvested after IPTG induction, crude enzyme extracts representing various induction times were assayed for activity. Fifty percent of the maximum observed enzyme activity was present in cells harvested after 0.5 h, and approximately 80% of maximum activity was reached in 1 h (Fig. 5-5).

To determine the intracellular localization of active glucosyltransferase, various cell fractions were assayed for activity. The glucosyltransferase activity was present in both the washed membrane fraction and the high speed supernatant fraction. In assays of the whole cell lysate as well as all supernatant fractions, the zeaxanthin diglucoside was produced in significantly higher amounts than the monoglucoside (Fig. 5-6). The specific enzymatic activity was higher in the crude extract than in the later fractions, indicating the loss of some activity during fractionation (Fig. 5-6).

C. Physical properties and sequence comparisons of the crtX gene product with other UDP-glucose binding proteins

A hydropathy plot of the derived amino acid sequence of E. herbicola CrtX was prepared according to the method of Kyte and Doolittle (1982), and is presented in figure 5-7. No unambiguous

membrane-spanning regions were identified using the criteria of Kyte and Doolittle (1982), however two regions were found which contain primarily hydrophobic residues. Both of these regions, highlighted with shaded bars (Fig. 5-7), contain stretches of amino acid residues which are of insufficient length to constitute trans-membrane helices.

The CrtX amino acid sequence has been aligned for comparison with the following: insect ecdysteroid UDP-glucosyltransferase (O'Reilly and Miller, 1989); maize UDP-glucose flavonoid glucosyl-transferase (Ralston et al., 1988); and human (Jackson et al., 1987), rat (MacKenzie, 1987) and mouse (Kimura and Owens, 1987) UDP-glucuronosyltransferases (Fig. 5-8). The three mammalian glucuronosyltransferases, H, R, and M, show good identity among themselves but not with the other proteins. Between the two species of Erwinia, there is 56% identity. The main region of homology present in all six sequences is the area marked by the shaded line in figure 5-8.

Figure 5-1. Construction of plasmid pAPUX. "Ap" represents the ampicillin resistance gene. crtX is the Erwinia herbicola zeaxanthin glucosyl transferase gene. Upstream of crtX gene at an appropriate distance is a T7 promoter and Shine Dalgarno sequence. Downstream of crtX is a strong T7 terminator.

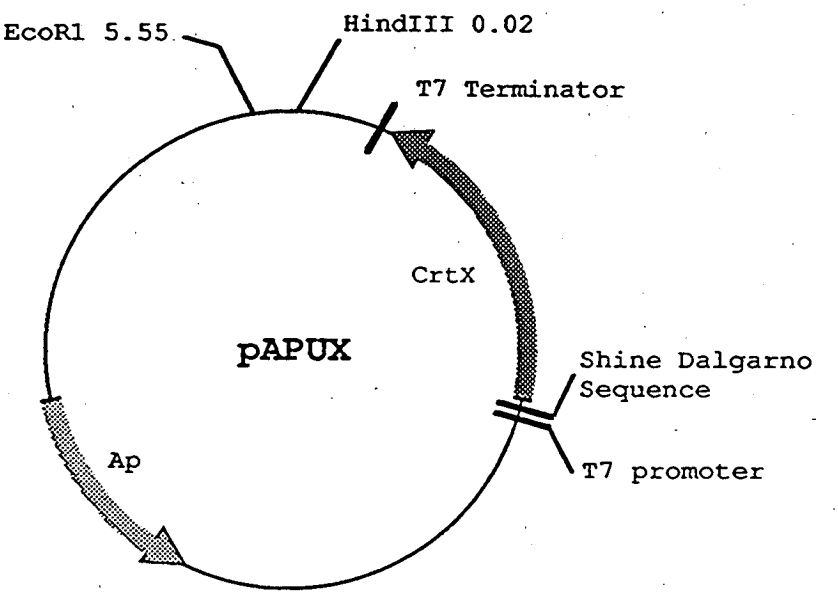

EcoR1 5.55
HindIII 0.02
T7 Terminator
CrtX
pAPUX
Shine Dalgarno
Sequence
T7 promoter
Ap

Figure 5-2. An SDS-10% PAGE gel showing expression of zeaxanthin glucosyltransferase. Lane 1, protein standard markers with MW 66 kD, 45 kD, 24 kD, and 18.4 kD; Lanes 2 and 3 represent total protein from uninduced and induced whole E. coli cells containing the plasmid pAPUX (note: Far less total protein was loaded in Lane 2.); Lane 4, 12.4 μg crude extract obtained by French Press; Lane 5, 4.2 μg inclusion body fraction ; Lane 6, 22.7 μg low speed (13,000x g) supernatant fraction; Lane 7, 30.2 μg high speed (100,000x g) supernatant fraction; Lane 8, 12.8 μg unwashed membrane fraction; and Lane 9, 3.4 μg washed membrane fraction.

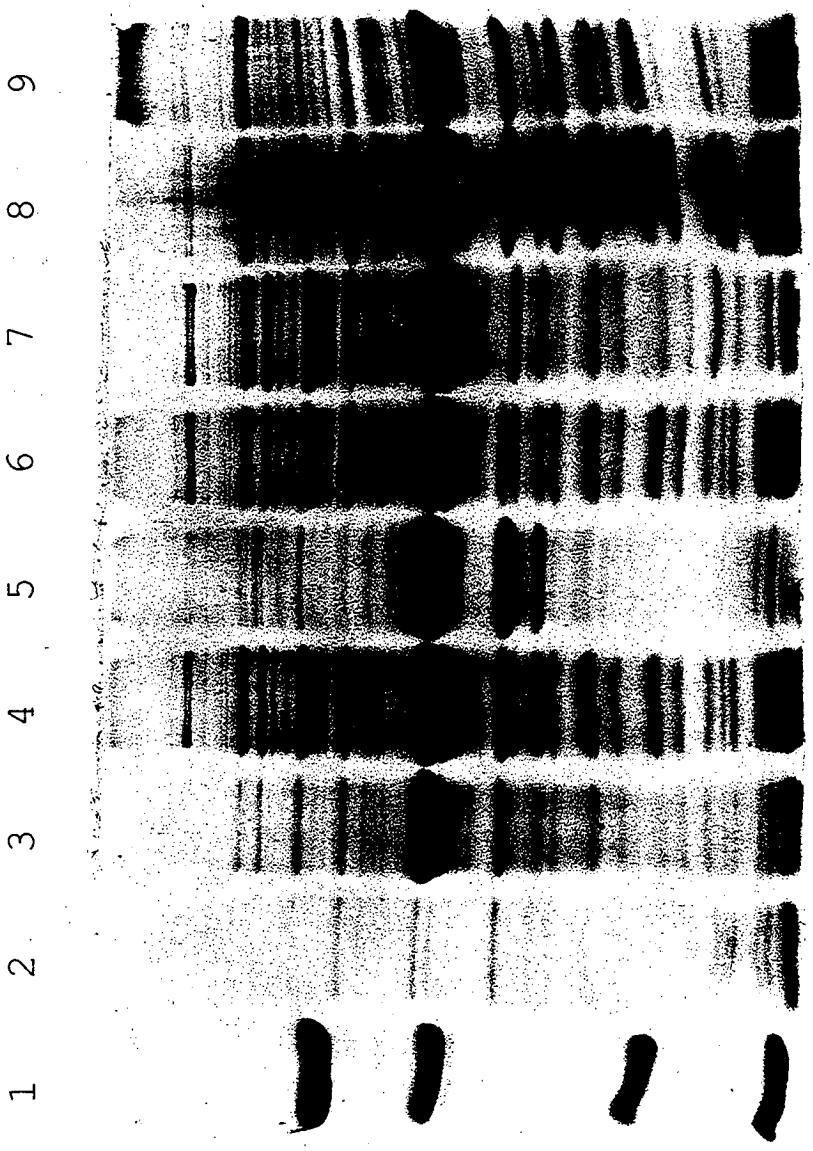

9
8
7
6
5
4
3
2
1

Figure 5-3. Zeaxanthin glucosyltransferase assay results showing formation of zeaxanthin mono- and diglucosides. (**A**) TLC plate, stained with iodine vapor: Lane 1, incubation of induced E. coli BL21(DE3) cell lysate containing pAPUX with zeaxanthin and UDP-[^{14}C]glucose showing formation of zeaxanthin mono- and diglucosides; lane 2, UDP-[^{14}C]glucose; Lane 3, zeaxanthin; Lane 4, E. coli(pPL376) carotenoid extract, containing mono- and diglucosides as major carotenoids. "Solvent A" and "Solvent B" indicate solvent fronts. (**B**) Phosphor-Image of the TLC plate showing: Lane 1, incorporation of ^{14}C label in both products, mono- and diglucosides; and Lane 2, standard UDP-[^{14}C]glucose remaining at the origin.

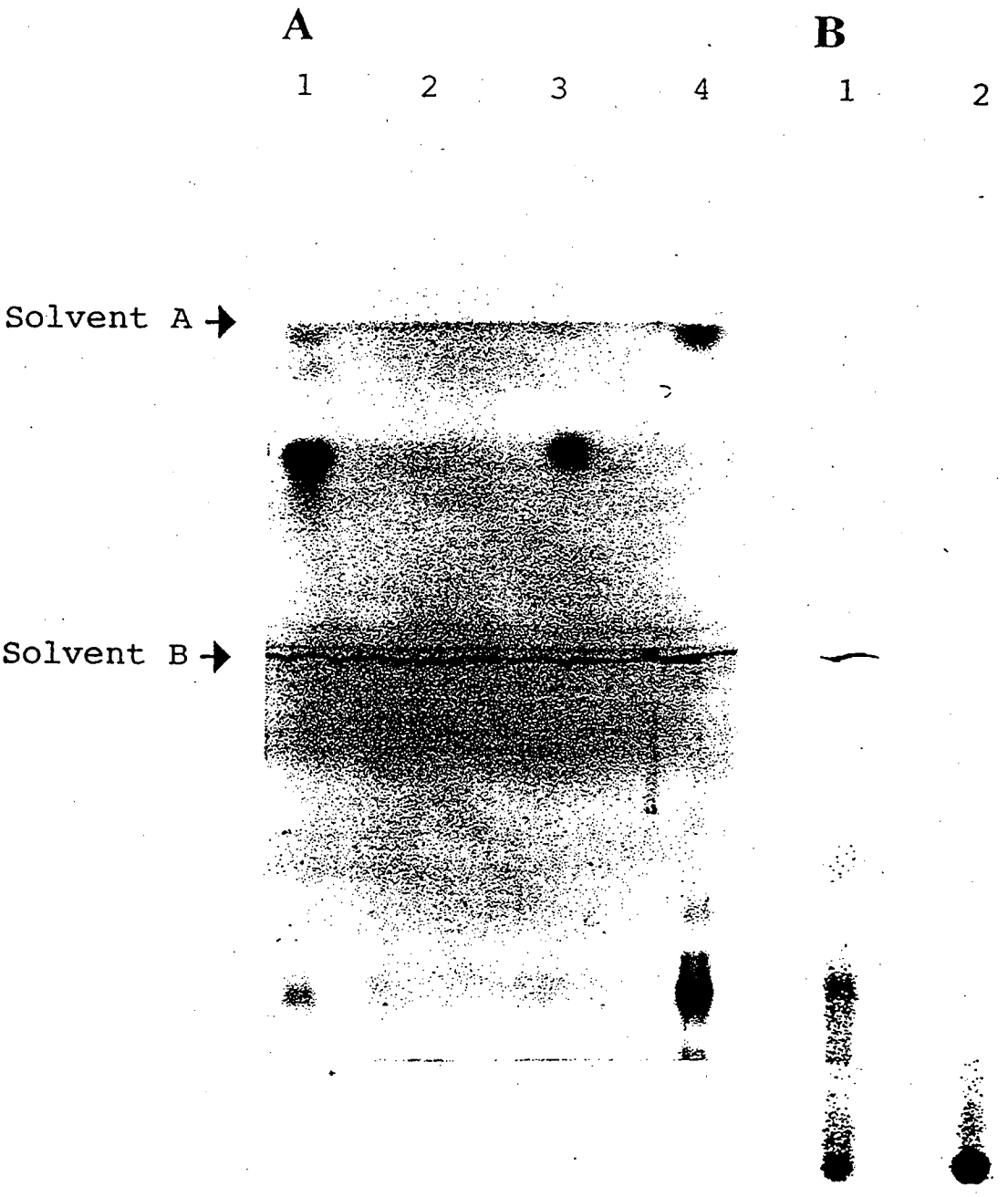

A
B
1
2
3
4
1
2
Solvent A
Solvent B

Figure 5-4. Plot of enzyme activity versus increasing amount of zeaxanthin. Enzyme activity was determined using 20 to 80 µl enzyme extract; zeaxanthin in 5 µl acetone (amount as indicated); 100 nmol UDP-glucose at a specific activity of 2.95 mCi/mmol ^{14}C, in 0.05 M Tris-HCl pH 7.5 containing 1mM βME to a volume of 100 µl. Yields of products were determined by the phosphorimaging of TLC plates (not shown) as described in figure 5-3, and comparison with a known ^{14}C standard.

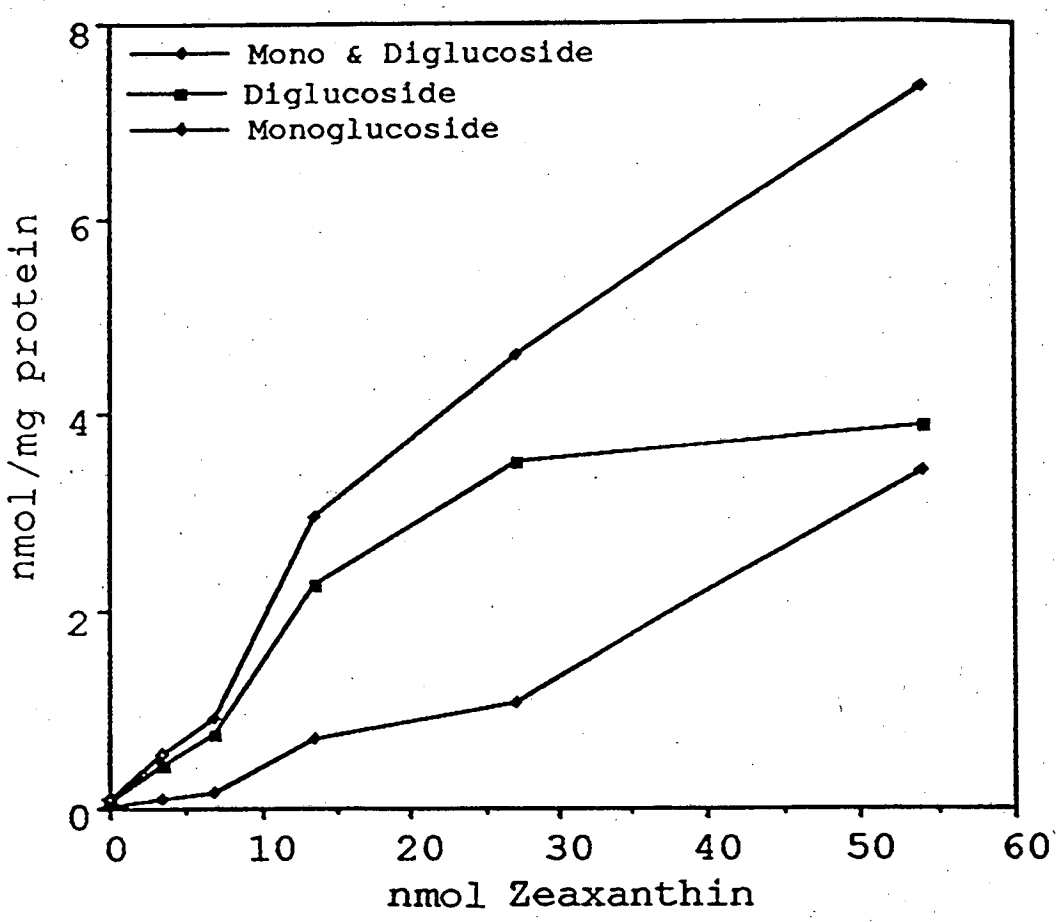

Mono & Diglucoside
Diglucoside
Monoglucoside
nmol/mg protein
0
2
4
6
8
0
10
20
30
40
50
60
nmol Zeaxanthin

Figure 5-5. Plot of enzyme activity versus IPTG induction time, after which cells were harvested. Enzyme activity was determined using 20 to 80 µl enzyme extract; 13.5 nmol zeaxanthin in 5 µl acetone; 100 nmol UDP-glucose at a specific activity of 2.95 mCi/mmol ^{14}C, in 0.05 M Tris-HCl pH 7.5 containing 1mM βME to a volume of 100 µl. Yields of products were determined by the phosphorimaging of TLC plates (not shown) as described in figure 5-3, and comparison with a known ^{14}C standard.

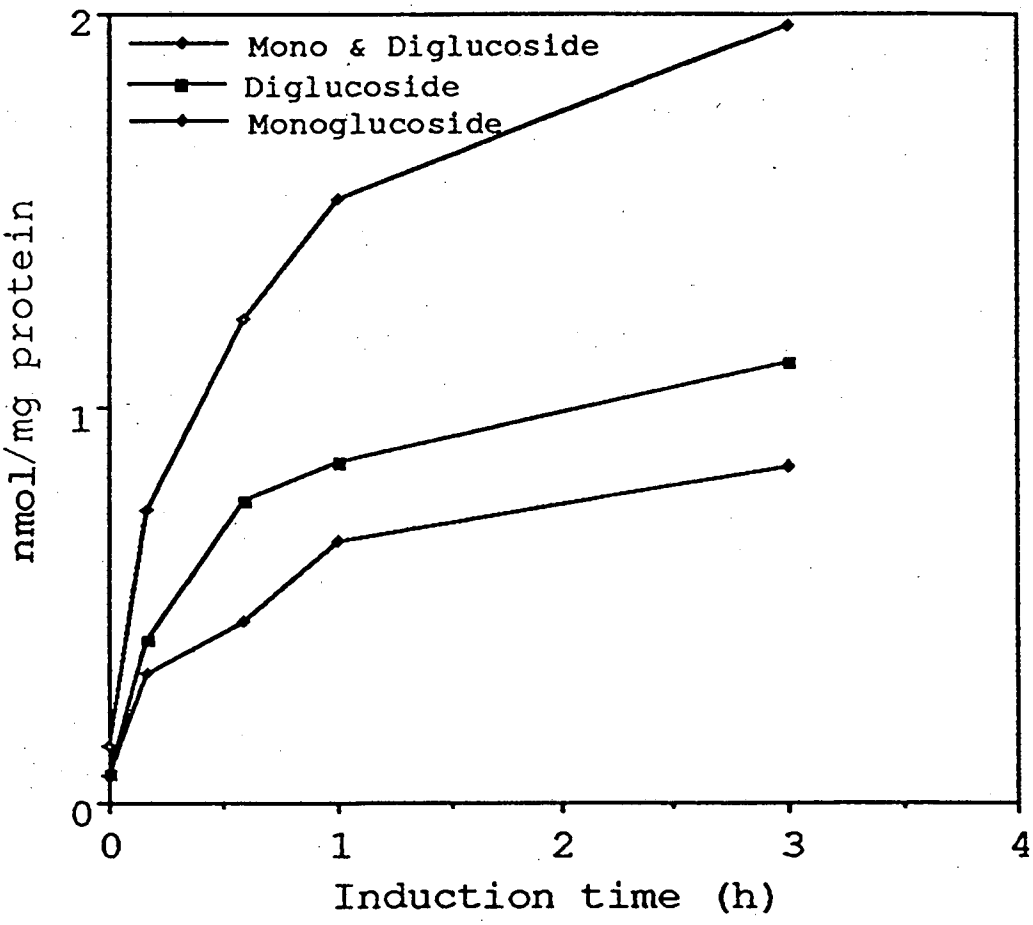

Mono & Diglucoside
Diglucoside
Monoglucoside
nmol/mg protein
0
1
2
0
1
2
3
4
Induction time (h)

Figure 5-6. Comparison of enzyme activity in cell lysate fractions 1) crude cell extract, 2) low speed supernatant, 3) high speed supernatant, 4) unwashed membranes and 5) washed membranes. Enzyme activity was determined using 20 to 80 µl enzyme extract; 13.5 nmol zeaxanthin in 5 µl acetone (or as indicated); 100 nmol UDP-glucose at a specific activity of 2.95 mCi/mmol ^{14}C, in 0.05 M Tris-HCl pH 7.5 containing 1mM βME to a volume of 100 µl. Yields of products were determined by the phosphorimaging of TLC plates (not shown) as described in figure 3, and comparison with a known ^{14}C standard.

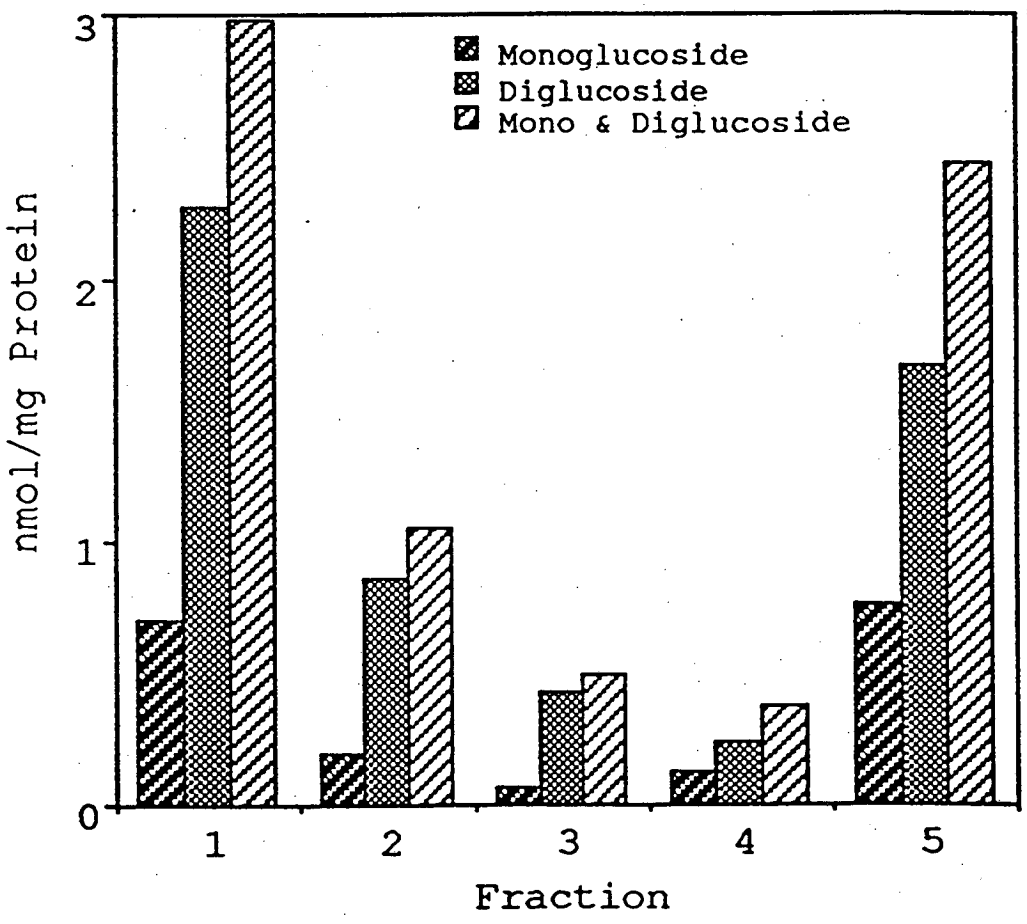

Monoglucoside
Diglucoside
Mono & Diglucoside
nmol/mg Protein
0
1
2
3
1
2
3
4
5
Fraction

Figure 5-7. Hydropathy profile for the deduced amino acid sequence of CrtX. The curve above the midline represents hydrophobicity and below represents hydrophilicity. The computer generated profile was determined using a moving window of 19 amino acid residues as recommended by Kyte and Doolittle (1982). A hydropathy index of +1.6 or greater of a segment containing at least 19 amino acids is indicative of a membrane spanning region. Shaded bars indicate stretches containing hydrophobic residues.

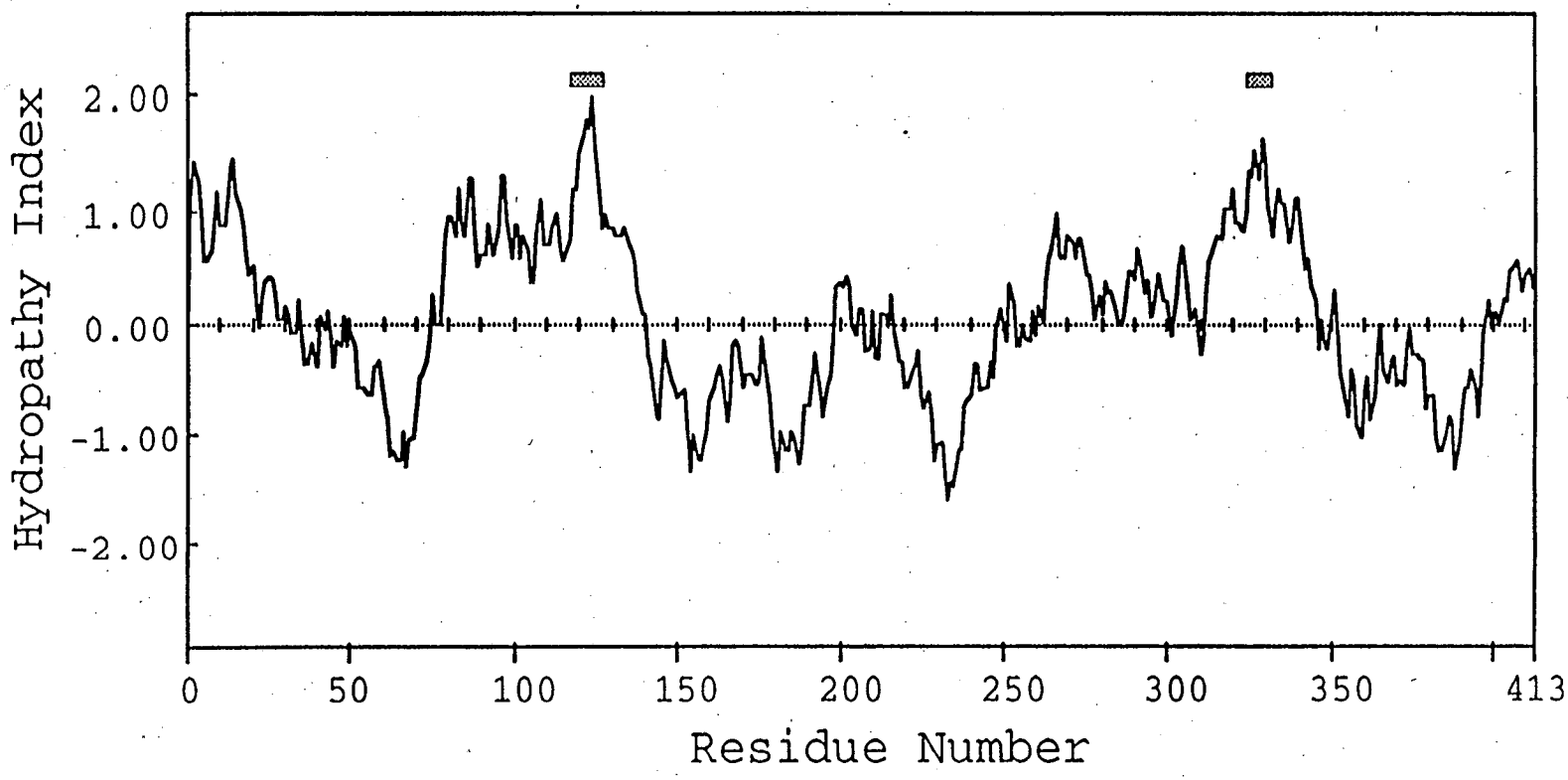
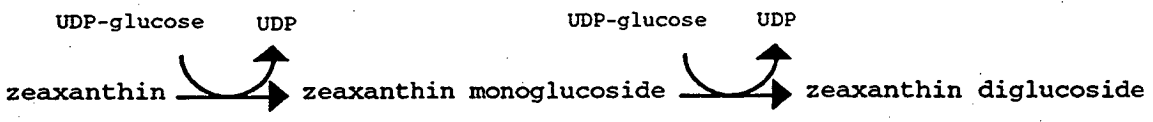

Hydropathy Index
2.00
1.00
0.00
-1.00
-2.00
0
50
100
150
200
250
300
350
413
Residue Number

Figure 5-8. Sequence similarities of the UDP-glucosyl-binding proteins. Protein sequences, sources and Swissprot accession numbers are as follows: I, Insect Ecdysteroid UDP-glucosyl-transferase, UDPE$NPVAC; Mz, Maize UDP-glucose flavonoid glucosyl-transferase, UFO1$MAIZE; crtX, E. herbicola zeaxanthin UDP-glucose transferase, GenBank Accession Number M87280; H, R and M represent, Human UDP1$HUMAN, Rat UDP4$RAT and Mouse UDP1$MOUSE, UDP-glucuronosyltransferases.

IV. Discussion

A. Zeaxanthin dependence of glucosyltransferase activity

The observation that [^{14}C]glucose is incorporated into two products is consistent with the carotenoid biosynthesis pathway for E. herbicola, in which a two step glucosylation of zeaxanthin was proposed based on the accumulation of the two pigments, zeaxanthin mono- and diglucosides (Hundle et al.,1991).



In a standard reaction mixture, 100 nmol of UDP-glucose and 13.5 nmol of zeaxanthin were incubated with 0.12 mg of protein from a crude enzyme extract in a volume of 0.1 ml. Increasing the amount of enzyme or UDP-glucose did not result in a significant increase in the amount of products. This indicates that, in this particular assay, neither the enzyme nor the UDP-glucose was the limiting factor. However, product yield was greatly increased by increasing the amount of zeaxanthin. Also, the zeaxanthin diglucoside was found in significantly higher amounts than the monoglucoside, despite the fact that the diglucoside is the product of the second condensation reaction. The monoglucoside, as an intermediate, never reaches very high levels compared with the original substrate, zeaxanthin. These results are consistent with a

model in which the binding of zeaxanthin to the enzyme is the rate limiting step for the reaction sequence. It is not known whether the monoglucoside intermediate must dissociate completely from the enzyme complex to undergo further reaction.

A significant increase in monoglucoside yield as a fraction of total products could be seen at high zeaxanthin levels (Fig. 5-4), as expected from a zeaxanthin limited reaction. Compared with the supernatants, membrane preparations produced relatively higher levels of the monoglucoside (Fig. 5-6). This increase in monoglucoside yield is most likely the result of the greater availability of zeaxanthin substrate, since this lipophilic compound is expected to partition into the membrane (non polar phase).

Despite the apparent solubility limitations, levels of zeaxanthin substrate as high as 50 nmol per 0.1 ml continued to produce an increase in the glucosylated products. The addition of zeaxanthin above the solubility limit results in the formation of a separate phase and does not cause an increase in the activity of zeaxanthin in the aqueous phase. Therefore, the glucosylation must occur in an oil phase, or more likely at an interface, such as a membrane, which does not depend on the substrate being water-solvated.

B. Enzyme activity is membrane associated

The glucosyltransferase activity was present in both the washed membrane fraction and the supernatant fraction. After fractionation, washed membranes exhibited over 5 fold higher activity per mg total protein when compared with the high speed

supernatant, indicating that the protein is membrane associated. A hydropathy plot of the enzyme's derived amino acid sequence does not show any clear membrane spanning regions. However, it is possible that one of the two hydrophobic domains (see Results), instead of being an internal region of the protein, could serve as an "anchor" to associate the protein with the membrane. Phospholipid dependence of another UDP-glucose requiring enzyme involved in maize sterol glucosylation has been demonstrated by Ury et al. (1989). Our conclusion is that CrtX is loosely associated with the membrane, and is partially dissociated from the membrane during fractionation.

Several factors could account for the enzyme activity remaining in the high-speed supernatant fraction. It is possible that the enzyme is still associated with small membrane fragments which are too small to be sedimented. Mowat and Arias (1970), using electron microscopy, have shown that the transfer of glucuronosyl-transferase activity from liver microsomes to the supernatant obtained by ultracentrifugation at 100,000x g, results merely from reduction in the size of membrane vesicles and not from free enzyme. Alternatively, it is possible that the enzyme is associated with the membrane in vivo, but that the membrane is not absolutely required for some basal level of activity. For maximum activity of the enzyme, membranes might be essential simply because the zeaxanthin substrate is a lipophilic compound located in the membrane.

We observe the formation of a second band on the SDS-PAGE protein gel upon centrifugation of the cell lysate to isolate

membranes (Fig. 5-2). It appears that this protein, which is only slightly smaller in molecular weight than the original overexpressed product, is the result of proteolytic cleavage of the original enzyme. This second protein is segregated in the supernatant and may be formed after some of the original enzyme is freed from the membrane, making it more accessible to proteases. The appearance of this second band is also concomitant with a net loss of enzyme activity observed upon fractionation. The combined enzyme activities of both the 100,000x g supernatant plus membrane fractions are less than that found in the original lysate (Fig. 5-6) on an activity per mg protein basis. Because this lower band is not found in any significant levels in cells lysed by boiling, it is possible that the release and proteolysis of the enzyme results from fractionation itself, and does not reflect a normal in vivo process.

C. A Proposed UDP-Glucose Binding Site Based on Amino Acid Homology

The deduced amino acid sequence of E. herbicola CrtX shows a negligible amount of overall homology with the other sequenced UDP-glucosyl- and glucuronosyltransferases (See results and Fig. 5-8). However, there is one distinct region of homology which is present in each of the six sequences (Fig. 5-8). Since the only common feature among these enzymes is that they require UDP-activated substrates, we conclude that the following sequence is most likely a portion of the UDP binding site:

[$QX_{13}TX_2GX_7LX_4PX_4PX_3DQX_4A$], where X represents a variable residue.

In summary, the enzyme, zeaxanthin glucosyltransferase from E. herbicola, can be overproduced and its activity demonstrated in vitro. The enzyme apparently is not an integral membrane protein, but is loosely attached to the membrane. We have proposed a UDP-binding sequence which is shared by the known transferases which require a UDP-activated second substrate. Further biochemical studies will be performed to establish the enzyme's kinetic properties and cofactor requirements.

Chapter 6:
General Summary and Future Directions

Carotenoid production occurs in all photosynthetic organisms as well as in many non-photosynthetic organisms. Despite the widespread distribution of carotenoid biosynthesis, we are now only beginning to understand both its biochemistry and the transcriptional regulation of photosynthetic genes. This work utilizes the two bacteria, Erwinia herbicola and Rhodobacter capsulatus as model systems suited for the study of enzyme biochemistry and transcriptional regulation, respectively. Using E. herbicola, the first reported in vitro overexpression of enzymatically active individual post-phytoene carotenoid biosynthesis genes is presented. Rhodobacter capsulatus, as a facultative phototroph with a 46 kb cluster of anaerobically induced photosynthetic genes, including all those required for carotenoid biosynthesis, was used for the study of transcriptional regulation.

In Chapter 2, the recurring palindrome sequence found upstream of many carotenoid and bacteriochlorophyll biosynthesis genes is confirmed as a regulatory sequence by the identification of a sequence-specific DNA binding protein. This protein and RegA, which has recently been found to regulate the puf and puh structural protein operons (Sganga et al., 1992), represent the first examples of trans-acting factors affecting transcription of the inducible R. capsulatus photosynthesis genes. Competition assays and mutational analyses indicate that the palindrome-binding

protein acts in a sequence-specific manner, and that DNA-protein complex formation is cooperatively enhanced by the presence of an upstream AT-rich DNA sequence acting in cis. Both the binding constant and dissociation rate are estimated.

Three lines of evidence indicate that the palindrome acts as a repressor of aerobic transcription. First, the binding of protein to the bchCXYZ DNA fragment in the mobility shift assay is approximately 2-3 times as strong when an aerobic cell lysate is used, compared to the corresponding culture shifted to anaerobic/ photosynthetic growth conditions, where transcription is reported to be 3-4 times greater (Armstrong, 1989b). Second, the placement of the palindrome relative to the σ^{70}- like promoter is much more like that of a repressor. Finally, the proposed role of the binding protein as a repressor under aerobic conditions is consistent with the results of a parallel in vivo study utilizing lacZ transcriptional fusions (Ma et al., 1992). Although mutations in the right, left, or both halves of the palindrome produced different basal levels of β-galactosidase expression, the differential effect of shifting to anaerobic growth was markedly reduced in each case in which the palindrome was disrupted.

More limited evidence indicates a possible second role for this factor as an activator under anaerobic/ photosynthetic conditions. Both the reduced expression of bchCXYZ lacZ fusions containing mutations in the right half of the palindrome (Ma et al., 1992) and the formation of an upper shifted band in a gel-shift assay, especially in the case of the highly induced crtEF operon (Chapter 2), point toward an activation mechanism. Although two crtEF

transcripts mapped earlier by Giuliano et al. (1988) appear to start just downstream of the perfect match of the consensus palindrome, the true promoter for the operon has not been conclusively identified.

The results of Chapter 2, when taken together with both the distribution of putative regulatory sites throughout the gene cluster and the transcriptional data of other researchers, lead us to a possible mechanism for transcriptional regulation in which the expression of these pigment biosynthesis genes is coordinated with that of the structural proteins of the photosynthetic apparatus, through the superoperonal organization of the gene cluster (reviewed by Wellington et al., 1991).

Whether the palindrome itself is sufficent to initiate transcription at crtEF, and whether other sequences also play a role, are subjects for further in vitro and in vivo mutational analyses (as described in Chapter 2 for the bchCXYZ operon). Another interesting region in which to apply these mutational and gel-shift assays is the sequence upstream of the puc operon which codes for the LH-II structural polypeptides. This operon contains another example of the same conserved palinrome and holds the additional advantages of being more highly expressed and heavily induced by a drop in oxygen tension (Youvan and Ismail, 1985; Armstrong, 1989b). Since LH-II is the most abundant pigment/ protein complex utilizing carotenoids under standard photosynthetic growth conditions, it would not be surprising to find similar regulatory mechanisms at work at this operon.

Finally, use of an enzymatic oxygen trap with in vitro enzyme

assays has been reported (Lam and Malkin, 1982), and this approach may provide a more rigorous method of comparing protein binding to the palindrome DNA in vitro under aerobic versus anaerobic conditions.

Chapter 3 focuses on the isolation of the transcription factor described in Chapter 2. Limited N-terminal amino acid sequence information was obtained, and this should permit the design of a redundant oligonucleotide probe to locate the corresponding gene. Over-expression of the gene could provide a better source of useful quantities of active protein. Analysis of transcripts from binding-factor deletion mutants should settle the questions of whether the palindrome can act as an activator of transcription and whether this sequence is absolutely essential for photosynthetic gene expression.

Unfortunately, isolation of the binding protein in an active form was complicated by protein denaturation and the fact that the optimum DNA binding site was large enough to be disturbed by crosslinking this DNA to a column. For future work, an alternate type of poly(bchC-200mer) DNA column should be constructed in which the active DNA concatemer is ligated in situ to compatible DNA previously crosslinked to the support. This would provide unhindered sites available for protein binding. Also, if denaturation of the factor at low total protein concentrations continues to be a problem, the addition of excess carrier protein (BSA, etc.) or DNA to all column wash solutions should be considered as ways of stabilizing the binding factor.

In Chapters 4 and 5, Erwinia herbicola, a non-photosynthetic plant pathogen, was selected for enzymology studies because of its

similarity to and compatibility with E. coli, and its lack of dependence on a specialized photosynthetic membrane or organelle. Use of this bacterium instead of more conventional photosynthetic sources resulted in the first reported overexpression of functionally active post-phytoene carotenoid biosynthesis enzymes.

Chapter 4 describes the separate overexpression in E. coli and in vitro analysis of two Erwinia enzymes involved in the main sequence of the carotenoid biosynthesis pathway, lycopene cyclase and β-carotene hydroxylase. In this study, each of the two enzymes could act separately and process exogenous substrates, indicating that neither the clustering of the carotenoid biosynthesis enzymes nor the channelling of intermediates was absolutely required, as had been proposed (Beyer et al., 1985; Candau et al., 1991). The levels of activity of both enzymes were found to be fairly low, particularly that of the lycopene cyclase. This result was not unexpected, considering the fact that both enzymes require extremely hydrophobic substrates which are introduced into a basically aqueous environment.

All-trans-lycopene was found to be the only acceptable in vitro substrate for this Erwinia lycopene cyclase, in contrast to the findings with Narcissus pseudonarcissus chromoplasts (Beyer et al., 1991), which converted only prolycopene or other cis-isomers of lycopene to β-carotene. The enhancement of β-carotene hydroxylase activity upon the addition of electron donors is consistent with a mixed-function oxygenase mechanism for carotene hydroxylation in which O_2 is first reduced to a more reactive species (e.g. a hydroperoxide).

Future strategies for increasing the activity of these two enzymes include: introducing the plasmid-borne genes back into Er. herbicola to use an Erwinia-based in vitro system to minimize any artifacts resulting from the E. coli system (particularly the segregation of the cyclase into inactive inclusion bodies); and the use of reconstituted liposomes from a cell-free source to remove other membrane-associated redox reactions from the assay (e.g. avoiding the membrane inhibition of hydroxylase activity while retaining a membrane site). In addition, this approach can also be applied to some of the other Erwinia carotenoid biosynthesis, such as crtI, the multi-step dehydrogenase.

One interesting practical application of these clustered Erwinia genes involves the possibility of introducing carotenoid biosynthesis into selected food crops to alleviate vitamin A deficiencies in populations dependent on a single major food source (e.g. rice). Besides the development of a better understanding of the biochemistry of the enzymes, this approach would require both blocking biosynthesis at crtZ because the xanthophylls are not biologically active, as well as the mobilization of this gene cluster on an appropriate vector.

Chapter 5 examines the overexpression and enzymology of functionally active zeaxanthin glucosyltransferase, an enzyme which carries out a more unusual transformation, converting a carotenoid into its more hydrophilic mono- and diglucoside derivatives. In addition, amino acid homology with other glucosyl and glucuronosyl transferases resulted in the identification of a putative binding site for the UDP-activated substrate. Although the

zeaxanthin glucosyl-transferase exhibited much more in vitro activity than either of the other two Erwinia enzymes, several interesting questions were raised by the results: How does the addition of zeaxaxnthin, well beyond its solubility limit, continue to increase the yield of products?; What is the true nature of the interaction of the enzyme with the membrane? Immuno-staining might be useful in determining if the normally expressed enzyme is associated with the membrane. Fractionation of the cell lysate in the presence of protease inhibitors is a possible way to determine if the partial loss of enzyme activity upon removal of the enzyme from the membrane is due to proteolysis.

A practical in vitro application of the active zeaxanthin glucosyltransferase is the production of carotenoids which are more soluble in water. This naturally occuring food-compatible pigment could be rendered even more water soluble if the transferase enzyme could be used to accept di- or polysaccharides, or if the normal diglucoside product is subsequently glycosylated.

In conclusion, the use of these two bacterial model system has furthered our general understanding of the biochemistry of carotenoid production, as well as the transcriptional regulation of inducible carotenoid biosynthesis genes.

References

Adams CW, Forrest ME, Cohen SN, and Beatty JT (1989) Structural and functional analysis of transcriptional control of the Rhodobacter capsulatus puf operon. J. Bacteriol. 171: 473-482

Alberti M and Hearst JE (1991) Sequence of the 46 kb photosynthesis gene cluster of Rhodobacter capsulatus. Photochem. and Photobiol. 53: 122S

Aragon CMG, Murillo FJ, De La Guardia MD, and Cerda-Olmedo E (1976) An enzyme complex for the dehydrogenation of phytoene in Phycomyces. Eur. J. Biochem. 63: 71-75

Armstrong GA, Alberti M, Leach F, and Hearst JE (1989) Nucleotide sequence, organization, and nature of protein products of the carotenoid biosynthesis gene cluster of Rhodobacter capsulatus. Mol. Gen. Genet. 216: 254-268

Armstrong GA (1989b) A molecular genetic analysis of carotenoid biosynthesis and the effects of carotenoid mutations on other photosynthetic genes in Rhodobacter capsulatus. Ph.D. Dissertation, University of California, Berkeley

Armstrong GA, Schmidt A, Sandmann G, Hearst JE (1990a) Genetic

and biochemical characterization of carotenoid biosynthesis mutants of Rhodobacter capsulatus. Journal of Biological Chemistry, 265: 8329-38

Armstrong GA, Alberti M, and Hearst JE (1990b) Conserved enzymes mediate the early reactions of carotenoid biosynthesis in nonphotosynthetic and photosynthetic prokaryotes. Proc. Natl. Acad. Sci. USA 87: 9975-9979

Bar-Peled M, Lewinson E, Fluhr R, and Gressel J (1991) UDP-rhamnose: flavanone-7-O-glucoside-2''-O-rhamnosyltransferase. Pur-ification and characterization of an enzyme catalyzing the production of bitter compounds in citrus. J. Biol. Chem. 266: 20953-20959

Bauer CE, Young DA, and Marrs BL (1988) Analysis of the Rhodobacter capsulatus puf operon. J. Biol. Chem. 263: 4820-4827

Bauer CE, Buggty JJ, Yang Z, and Marrs BL (1991) The superoperonal organization of genes for pigment biosynthesis and reaction center proteins is a conserved feature in Rhodobacter capsulatus: analysis of overlapping bchB and puhA transcripts. Mol. Gen. Genet. 228: 433-444

Baumann EA and Hand R, (1982) Phosphorylation and dephos-phorylation alter the structure of D2 hybrid T antigen. J. Virol. 44: 78-87

Beyer P, Weiss G, and Kleinig H (1985) Solubilization and reconstitution of the membrane-bound carotenogenic enzymes from

daffodil chromoplasts. Eur. J. Biochem. 153: 341-346

Beyer P, Mayer M, and Kleinig H (1989) Molecular oxygen and the state of geometric isomerism of intermediates are essential in the carotene desaturation and cyclization reactions in daffodil chromoplasts. Eur. J. Biochem. 184: 141-150

Beyer P, Kroncke U, Nievelstein V., and Kleinig H. (1991) On the mechanism of the lycopene isomerase/cyclase reaction in Narcissus pseudonarcissus L. chromoplasts. J. Biol. Chem. 266: 17072-17078

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72: 248-254

Bramley PM and MacKenzie A (1988) The regulation of carotenoid biosynthesis, in: Current Topics in Cellular Regulation 29: 291-343, Horecker BL and Stadtman ER (eds), Academic Press, London, .

Bramley PM and Sandmann G (1985) In vitro and in vivo biosynthesis of xanthophylls by the cyanobacterium Aphanocapsa. Phytochemistry 24: 2919-2922

Britton G (1983) The Biochemistry of Natural Pigments, pp1-92. Cambridge University Press, Cambridge

Brutlag, DL, Deutricourt JP, Maulik S, and Relph J (1990) Improved sensitivity of biological sequence database searches. Comp. App. Bio. Sci. 6: 237-245

Buck M, Miller S, Drummond M, and Dixon R (1986) Upstream

activator sequences are present in the promoters of nitrogen fixation genes. Nature 320: 374-378

Burke DH, Alberti, and Hearst JE (1992)..Chlorophyll Fe proteins and other chlorophyll biosynthesis genes from Rhodobacter capsulatus to higher plants. J. Bacteriol., submitted

Candau R., Bejarano ER, and Cerda-Olmedo E (1991) In vivo channelling of substrates in an enzyme aggregate for β-carotene biosynthesis. Proc. Natl. Acad. Sci. USA 88: 4936-4940

Carattoli A, Romano N, Ballario P, Morelli G, and Macino G (1991) The Neurospora crassa carotenoid biosynthetic gene (Albino-3) reveals highly conserved regions among prenyltransferases. J. Biol. Chem. 266: 5854

Clark WG, Davidson E, and Marrs BL (1984) Variations of levels of mRNA coding for antenna and reaction center polypeptides in Rhodopseudomonas capsulata in response to changes in oxygen concentration. J. Bacteriol. 157: 945-948

Cleveland DW, Fischer SG, Kirschner MW, and Laemmli UK (1977) Peptide mapping by limited proteolysis in sodium dodecyl sulfate and analysis by gel electrophoresis. J. Biol. Chem. 252: 1102-1106

Cogdell RJ and Frank HA (1987) How carotenoids function in photosynthetic bacteria. Biochem. Biophys. Acta 895: 63-79

Cogdell RJ, Hipkins MF, MacDonald W, and Truscott TG (1981) Energy transfer between the carotenoid and the bacteriochlorophyll

within the B-800-850 light harvesting pigment-protein complex of Rhodopseudomonas sphaeroides. Biochim. Biophys. Acta 634: 191-202

Cohen-Bazire G, Sistrom WR, and Stanier RY (1957) Kinetic studies of pigment synthesis by non-sulfur purple bacteria. J. Cell. Comp. Physiol. 49: 25-68

Colladovides J, Magasanik B, and Gralla JD, (1991) Control site location and transcriptional regulation in Escherichia coli. Microbiol. Rev. 55: 371-394

Cook DN, Armstrong GA, and Hearst JE, (1989) Induction of anaerobic gene expression in Rhodobacter capsulatus is not accompanied by a local change in supercoiling as measured by a novel assay. J. Bacteriol. 171: 4836-4843

Davies BH (1976) Carotenoids. in: Chemistry and Biochemistry of Plant Pigments, Second Edition, vol.2 pp 38-165, Goodwin TW (ed), Academic Press, London

Deisenhofer J and Michel H (1988) in: The Photosynthetic Bacterial Reaction Center, Breton J and Vermeglio A (eds), NATO ASI Series A: Life Sciences, pp 1-4, Plenum Press, New York

Dogbo O, Laferriere A, d'Harlingue A, and Camara B (1988) Carotenoid biosynthesis: isolation and characterization of a bifunctional enzyme catalyzing the synthesis of phytoene. Proc. Natl. Acad. Sci. USA 85: 7054-7058

Drews G and Oelze J (1981) Organization and differentiation of membranes of phototrophic bacteria, in: Advances in Microbial Physiology, vol.22 pp 1-92, Rose AH, Morris JG (eds)

Drews G, Dierstein R, and Schumacher A (1976) Genetic transfer of the capacity to form bacteriochlorophyll-protein complexes in Rhodopseudomonas capsulata. FEBS Lett. 68: 132-136

Dutton GJ (1980) Glucuronidation of Drugs and Other Compounds, pp 1-78, CRC Press, Boca Raton

Fox BG, Froland WA, Dege JE, and Lipscomb, JD (1989) Methane monooxygenase from Methylosinus trichosporium OB3b. J. Biol. Chem. 264: 10023-10033

Gicquel-Sanzey B and Cossart P, (1982) Homologies between different prokaryotic DNA-binding regulatory proteins and between their sites of action. EMBO J. 1: 591-595

Giuliano G, Pollock D, and Scolnik PA (1986) The crtI gene mediates the conversion of phytoene into colored carotenoids in Rhodopseudomonas capsulata. J. Biol. Chem. 261: 12925-12929

Giuliano G, Pollock D, Stapp H, and Scolnik PA (1988) A genetic-physical map of the Rhodobacter capsulatus carotenoid biosynthesis gene cluster. Mol. Gen. Genet. 213: 78-83

Goodwin TW (1958) Studies in carotenogenesis: The changes in carotenoid and chlorophyll pigments in the leaves of deciduous trees during autumn necrosis. Biochem. J. 68: 503-511

Goodwin TW (1980) The Biochemistry of Carotenoids, vol.1: Plants, Chapman and Hall, Ltd., London

Ho SN, Hunt HD, Horton RM, Pullen JK, and Pease LR (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene 77: 51-59

Hundle BS, Beyer P, Kleinig H, Englert G, and Hearst JE (1991) Carotenoids of Erwinia herbicola and an Escherichia coli HB101 strain carrying the Erwinia herbicola carotenoid gene cluster. Photochem. Photobiol. 54: 89-93

Hundle BS, Alberti A, Beyer P, Neivelstein V, Kleinig H, Englert G, Armstrong GA, Lee KA, Burke D (1992a) DNA sequence and assignment of gene products to the carotenoid biosynthesis pathway of an Erwinia herbicola gene cluster expressed in Escherichia coli., in preparation

Hundle BS, O'Brien DA, Alberti M, Beyer P, and Hearst JE (1992b) Functional expression of zeaxanthin glucosyltransferase from Erwinia herbicola and a proposed uridine diphosphate binding site. Proc. Natl. Acad. Sci. USA 89: 9321-9325

Jackson MR, McCarthy LR, Harding D, Wilson S, Coughtrie MWH, and Burchell B (1987) Cloning of human liver UDP-glucuronosyl-transferase cDNA. Biochem. J. 242: 581-588

Joyard J, Block MA, and Dounce R (1991) Molecular aspects of plastid envelope biochemistry. Eur. J. Biochem. 199: 489- 509

Kadonaga JT and Tjian R (1986) Affinity purification of sequence-specific DNA binding proteins. Proc. Natl. Acad. Sci. USA 83: 5889-5893

Keener J and Kustu S (1988) Protein kinase and phosphoprotein phosphatase activities of nitrogen regulatory proteins NTRB and NTRC of enteric bacteria: Roles of the conserved amino-terminal domain of NTRC. Proc. Natl. Acad. Sci. USA 85: 4976-4980

Kimura T and Owens IS (1987) Mouse UDP-glucuronosyltransferase: cDNA and complete amino acid sequence and regulation. J. Biochem 168: 515-521

Kleinig H (1989) The role of plastids in isoprenoid biosynthesis. Annu. Rev. Plant Physiol. Plant Mol. Biol. 40: 39-59

Klug G, Kaufmann N, and Drews G (1985) Gene expression of pigment-binding proteins of the bacterial photosynthetic apparatus: transcription and assembly in the membrane of Rhodopseudomonas capsulata. Proc. Natl. Acad. Sci. USA 82: 6485-6489

Krinsky NI (1979) Carotenoid protection against oxidation. Pure Appl. Chem. 51: 649-660

Kushwaha SC, Subbarayan C, Beeler DA, and Porter JW (1969) The conversion of lycopene-15,15'-^{3}H to cyclic carotenes by soluble extracts of higher plant plastids. J. Biol. Chem. 244: 3635-3642

Kushwaha SC, Kates M, and Porter JW (1976) Enzymatic synthesis of C_{40} carotenes by a cell-free preparation from Halobacterium

cutirubrum. Can. J. Biochem. 54: 816-823

Kyte J and Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157: 105-132

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680-685

Lam E and Malkin R (1982) Ferredoxin-mediated reduction of cytochrome b-563 in a chloroplast cytochrome b-563/f complex. FEBS Lett. 141: 98-101

Lee JK and Kaplan S (1992) Cis-acting regulatory elements involved in oxygen and light control of puc. J. Bacteriol. 174: 1146-1157

Loh EY, Elliot JF, Cwirla S, Lanier LL, and Davis M (1989) Polymerase chain reaction with single-sided specificity: analysis of T cell receptor delta chain. Science 243: 217-220

Ma D, Cook DN, O'Brien DA, and Hearst JE (1992) Analysis of the promoter and regulatory sequences of an oxygen- regulated bch promoter in Rhodobacter capsulatus by site-directed mutagenesis. J. Bacteriol., submitted

Mackenzie PI (1987) Rat liver UDP-glucuronosyltransferase. Identification of cDNA's encoding two enzymes which glucuronidate testerone, dihydrotesterone, and β-estradiol. J. Biol. Chem. 262: 9744-9749

Marrs B (1981) Mobilization of the genes for photosynthesis from

Rhodopseudomonas capsulata by a promiscuous plasmid. J. Bacteriol. 146: 1003-1012

Math SK, Hearst JE, and Poulter CD (1992) The crtE gene in Erwinia herbicola encodes geranylgeranyl diphosphate synthase. Proc. Natl. Acad. Sci. USA 89: 6761-6764

Mathews-Roth MM (1987) Photoprotection by carotenoids. Federation Proc. 46: 1890-1893

Mayer MP, Bartlett DL, Beyer P, and Kleinig H (1989) The in vitro mode of action of bleaching herbicides on the desaturation of 15-cis-phytoene and 15-cis-ζ-carotene in isolated daffodil chromoplasts. Pestic. Biochem. Physiol. 34: 111-117

McDermott JBC, Brown DJ, Britton G, and Goodwin T W (1974) Alternative path of zeaxanthin biosynthesis in a flavobacterium species. Experiments with nicotine as an inhibitor. Biochem. J. 144: 231-243

Misawa N, Nakagawa M, Kobayashi K, Yamano S, Izawa Y, Nakamura K, and Harashima K (1990) Elucidation of the Erwinia uredovora carotenoid biosynthetic pathway by functional analysis of gene products expressed in Escherichia coli. J. Bacteriol. 172: 6704-6712

Morrissey JH (1981) Silver stain for proteins in polyacrylamide gels: A modified procedure with enhanced uniformed sensitivity. Anal. Biochem. 117: 307-310

Mowat AP, and Arias IM (1970) Partial purification of hepatic UDP-

glucuronosyltransferase. Studies of some its properties. Biochim. Biophys. Acta. 212: 65-78

Nelson HCM, Finch JT, Luisi BF, and Klug A (1987) The structure of an oligo(dA)-oligo(dT) tract and its biological implications. Nature 330: 221-226

Nelson MA, Morelli G, Carattoli A, Romano N, and Macino G (1989) Molecular cloning of a Neurospora crassa carotenoid biosynthetic gene (albino-3) regulated by blue light and the products of the white collar genes. Mol. Cell. Biol. 9:-1271-1276

Nybraaten G and Liaaen-Jensen S (1974) Bacterial carotenoids. XLIV. Zeaxanthin mono- and dirhamnoside. Acta Chem. Scand. B 28: 1219-1224

O'Reilly DR and Miller LK (1990) Regulation of expression of a baculovirus ecdysteroid UDP-glucosyltransferase gene. J. Virol. 64: 1321-1328

Orser C, Staskawicz BL, Panopoulous N, Dahlbeck D, and Lindow S (1985) Cloning and expression of bacterial ice nucleation genes in Escherichia coli. J. Bacteriol. 164: 359-366

Perry KL, Simonitch TA, Harrison-Lavoie KJ, and Liu ST (1986) Cloning and regulation of Erwinia herbicola pigment genes. J. Bacteriol. 168: 607-612

Pfander VH and Hodler M (1974) Carotenoid glycosides. 1: Partial synthesis and characterization of zeaxanthin mono- and di-

glucosides. Helvetica Chim. Acta 57: 1641-1651

Porter JW and Lincoln RE (1950) Lycopersicon selections containing high contents of carotenes and colorless polyenes: II. The mechanism of carotene biosynthesis. Arch. Biochem. Biophys. 27: 390-403

Poulter CD and Rilling HC (1983) Prenyl transferases and isomerases and isomerase. in: Biosynthesis of Isoprenoid Compounds, vol.1, pp 161-224, Porter JW and Spurgeon SL (eds), John Wiley and Sons, New York

Qureshi N and Porter JW (1983) Conversion of acetyl CoA to isopentenyl pyrophosphate. in: Biosynthesis of Isoprenoid Compounds, vol 1, pp 47-94, Porter JW and Spurgeon SL (eds), John Wiley and Sons, New York

Ralston EJ, English JJ, and Dooner HK (1988) Sequence of three bronz alleles of maize and correlation with the genetic fine structure. Genetics 119: 185-197

Riggs MG and McLachlan A (1986) A simplified plasmid mini-prep protocol. Bio-Techniques 4: 312-313

Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB and Erlich HA (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239: 487-491

Sambrook J, Frisch EF, and Maniatis T (1989) Molecular Cloning: A Laboratory Manual., Cold Spring Harbor Laboratory Press, 1989,

Cold Spring Harbor, NY

Sandmann G and Boger P (1989) Inhibition of carotenoid biosynthesis by herbicides. in: Target Sites of Herbicide Action, pp 25-44, Boger P and Sandmann G, (eds), CRC Press, Boca Raton

Sandmann G and Bramley P (1985) Carotenoid biosynthesis by Aphanocapsa homogenates coupled to a phytoene-generating system from Phycomyces blakesleeanus. Planta 164: 259-263

Sandmann G and Kowalczyk S (1989) In vitro carotenogenesis and characterization of phytoene desaturase reaction in Anacystis. Biochem. Biophys. Res. Commun. 163: 916-921

Schmidt K (1980) A comparative study on the composition of chlorosomes [chlorobium vesicles] and cytoplasmic membranes from Chloroflexus aurantiacus strain Ok-70-fl and Chlorobium limicola f thiosulfatophilum strain 6230. Arch. Microbiol. 124: 21-31

Schneour EA (1962) Carotenoid pigment conversion in Rhodopseudo-monas sphaeroides. Biochem. Biophys. Acta 62: 534-540

Schultz JE and Weaver PF (1982) Fermentation and anaerobic respiration by Rhodospirillum rubrum and Rhodopseudomonas capsulata. J. Bacteriol. 149: 181-190

Scolnik PA, Walker MA, and Marrs BL (1980) Biosynthesis of carotenoids derived from neurosporene in Rhodopseudomonas

capsulata. J. Biol. Chem. 255: 2427-2432

Sganga MW and Bauer CE (1992) Regulatory factors controlling photosynthetic reaction center and light harvesting gene expression in Rhodobacter capsulatus. Cell 68: 945-954

Siefermann D and Yamamoto HY (1975) Light induced de-epoxidation of violaxanthin in lettuce chloroplasts IV. The effect of electron transport conditions on violaxanthin availability. Biochim. Biophys. Acta 387: 149-158

Siefermann-Harms D (1984) Carotenoids in photosynthesis: I. location in photosynthetic membranes and light-harvesting function. Biochim. Biophys. Acta 811: 325-355

Simmons DT, Chou W, and Rodgers K (1986) Phosphorylation downregulates the DNA-binding activity of simian virus 40 T antigen. J. Virol. 60: 888-894

Starr MP (1981) The genus Erwinia, in: The Prokaryotes, pp1260-1271, Starr MP, Stolp H, Truper G, Balows A, and Schlegel HG (eds), Springer Verlag, Heidelberg

Stock JB, Ninfa AJ, Stock AM (1989) Protein phosphorylation and regulation of adaptive responses in bacteria. Microbiol. Rev. 53: 450-90

Straub, O. (1987) in: Key to Carotenoids, Pfander H, Gerspacher M, Rychener M, and Schwabe R (eds), Birkhäuser Verlag, Basel

Studier WF, Rosenberg AH, Dunn JJ, and Dubendorff JW (1990) Use

of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymology 185: 60-89

Taremi SS and Marrs BL (1990) Regulation of Gene Expression by Oxygen: Phototrophic bacteria. Proceedings of the 41st Mosbach Colloquium, in: The Molecular Basis of Bacterial Metabolism, Springer Verlag, Heidelberg

Taylor DP, Cohen SN, Clark WG, and Marrs BL (1983) Alignment of the genetic and restriction maps of the photosynthesis region of the Rhodopseudomonas capsulata chromosome by a conjugation-mediated marker rescue technique. J. Bacteriol. 154: 580-590

Tuveson RW, Larson RA, and Kagan J (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

Ury A, Benveniste P, and Bouvier-Nave P (1989) Phospholipid-dependence of plant UDP-glucose sterol β-D-glucosyltransferase. IV. Reconstitution into small unilamellar vesicles. Plant. Physiol. 91: 567-573

Van der Rest M and Gingras G (1974) Pigment complement of the photosynthetic reaction center isolated from Rhodospirillum rubrum. J. Biol. Chem. 249: 6446-6455

Walsh C (1976) Enzymatic Reaction Mechanisms pp 358-522, WH Freeman and Co., San Francisco

Wellington CL, and Beatty JT (1989) Promoter mapping and nucleotide sequence of the bchC bacteriochlorophyll biosynthesis gene from Rhodobacter capsulatus. Gene 83: 251-261

Wellington CL, Taggart AKP, and Beatty JT (1991) Functional significance of overlapping transcripts of crtEF, bchCA, and puf photosynthesis gene operons in Rhodobacter capsulatus. J. Bacteriol. 173: 2954-2961

Wellington CL, Bauer CE, Beatty JT (1992) Photosynthesis gene superoperons in purple nonsulfur bacteria- The tip of the iceberg? Can. J. Microbiol. 38: 20-27

Wilson AC, Ochman H, and Prager EM (1987) Molecular time scale for evolution. Trends Genet. 3: 241-247

Yamamoto HY, Chichester CO, and Nakayama TO (1962) Biosynthetic origin of oxygen in the leaf xanthophylls. Arch. Biochem. Biophys. 96: 645-649

Yeates TO, Komiya H, Chirino A, Rees DC, Allen JP, and Feher G (1988) Structure of the reaction center from Rhodobacter sphaeroides R-26 and 2.4.1: protein cofactor (bacteriochlorophyll, bacteriopheophytin, and carotenoid) interactions. Proc. Natl. Acad. Sci. USA 85: 7993-7997

Yen HC and Marrs BL (1976) Map of genes for carotenoid and bacteriochlorophyll biosynthesis in Rhodopseudomonas capsulata. J. Bacteriol. 126: 619-629

Yen HC and Marrs BL (1977) Growth of Rhodopseudomonas capsulata under anaerobic dark conditions with dimethylsulfoxide. Arch. Biochem. Biophys. 181: 411-418

Young DA, Bauer CE, Williams JC, and Marrs BL (1989) Genetic evidence for superoperonal organization of genes for photosynthetic pigments and pigment-binding proteins in Rhodobacter capsulatus. Mol. Gen. Genet. 218: 1-12

Youvan DC and Ismail S (1985) Light-harvesting II (B800-B850 complex) structural genes from Rhodopseudomonas capsulata. Proc. Natl. Acad. Sci. USA 82: 58-62

Zhu YS and Hearst JE (1986) Regulation of the expression of genes for light-harvesting antenna proteins LH-I and LH-II; reaction center polypeptides RC-L, RC-M, and RC-H; and enzymes for bacterio-chlorophyll and carotenoid biosynthesis in Rhodobacter capsulatus by light and oxygen. Proc. Natl. Acad. Sci. USA 83: 7613-7617

Zsebo KM and Hearst JE (1984) Genetic-physical mapping of a photosynthetic gene cluster from R. capsulata. Cell 37: 937-947

LAWRENCE BERKELEY LABORATORY
TECHNICAL INFORMATION DEPARTMENT
1 CYCLOTRON ROAD
BERKELEY, CALIFORNIA 94720