Feedback and Molecular Interactions in the Process of Light-Induced Carotenogenesis in *Myxococcus xanthus.*

David E. Whitworth B.A. (Hons.) oxon.

A Thesis submitted for the degree of Doctor of Philosophy in the University of Warwick.

Department of Biological Sciences
March 1999

Table of Contents

List of Figures

List of Tables

Acknowledgements

Here we go then, the only opportunity in this thesis where I get to say things as me rather than as microbiologist David Whitworth BA. First and foremost I would like to thank my supervisor Dave Hodgson without whom this thesis would be a pamphlet. Now comes the obligatory but tedious bit where I mention the other people in the lab and try to embarass them or be generally obnoxious to them. Okay, thanks go to: Sax for discussion, sharing the PhD experience from day one and not being too Welsh. Dr. Ges Owenson for not being too Scottish. Chris Hemingwaj, the ginger badger, for his 'different' perception of reality and Pete Watson the goober king with his protege the monkey boy. Thanks also to my lab ancestors Andy Berry and Doug Browning from whom I learned things. To everyone else who passed through micro II during my sojourn including the Salmond, Mann, and Scanlan groups en masse. Oh, nearly forgot Nazrin. And I was trying so hard too. Thanks also go to Dr. N. Mann for an introduction to the 'pleasures' of gapped PSI-BLAST and other molecular biology predictive tools.

Now must mention my parents and Anna for their support and love. At the times when everything around me was failing and feelings of utter despair crept in, I couldn't have done it without them. Anna must be thanked especially as she put up with a lot during the last three years. I hope it was worth it.

PS. This thesis contains in parts, a great deal of speculation. Please do not hold it against me, I can not defend, and could not resist any of it. I throw myself on the mercy of my examiners.

Declaration

This thesis is my own work unless otherwise acknowleded and at no time has been submitted for another degree.

. . . . . .

David Edward Whitworth.

I certify this statement to be correct.

. . . . . .

David A. Hodgson.

Abstract

Myxococcus xanthus is a soil-dwelling bacterium which produces carotenoids upon irradiation with blue light. Genetic analysis has allowed elucidation of transduction of the light signal to the carotenogenic machinery within the cell.
The primary element within the carotenogenic regulon is the genetic switch manifested by CarR and CarQ. CarR is an integral membrane protein which binds to the sigma factor CarQ and holds it in an inactive state at the cell membrane. Illumination of the cell with blue light excites the photosensitiser protoporphyrin IX (PPIX) within the bacterial membrane, which then excites molecular oxygen to the excited singlet state. Both singlet oxygen and excited triplet state PPIX can cause large amounts of cellular damage. Carotenoids prevent this damage by absorbing the excess energy from these excited species and dissipating it harmlessly as heat.
The presence of singlet oxygen within the bacterial membrane causes the inactivation/degradation of CarR. Removal of CarR releases CarQ from the membrane enabling it to mediate transcription from various promoters. CarQ causes transcription of the *crtI* gene and of the *carQRS* operon which produces further CarQ and CarS. CarS causes de-repression of the *crtEBDC* cluster. The carotenogenic enzymes encoded by *crtI* and the *crtEBDC* cluster catalyse the production of carotenoids which quench the initial signalling molecules, singlet oxygen and triplet PPIX. This causes down-regulation of the regulon as a whole as CarR is no longer degraded and once again carries nascent CarQ to the membrane in an inactive state.
The negative feedback loop described above is an important consideration when assessing mutants which produce carotenoids either constitutively (Car^C phenotype), or under no conditions (Car^- phenotype). This work investigates the consequences of Car^C and Car^- mutations on the activity of promoters within the Car regulon in order to clarify the roles of various genetic loci. It is demonstrated that CarA has no regulatory role in expression of *crtI* or *carQRS* and that the expression of *crtI* has no regulatory consequences. Sequencing downstream of *crtI* revealed a novel gene *gufB* (*g*ene of *u*nknown *f*unction *B*) which has homologues of no known function.
The critical event in the activation of the carotenogenic system is expression of the *carQRS* operon allowed by the release of CarQ from its complex with CarR at the membrane. Attempts were made to extract information about the interaction of CarQ with its cognate promoter at *carQRS* through a variety of *in vivo* and *in vitro* molecular and genetic techniques. Site-directed mutations within P^{carQRS} were assessed *in vivo* through the use of *lacZ* transcriptional fusions, enabling identification of important regions within the *carQRS* promoter. *In vitro* experiments provided information about the possibility of using molecular methods to assess interactions between CarQ and the P^{carQRS} promoter.

List of Abbreviations

$^1O_2$	singlet oxygen
$^3O_2$	molecular oxygen
^{3}PPIX	triplet protoporphyrin IX
bp	base pair(s)
CTP	cytidine triphosphate
DNA	deoxyribonucleic acid
DTT	dithiothreitol
dATP	deoxyadenosine triphosphate
dGTP	deoxyguanosine triphosphate
dCTP	deoxycytidine triphosphate
DMSO	dimethyl sulfoxide
dNTP	deoxyribonucleoside triphosphate
dTTP	deoxythymidine triphosphate
E. coli	*Escherichia coli*
EDTA	ethylenediamine tetra-acetic disodium salt
GGPP	geranylgeranyl diphosphate
GST	glutathione-S-transferase
HTH	helix-turn-helix
IPTG	isopropyl-β-D-thiogalactopyranoside
kb	kilo basepair(s)
KDa	kilo Dalton(s)
M. xanthus	*Myxococcus xanthus*
MOPS	3-[N-Morpholino] propanesulfonic acid
MW	molecular weight
$O_2^{\cdot}$	superoxide radical
$OD_{\#}$	optical density$_{(\text{nanometres})}$
ONPG	o-nitrophenyl-β-D-galactose pyranoside
ORF	open reading frame
P^{carQRS}	promoter of *carQRS*
P^{crtI}	promoter of *crtI*.
PCR	polymerase chain reaction
PPIX	protoporphyrin IX
rNTP	ribonucleoside triphosphate
SDS	sodium dodecyl sulphate
SOD	superoxide dismutase
TM	transmembrane
Tris	2-amino-2(hydroxymethyl)-1,3-propane diol
TSS	transformation and storage system
UV	ultra violet radiation
X-Gal	5-bromo-4-chloro-3-indolyl-galactoside

Abbreviations for amino acids.

Amino Acid	One-Letter Symbol	Three-Letter Symbol
Alanine	A	Ala
Arginine	R	Arg
Asparagine	N	Asn
Aspartic Acid	D	Asp
Cysteine	C	Cys
Glutamine	Q	Gln
Glutamic Acid	E	Glu
Glycine	G	Gly
Histidine	H	His
Isoleucine	I	Ile
Leucine	L	Leu
Lysine	K	Lys
Methionine	M	Met
Phenylalanine	F	Phe
Proline	P	Pro
Serine	S	Ser
Threonine	T	Thr
Tryptophan	W	Trp
Tyrosine	Y	Tyr
Valine	V	Val

1. Introduction

1.1. The Myxobacteria

The myxobacteria as a group of organisms were first described by Roland Thaxter in 1892 (Thaxter, 1892). They are found in almost all environments including Antarctica (Dawid, 1988), but their main habitat seems to be the temperate topsoil and rotting vegetation (Reichenbach, 1993). There are several physiological features of the myxobacteria that have prompted the interest of researchers and most of these aspects involve social interactions between individual cells (Dworkin, 1996). Myxobacterial colonies can exhibit a swarming motility which allows 'wolf-pack' behaviour as myxobacteria predates on other bacteria. Their predatory life-style involves the production of lytic exoenzymes and the lysis products of other bacteria are sufficient to sustain growth (Dworkin, 1962). Myxobacteria are also capable of undergoing multicellular development to form fruiting bodies, morphologically complicated cellular aggregates which contain resistant differentiated cells termed myxospores (Sudo and Dworkin, 1969; Dworkin, 1996). The production of a population of myxospores within a fruiting body, sometimes within a hardened sporangiole, means that on relief of starvation, a population of germinants is released which is more efficient in terms of exoenzyme production and predation on other bacteria than a single germinating cell.

The order of Myxococcales is divided into the sub-orders of Cystobacterineae and Sorangineae each of which are divided further into two families. The Myxococcaceae family within the Cystobacterineae contains the organism dealt with in this work, *Myxococcus xanthus*. The other myxobacterium which is widely studied is *Stigmatella aurantiaca* which is a member of the Cystobacteraceae family within the Cystobacterineae. The genus *Myxococcus* contains four species, *M. xanthus*, *M. fulvus*, *M. virescens* and *M. stipitatus*. All four species form fruiting bodies which are small mounds which differ mainly in their pigmentation and in the fruiting body stalks.

Species	Fruiting body stalk	Colouration	Other features
M. xanthus	none	yellow/orange	bright orange fruiting bodies
M. stipitatus	stalked	white/fawn	yellow fluorescence under UV
M. fulvus	constriction at base	white/pink/red	
M. virescencs	none	green/yellow/grey	

Table 1-1 Species of the genus Myxococcus *and their main similarities/differences.*

It must be pointed out that *M. xanthus* should not be considered a typical member of the *Myxococcus* genus, let alone of the entire Myxococcales order. It forms a very simple fruiting body in comparison to other myxobacteria such as *Stigmatella aurantiaca* and *Nannocystis elegans* and it also possesses some unique features. For example, it is the only myxobacterium for which illumination inhibits fuiting body formation rather than being a prerequisite for fruiting. For a general review of the classification and early characterisation of the myxobacteria see Reichenbach (1993).

1.1.1. *Myxococcus xanthus*

Most workers who research the myxobacteria and their features study *Myxococcus xanthus*. It is an easily culturable, genetically amenable organism with several generalised tranducing phage, a coliphage P1-based plasmid transduction system (O'Connor and Zusman, 1983) and it can be electroporated with high efficiency. It is readily studied using *lacZ* reporter genes and can support the introduction of transposons (For a review see Gill and Shimkets, 1993).

In the vegetative mode of growth *M. xanthus* cells are typically rod-shaped of dimensions 0.7-1.2 by 3-12 μm, with a doubling time of ~5 hours. Under conditions of limiting nutrients on a solid substrate, around 100,000 cells of *M.*

xanthus aggregate to form a mound of cells (a fruiting body) within which a subset of cells differentiate to form relatively resistant dormant cell forms or myxospores (Dworkin, 1996; Sudo and Dworkin, 1969). This process in *M.* xanthus has been shown to require at least five separate intercellular signals, cell-cell contact, extracellular fibrils, an active motility system etc. (Downard *et al.*, 1993).

Other interesting features of non-flagellate *M. xanthus* include its gliding motility system. Motion has two modes, adventurous, single cell motion (A-motility) and social, group movement (S-motility). S-motility requires cell-cell contacts and appears to be mediated by extracellular fibrils and pili. Both forms of motility require the production of extracellular polysaccharide in the form of slime which can be seen as a trail left behind by cells moving adventurously. Most mutants defective in motility appeared to be hindered in either A-motility or S-motility but not both (except for the *mglA* gene for mutual gliding; Stephens *et al.*, 1989). As well as the A (*agl* adventurous gliding) and S (*sgl* social gliding) genes (Hodgkin and Kaiser, 1979), many other loci have been implicated in motility, for instance the *frz* chemotactic system, the *dsp* genes and the *dif* chemotaxis homologues (Dworkin, 1996; Yang *et al.*, 1998).

The organism also exhibits a number of characteristics which had conventionally been regarded as present only in eukaryotes. For instance, the presence of serine/threonine kinases (Munoz-Dorado *et al.*, 1991), conserved 'eukaryotic' protein sub-domains (e.g. HMGI(Y) subdomains of CarD) and the ability to undergo complex co-ordinated multicellular morphogenesis. *M. xanthus* was one of the two myxobacteria (the other being *Stigmatella aurantiaca*) in which the first bacterial retron elements and their reverse transcriptases were found (Lampson, 1993). The genome size of *M. xanthus* is also unusually large at 9,454 kb, nearly double that of *E. coli*.

Myxobacteria are producers of a wide range of secondary metabolites including, for *M. xanthus*, the antibiotic TA (Varon *et al.*,1997), saframycin, althiamycin, myxovalargin and myxovirescin (for review, see Reichenbach and Hofle, 1993; Foster *et al.*, 1992).

The feature of *M. xanthus* biology which concerns the rest of this work is the production of coloured carotenoids as a response to illumination with blue light. This is the most apparent phenomenon associated with irradiation, but light has also been shown to inhibit fruiting body formation at an early stage in *M. xanthus* (Shimkets, L., pers. comm.). As mentioned earlier, this is the only member of the myxobacteria for which this is the case.

1.2. Biology and Light.

Light has a wide ranging and important impact on biological systems. It is obviously an integral requirement for photosynthesis and visual perception. Other biological roles for light include bioluminescence and vitamin D synthesis in animals, the establishment of circadian rhythms in a wide range of organisms and morphogenesis in fungi (Linden and Macino, 1997). The alga *Chlamydomonas reinhardtii* requires blue light to differentiate immature pregametes into mature gametes (Pan *et al.,* 1996). However, illumination also has less beneficial consequences. UV light is a potent carcinogen as a result of various modes of photo-induced DNA damage, including formation of pyrimidine dimers and hydroxylation of guanosine residues (Kohen *et al.,* 1995). Adsorption of radiation in aerobic conditions is also linked to the production of high-energy oxygen species which are capable of causing potentially lethal damage and the prevention of this oxidative damage is of prime concern to photosynthetic organisms. In general most organisms seem to have evolved coordinated mechanisms for protection against illumination and oxidative damage, for example the SOS system (Humayun, 1998) and the SoxR and OxyR systems in *E. coli* and *S. typhimurium* (Ahern and Cunningham, 1995).

1.2.1. Photochemistry

Light energy can be absorbed by many molecules with the colour of light absorbed being dependent on the chemical and physical nature of the molecule absorbing the light. For instance bacteriorhodopsin appears red because it absorbs blue light. It is capable of doing this because its conjugated retinal molecule contains a tract of six conjugated double bonds which delocalise their electric fields over the region of conjugation.

Molecules which absorb the energy of light have their own intrinsic energy increased as electrons are raised from the ground state to an excited state. The second law of thermodynamics dictates that the system must return to its lowest energetic state if possible, reverting to ground state. This results in a number of possibilities for the excited molecule. It may re-emit its excitation energy as light due to electronic transitions, giving rise to fluorescence. Alternatively, it may emit the energy as heat due to vibrational relaxation, or in some cases the excitation energy is released by chemical reaction involving the excited molecule.

If the electron that is excited is one of a pair of electrons within a particular bonding or non-bonding orbital, then in the ground state the electrons will have opposite spin and exist in a singlet state. (Spin states of electrons are governed by a quantum number which may have the values 1/2 or -1/2, each denoting a 'direction' of spin. The state of the electron pair is denoted by S=2I+1, where I is net spin, and therefore if the electrons have opposite spin, S=1 {2[1/2+-1/2]+1}, i.e. a singlet state. If the electrons are spinning in the same direction S=3 i.e. a triplet state.) On excitation one of the electrons is removed to an orbital of higher energy to form an excited singlet state. Since the electrons are now in separate orbitals the excited electron may change its spin state, a process which is now no longer forbidden by the Pauli exclusion principle. This spin flipping is termed intersystem cross-over and results in a metastable triplet state.

Excited triplet states are relatively stable since the excited electron must first return to its original spin state before it can lose its excess energy and return to the ground state. This leads to an extended half-life of excitation and the delayed

remission of absorbed light is termed phosphorescence rather than fluorescence. Figure 1-1 demonstrates these possible transitions for a pair of electrons within a π bond.

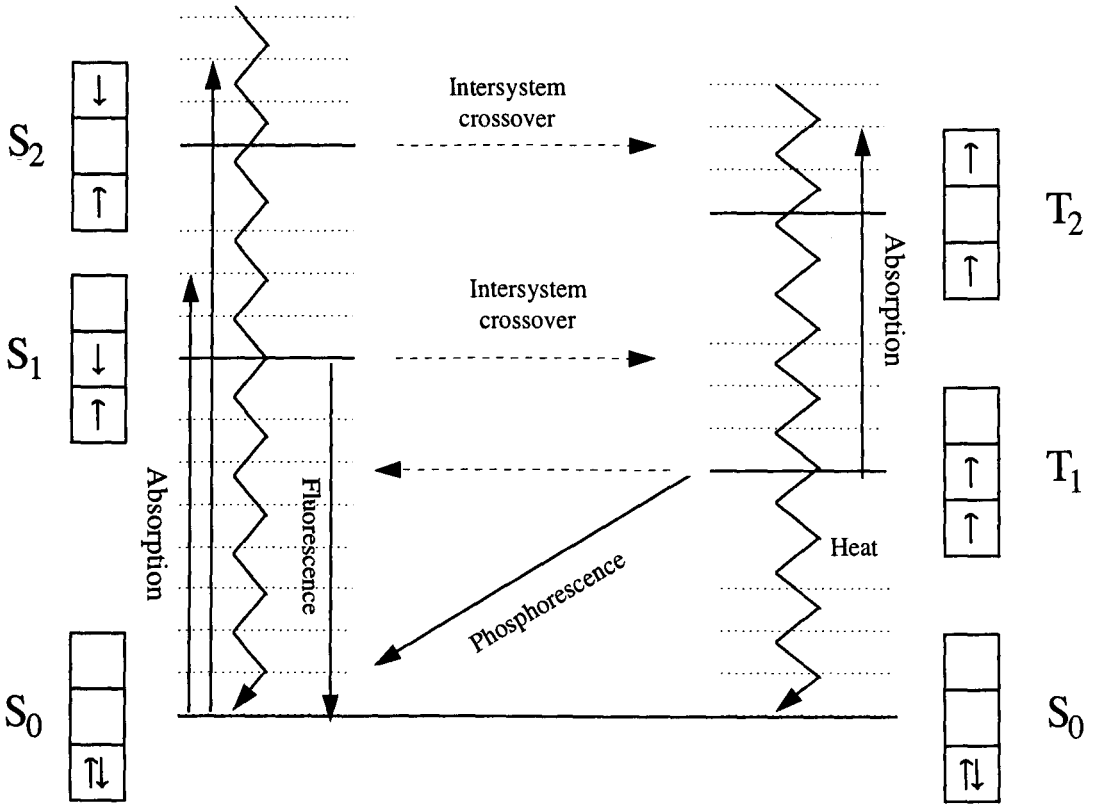


Figure 1-1 Jablonski diagram of electronic transitions of π electron pairs.

The ground state is a singlet state with electrons having different spin states. On absorption of energy an electron is promoted to a higher electronic orbital (S_1 and S_2). Intersystem crossover may occur flipping the spin state of the excited electron and producing an excited triplet state (T_1 and T_2). Relaxation mechanisms described include fluorescence, phosphorescence and vibrational relaxation which emits heat. Adapted from Kohen et al., *1995.*

1.2.2. Photosensitisation

Photosensitisers are molecules which absorb incident light to cause a deleterious effect within the cell. Excitation of photosensitising molecules is responsible for a great deal of light-induced cellular damage. Damage is mainly caused by the last of the relaxation mechanisms described in the previous secton - through the excitation of a second molecule, either a cellular biomolecule, leading to formation of photosensitiser adducts, or molecular oxygen, which is excited to the potent oxidiser, singlet oxygen.

In the myxobacterial membrane is the photosensitiser protoporphyrin IX which is the direct biosynthetic precursor to haem (Burchard *et al.*, 1966; Burchard and Hendricks, 1969). The molecule is ubiquitous within cellular membranes but its levels increase around 16-fold as the cells reach stationary phase. It is a porphyrin as its name suggests and possesses a large resonance energy at the same wavelength as blue light. Blue light incident on protoporphyrin IX causes the excitation of the photosensitiser to an excited singlet state from the ground singlet state. The excited singlet state PPIX can undergo inter-system cross-over (a forbidden transition) and form metastable excited triplet state PPIX. This triplet state can cause cellular damage through two types of mechanism.

In type II, or oxygen-independent photosensitisation, damage is caused directly by excited photosensitiser, through oxidation and porphyrin adduct formation with macromolecules (Foote, 1991). In type I, oxygen-dependent photosensitisation, the photosensitiser triplet state is sufficiently long-lived to cause the excitation of ground state triplet oxygen to an excited singlet state. Singlet oxygen can then itself cause cellular damage by causing peroxidation of lipids (a radical chain reaction), cross-linking of proteins, cleavage of DNA (Fiel *et al.*, 1981) and electrophilic or nucleophilic oxidation of susceptible biomolecules (principle oxidation reactions of singlet oxygen are shown in Figure 1-2). Singlet oxygen is a particular threat to the bacterium as it is relatively long-lived in the hydrophobic environment of the bacterial membrane with a half-life of up to 100μs (Knox and

Dodge, 1985; Suwa, Kimura and Schaap, 1977) and is primarily responsible for damage within membranes (for review see Valenzeno, 1987).

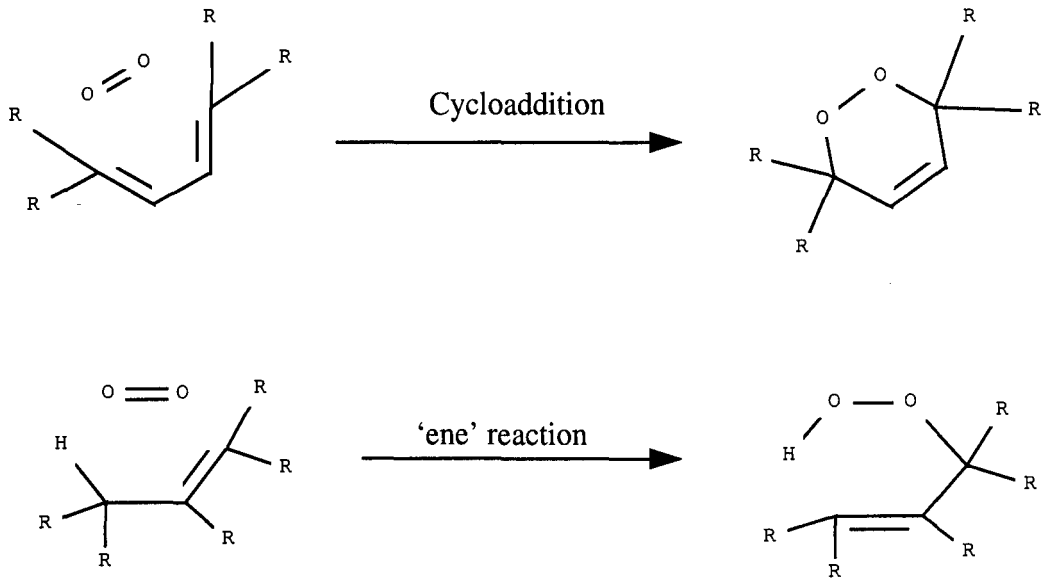


Figure 1-2 Reactions between singlet oxygen and susceptible functionalities.

In the 'ene' reaction and the cycloaddition reaction, introduction of oxygen into the target molecule results in the formation of organic peroxides.

1.2.3. Oxidative Damage

Oxygen can be toxic to most lifeforms and is most potent as a set of energised oxygen-containing species which includes singlet oxygen and peroxide radicals. The most destructive of these is the radical superoxide ion ($O_2^{\cdot}$-) which can attack the macromolecules within a bacterial cell causing mutation or death. Superoxide can also undergo reduction to form hydrogen peroxide (H_2O_2) and hydroxyl radicals ($OH^{\cdot}$) which are also potent oxidants (Fridovich, 1978). These species are

encountered due to sequential monovalent reduction of oxygen, a series of reactions caused by many metabolic enzymes and processes (Fridovitch, 1978).

The other main source of high energy oxygen species is directly through illumination via photosensitisers. Photosensitisation can be mediated by any molecule with π orbitals, but is primarily caused by molecules that contain delocalised electrons, such as found in tetrapyrroles, aromatics and polyenes. This is because delocalised electrons require less energy to excite than those bound rigidly in localised orbitals. Photosynthetic bacteria generate high energy oxygen species as a result of their capture of energy from light and its subsequent conversion into proton motive force, due to the presence of chlorophyll which acts as a photosensitiser. *Rhodobacter* sp. have avoided this problem by only engaging in photosynthesis in anerobic conditions.

As a general source of oxidative damage in cells, similar means are often employed to prevent oxidative damage caused by light and oxidative damage caused through the generation of activated oxygen species by other pathways.

1.2.4. Prevention of Oxidative Damage

The reponses to oxidative and photooxidative damage can be broadly classified into two general mechanisms. Firstly, the production of primary antioxidants, which cause the direct removal of the high energy oxygen species. These species include the enzymes catalase and superoxide dismutase (which catalytically remove the oxidants H_2O_2 and $O_2^{\cdot}$ respectively) and small antioxidant molecules such as glutathione and α-tocopherol. The second class of antioxidant reponses comprise a range of activities, including the repair of biomolecules already damaged by oxidation and the induction secondary defenses which maintain an environment capable of supporting the requirements of the primary antioxidants, e.g. glutathione reductase and glutathione s-transferase which maintain glutathione in its active reduced form (Ahmad, 1995).

Antioxidant defenses are co-ordinately expressed in regulons which typically respond to a precise antioxidant species, although some antioxidant responses are activated generally by 'oxidising conditions'.

The SoxR protein is a regulator of gene expression in response to the presence of superoxide radicals (Hidalgo *et al.*, 1998) in *Escherichia coli.* The presence of superoxide causes SoxR to activate expression of *soxS* which in turn activates nine *sox* genes including MnSOD (manganese-containing superoxide dismutase) and endonuclease IV (Nunoshiba *et al.*, 1992). The response to H_2O_2 in *E. coli* and *S. typhimurium* involves the regulator OxyR. The OxyR protein represses transcription of its own gene in all conditions. In the presence of peroxides however, it also causes the activation of nine genes including those encoding KatG (the stationary phase sigma factor, SigS) and AphC (alkyl hydroperoxide reductase), which are involved in the antioxidant reponse (Tartaglia *et al.*, 1992). The activation and deactivation of OxyR appears to be modulated by a pair of cysteine residues which can be covalently linked in a disulfide bond under oxidising conditions and which exist as thiols in reducing conditions (Zheng *et al.*, 1998). Thus OxyR is intrinsically receptive to the redox state of the cell.

The SOS response elucidated in *E. coli* and *S. typhimurium* is a co-ordinately regulated set of induced functions (Walker, 1984), induced as a response to UV-induced DNA damage. These functions include inhibition of cell division and post-irradiation DNA degradation, induced mutagenesis of DNA and induced lysis and mutagenesis of lysogens (Weigle, 1953). In the non-stressed cell, LexR represses expression of RecA and the SOS regulon. It is thought that replication cannot proceed regions of UV-induced DNA damage, and the stalled replicasome exposes single stranded DNA at the replication fork to RecA, which is conformationally activated to RecA*. RecA* induces LexA to undergo autoproteolysis, relieving repression of *recA* and the SOS regulon. Part of the SOS response seems to cause modification of DNA polymerases so that they are capable of replicating past damaged regions of DNA, although introducing errors at higher frequency (Humayun, 1998). Other features of the response include the

induction of DNA repair mechanisms, including excision, recombination and mutagenic repair mechanisms (Sancar and Sancar, 1988).

A further reponse to UV-induced DNA damage is the removal of thymine dimers by photolyase (Sancar and Sancar, 1988; Sancar, 1994). Photolyase utilises the energy of light to repair the DNA damage caused by light. Binding of enzyme to its substrate results in a complex, which upon absorption of light catalyses cleavage of the thymine dimer (Schieferstein and Thoma, 1998). The photolyase of *M. xanthus* has recently been identified and shown to be more similar to the photolyase of eukaryotes than to those of other eubacteria (O'Connor *et al.*, 1996).

Blue light causes the photosensitisation of susceptible organisms through the production of singlet oxygen. Production of singlet oxygen induces various anti-oxidant defenses, however, while overlapping with the antioxidant responses to superoxide and peroxides (e.g. alkyl hydroperoxidase induction), there does not seem to be a co-ordinated 'general antioxidant response' which is induced by all oxidants.

Bacteria have evolved many solutions to the problems of cellular damage caused by oxidative mechanisms such as those detailed above. In general, primary antioxidant responses fall into two broad classes: enzymatic processes such as those catalysed by catalase and superoxide dismutase, or the production of small antioxidants including α-tocopherol, glutathione and carotenoids. The production of carotenoids is a major subset of responses to singlet oxygen-mediated photo-induced oxidative damage.

1.2.5. Role of carotenoids as photoprotectors

Carotenoids are secondary metabolites produced mainly in conditions of arrested growth or limiting food. They are typically C_{40} conjugated polyenes and are a reddish-orange colour. Carotenoids fall into two chemical classes, the carotenes and the xanthophylls, which are oxygen containing carotenes.

Carotenoids protect against photosensitisation damage since they are capable of quenching excited photosensitisers, singlet oxygen and are also able to directly absorb incident light. The excess energy absorbed by the carotenoids as they quench the excited species is shuttled backwards and forwards along the conjugated carbon backbone of the carotenoid molecule and is dispersed harmlessly as thermal energy (through vibrational rather than electronic relaxation). This requires a tract of at least seven conjugated double-bonds (Fiel *et al.*, 1981). The excited singlet state of the typical carotenoid, β-carotene has a half-life of around 8 ps and this means that quenching of singlet oxygen by carotenoids is limited only by diffusion (Cogdell and Frank, 1987). The occurrence of carotenoids in the hydrophobic membrane means that they are also ideally situated for photoprotection, as they reside in the same environment as the triplet PPIX and singlet oxygen they quench.

In the plant and bacterial photosynthetic reaction centre, carotenoids are positioned to protect the integral chlorophyll molecules from oxidation by singlet oxygen, to protect the cell against photosensitisation and to act as accessary light-collecting pigments (Kuhlbrandt *et al.*, 1994; Cogdell and Frank, 1987). In *Rhodobacter sphaeroides,* a blue-green mutant which didn't produce end-product carotenoids was shown to grow normally under anaerobic, photosynthetic conditions. However, when illuminated in aerobic conditions the mutant underwent rapid cell death due to photosensitisation by bacteriochlorophyll (Dworkin, 1958, cited in Dworkin, 1959). Zhu and Hearst (1986) showed that expression of genes coding for reaction centre and light-harvesting complex proteins was down-regulated by high light intensity and oxygen concentration in *R. capsulatus*. Expression of *crt* genes however was enhanced by increasing levels of light and oxygen consistant with a role in protection from photosensitisation.

Additionally, carotenoids are produced as a protective measure against illumination with blue light among non-photosynthetic bacteria. In some cases this is independent of light (e.g. *Erwinia* {Armstrong *et. al.*, 1990 and To *et al.*, 1994}) although light still up-regulates carotenoid production, whereas in others, illumination is absolutely required (e.g. *Myxococcus xanthus*). In some cases

carotenoid production can even be cryptic, which is the case for *Streptomyces griseus* (Schumann *et al.*, 1996).

In *M. xanthus*, photodynamic sensitisation through protoporphyrin IX has an action spectrum similar to both the absorption spectrum of porphyrins and to the spectrum for the induction of carotenogenesis (Burchard and Hendricks, 1969). Excited protoporphyrin IX and singlet oxygen are quenched by the carotenoids produced by *M. xanthus*, and they therefore prevent cell damage (Hodgson and Murillo, 1993). It seems that in light there is competition between the processes of photolysis and carotenoid induction, determining the fate of illuminated cells (Burchard and Hendricks, 1969). This competition is affected by many diverse factors including growth phase, temperature and metal ions, with monovalent ions enhancing and divalent ions inhibiting lysis (Burchard and Dworkin, 1966).

It should be noted that production of carotenoids as protectors against photooxidative damage is not restricted to the bacteria. Other examples include *Phaffia rhodozyma*, *Neurospora crassa* and all photosynthetic organisms. Carotenoid synthesis is found in fungi, yeast and plants. Although animals can not synthesise carotenoids *de novo*, they are capable of modifying ingested carotenoids.

1.3. Carotenoids

1.3.1. Carotenoids of *M. xanthus*.

The carotenoids synthesised by *M. xanthus* number 50-60, but most are produced in relatively low amounts. They exhibit varying degrees of unsaturation and may be covalently modified by cyclisation, hydroxylation, ketonation or esterification via sugar residues to fatty acids. The majority of carotenoids found in *M. fulvus* are xanthophylls (oxygenated carotenoids) in the form of myxobactin and myxobacton esters (>70%), and 4-keto-torulene glucoside (Reichenbach and Kleinig, 1984; Hodgson and Murillo, 1993 and shown in Figure 1-3). They cause

an orange/red colouration, with significant carotenoid synthesis occuring only upon illumination and when the cells have entered stationary phase when levels of protoporphyrin IX rise around sixteen-fold (Burchard and Dworkin, 1966; Fontes *et al.*, 1993).

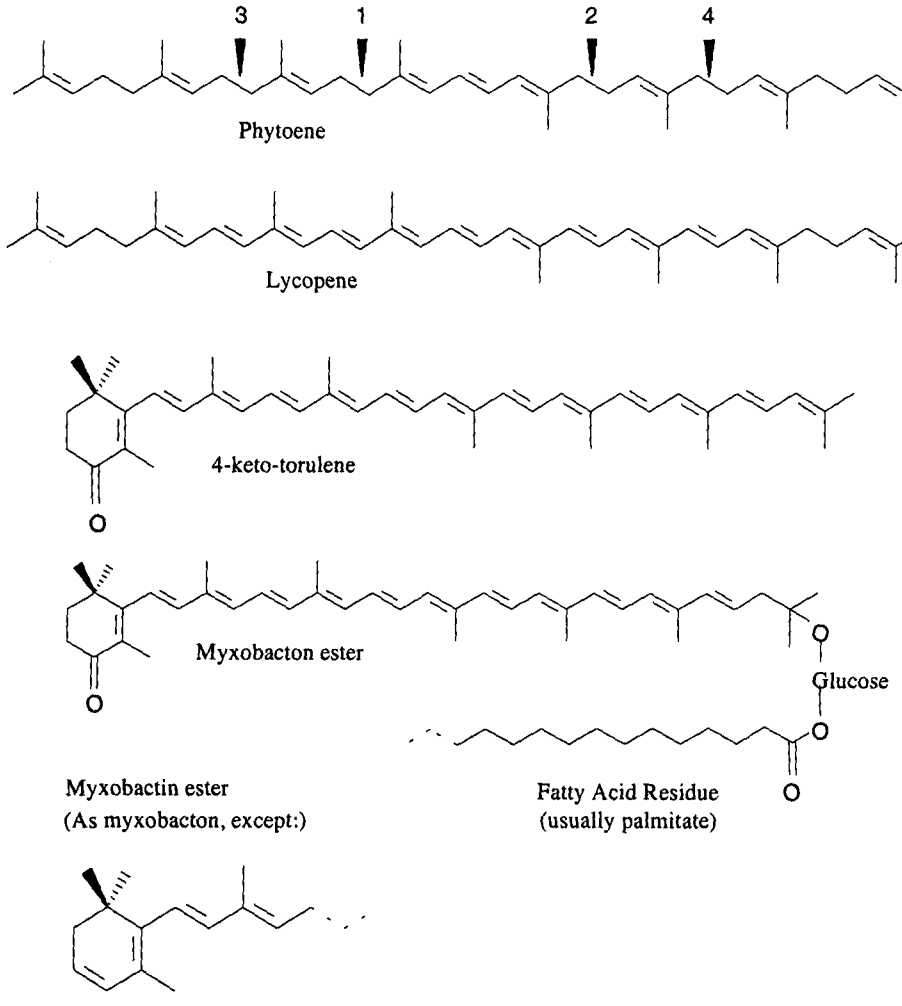


Figure 1-3 Phytoene, Lycopene and End-Product Myxobacterial Carotenoids.

The four bonds of phytoene which are oxidised to from lycopene are indicated with arrows. The order of reactions according to the Porter-Lincoln series is denoted numerically.

1.3.2. Carotenoid synthesis

The pathway of carotenoid production shares a common pathway with steroid and terpenoid production, with condensation of isoprenoid units leading to the production of C_{20} geranylgeranyl-diphosphate (GGPP). In carotenogenesis, two molecules of GGPP are condensed to from C_{40} phytoene via the C_{40} intermediate prephytoene PP_i by CrtB. Phytoene may then be desaturated in a series of four reactions catalysed by CrtI that produce consecutively, phytofluene, ζ-carotene, neurosporene and finally lycopene which is a pink/red colour, according to the series described by Porter and Lincoln (1950). Lycopene is a very potent scavenger of singlet oxygen, acting 120 times more efficiently than α-tocopherol (Di Mascio *et al.,* 1990). Subsequent steps in the carotenogenic pathway are organism-specific as are the end-products of carotenogenesis.

In *M. fulvus*, there is a split in the carotenogenic pathway after the third desaturation of phytoene, with neurosporene either being dehydrogenated to from lycopene or alternatively undergoing hydroxylation to form hydroxyneurosporene. Hydroxyneurosporene is then converted to 3,4 dehydro-rhodopin glucoside ester by two dehydrogenations and an esterification and further modified to myxobacton esters by cyclisation and oxidation. Lycopene may also be converted into 3,4 dehydro-rhodopin glucoside ester by a series of reactions, alternatively it is sequentially converted to 4 keto-torulene via γ-carotene and 4-keto γ-carotene (Kleinig *et al.,* 1975). See Figure 1-4 for a schematic representation of the synthetic pathway, with the enzymatic activities denoted.

It should be emphasised that although it is generally assumed that this pathway is also true for *M. xanthus* this has not yet been proven to be the case. Indeed, the presence of 4-keto-torulene cannot be detected in wild-type *M. xanthus* strains, even in the light (Ruiz-Vazquez *et al.,* 1993).

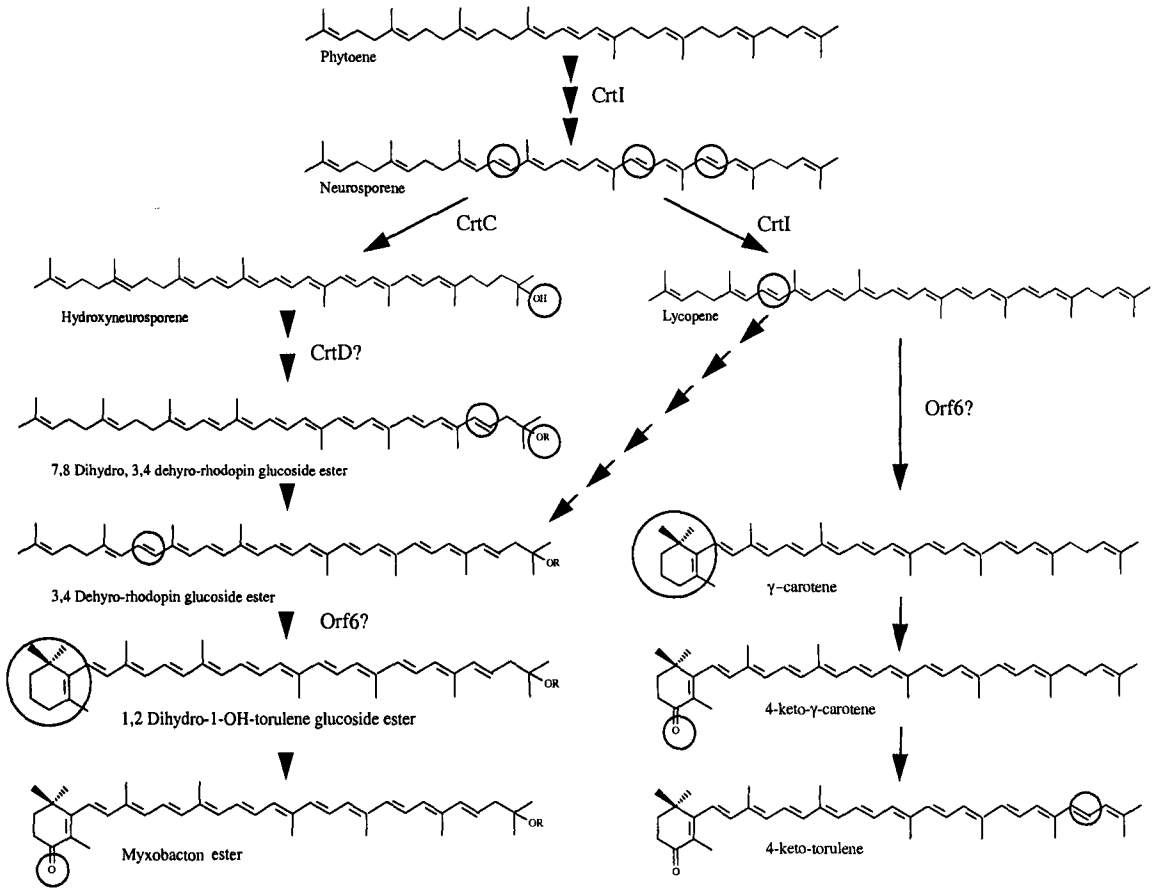
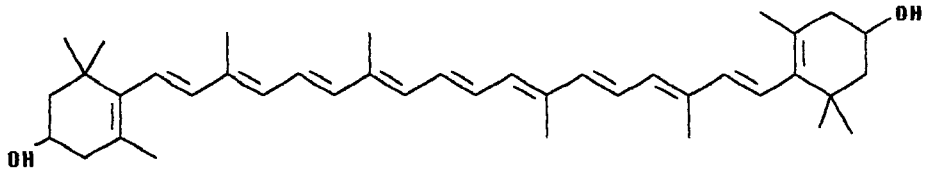


Figure 1-4 Pathway of carotenoid production in M. fulvus.

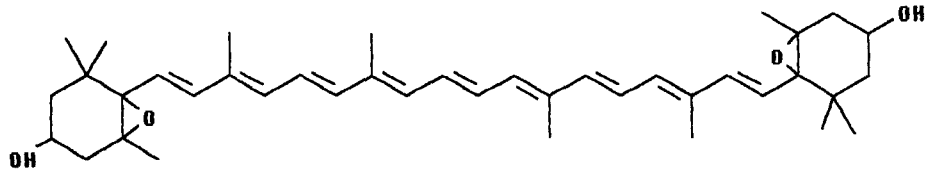
Adapted from Hodgson and Murillo, 1993. Enzymatic steps are indicated by circles around the functionalities introduced.

In *M. xanthus* the carotenogenic response is switched on by light. This could be rationalised since there could conceivably be a huge metabolic burden inherent in the production of carotenoids. Conversely, there seems to be very little effect on the physiology of cells actively producing carotenoids with no detectable reduction in growth rate during carotenoid production. However, the growth of cultures in the laboratory is not reflective of the situation in nature where nutrients may well be limiting. Nevertheless, the necessity for inducibility of carotenogenesis in *M. xanthus* remains unclear.

Carotenoids may also have a role in photoprotection other than the quenching of excited singlet oxygen and photosensititsers. In the thylakoid membrane of plants the constitutuent phospholipids are highly saturated and there is no cholesterol or sterols, which serve to maintain the membrane in a very fluid state which is necessary for photosynthetic processes involving membrane diffusion. This makes the membrane particularly vulnerable to photooxidative and thermal damage. Peroxidation of lipids is greater in highly saturated lipids and thermal membrane damage is associated with hyperfluidity and leads ultimately to membrane protein denaturation. Since the thylakoid membrane requires high fluidity for function it is highly susceptible to both forms of damage. The thylakoid membrane contains high levels of the carotenoid violaxanthin, a bi-cyclic doubly-epoxidated derivative of zeaxanthin. (Figure 1-5).



Zeaxanthin



Violaxanthin

Figure 1-5 Carotenoids of the xanthophyll cycle.

In the thylakoid lumen resides an enzyme capable of de-epoxidating violaxanthin to zeaxanthin. This activity is only present *in vivo* in high light conditions and may require release of bound violaxanthin from the light harvesting complexes, due to conformational changes in the LHC on sudden illumination. The zeaxanthin may rebind to the LHC or it may remain in the membrane lipid phase. Reduction whilst under illumination causes zeaxanthin to be slowly converted back to violaxanthin. This comprises the xanthophyll cycle. The function of zeaxanthin is obscure but hints as to its role exist. As well as its ability to quench high-energy species, zeaxanthin is dipolar, with a length corresponding to the width of the thylakoid membrane and is likely to exist perpendicular to the plane of the membrane. This leads to an increase in the viscosity of the membrane and also reduces the permeability of the bilayer to oxygen species. Thus the membrane is protected from both lipid peroxidation and also from thermal damage, which will usually go hand-in-hand with high illumination. Zeaxanthin appears to have additional roles in the plant as well, with evidence suggesting a role in the transduction of a blue light signal to stomatal opening and coleoptile binding (Quinones and Zeiger, 1994). Therefore carotenoids can have a role in modulation of membrane fluidity as well as the quenching of excited species in the prevention of photooxidative damage (Havaux, 1998).

1.3.3. Bacterial carotenogenesis in other organisms.

In the Gram negative, facultative phototroph *Rhodobacter capsulatus* and the phytopathogen *Erwinia uredevora*, carotenoid biosynthesis genes are clustered. Eight carotenoid biosynthesis genes are clustered together in *R. capsulatus* in a minimum of four operons, *crtA*, *crtIBK*, *crtDC* and *crtEF* (Armstrong *et al.*, 1989). In *Rhodobacter capsulatus* the major carotenoid end-product is acyclic spheroidene in anaerobic cultures and its keto-derivative spheroidenone in aerobic cultures (Figure 1-6). The gene products of *crtD* and *crtI* were found to share homology, and both act as dehydrogenases. It is thought that *crtD* acts to dehydrogenate methoxyneurosporene and hydroxyneurosporene to spheroidene

and demethylspheroidene respectively. In the related species *R. sphaeroides,* there is also clustering of carotenoid genes with *crtA, crtB, crtC, crtD, crtE* and *crtF* mapping together.

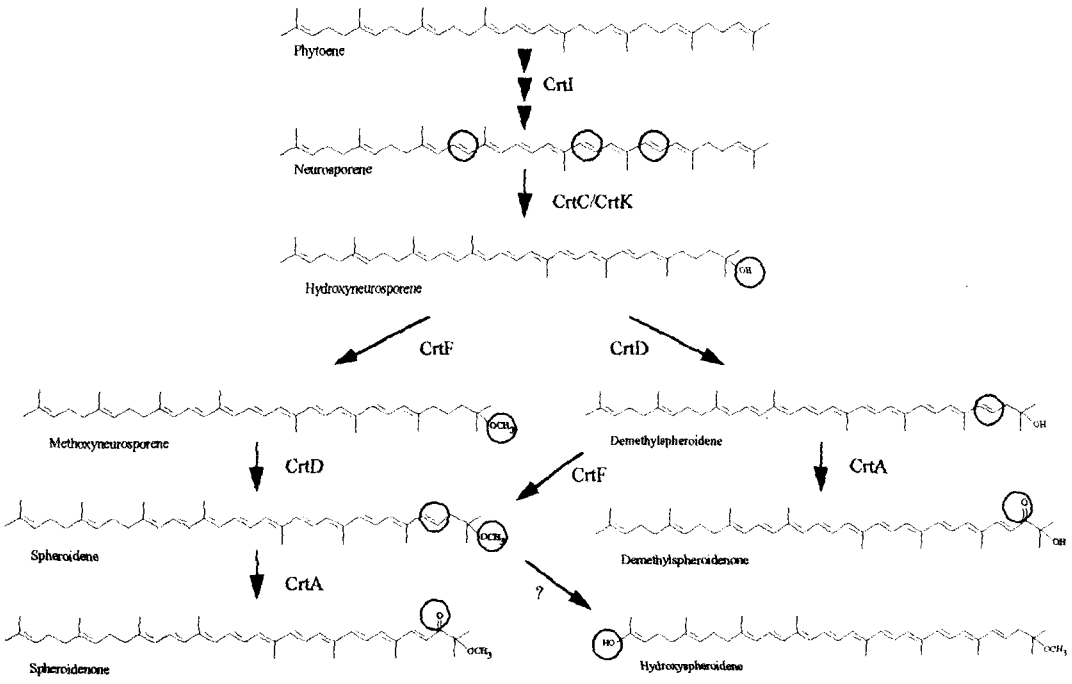


Figure 1-6 Carotenogenic pathway of R. capsulatus.

Adapted from Armstrong et al. *(1989). Enzymatic steps are indicated and introduced functional groups circled.*

In *Erwinia uredevora* six carotenoid genes are found in a cluster in a minimum of two operons, *crtEXYIB* and *crtZ* (Misawa *et al.,* 1990). Genes *crtY,Z* and *X* have roles not found in *R. capsulatus*, being responsible for cyclisation of lycopene, hydroxylation of β-carotene and esterification to sugars respectively (Figure 1-7).

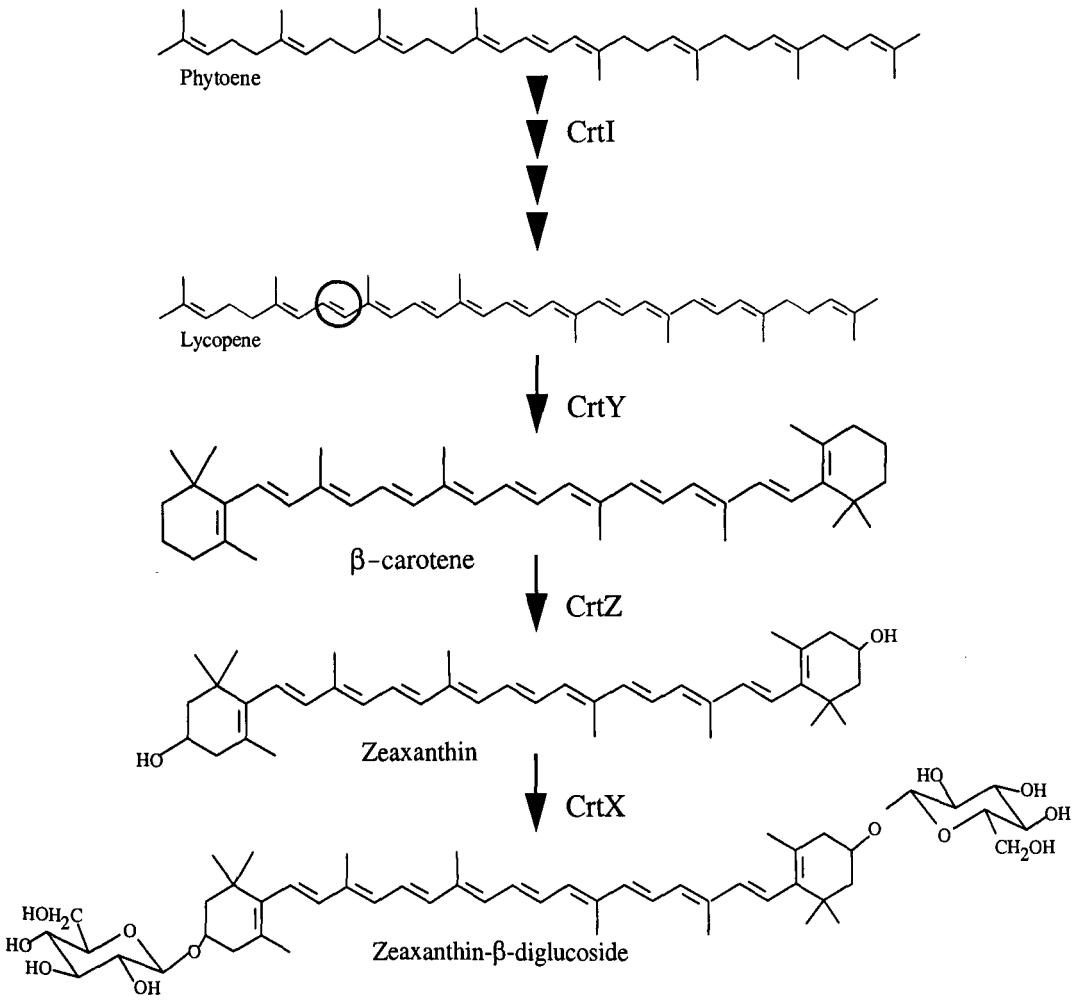


Figure 1-7 Carotenogenic Pathway of Erwinia uredovora.

Adapted from Misawa et al. *(1990). Enzymatic steps are indicated and introduced functional groups circled.*

Light-induced carotenogenesis is observed in a variety of species of actinomycetes. In *Mycobacterium marinum* carotenogenesis is light-induced and the biosynthetic genes *crtI* and *crtB* have been isolated (Ramakrishnan *et al.*, 1997). In *Mycobacterium vaccae* the expression of carotenogenic genes is repressed in the dark, with light relieving the repression (Houssaini-Iraqui *et al.*, 1992). In *Streptomyces setonii* ISP5395 a sigma factor homologue, *crtS* is responsible for expression of carotenogenic genes. *S. setonii* normally does not produce carotenoids, however, it contains cryptic genes for carotenogenesis whose

expression can be induced in mutants obtained through protoplast regeneration (Kato *et al.*, 1995).

Unusually, *M. xanthus* has a *crtI* gene distinct from the other carotenogenesis structural genes. In other bacteria, all *crt* genes are clustered into operons, including *crtI* (Armstrong, 1997).

1.3.4. Eukaryotic carotenoid production.

In eukaryotes, production of carotenoids occurs as a response to different stimuli than in non-photosynthetic bacteria. For instance, plants produce carotenoids as a response to red light which stimulates the production of chloroplasts (Hader and Tevini, 1987). Clustering of carotenoid biosynthesis genes is not the case for eukaryotes. In the filamentous ascomycete *Neurospora crassa*, during mycelial growth, carotenogenesis is light-induced while during the developmental pathway, carotenogenesis can be independent of light. The identified genes for carotenogenesis in *Neurospora crassa* are unlinked. The *al-1* gene has been shown to code for a product with phytoene dehydrogenase activity (Schmidthauser *et al.*, 1990) and is homologous to *crtI* and *crtD* from *R. capsulatus*. Genes *al-2* and *al-3* have been identified and *al-3* has been shown to encode GGPP synthase (Carattoli *et al.*, 1991). Expression of the albino genes is under the control of a pair of DNA binding proteins, white collar 1 and white collar 2 (*wc-1* and *wc-2*). Both genes code for proteins which contain Zn-finger motifs, dimerisation domains and transcriptional activation domains (Linden and Macino, 1997). How WC-1 and WC-2 interact with the flavin thought to perceive the light signal (Paietta and Sargent, 1981), is unclear.

1.4. Carotenoid biosynthetic genes in *Myxococcus*.

In *M. xanthus* the genes responsible for carotenogenesis are distributed between three genetically unlinked loci. The three central loci are the *carQRS* operon, the *carBA* cluster and the *carC* gene. Recently, the nomenclature for the Car genes has

been changed (Botella *et al.*, 1995). The structural genes for carotenogenic enzymes have been given the *crt* designation based on their homologies to known carotenogenic enzymes from other organisms. Regulatory genes however retain their *car* designation with *car* being an acronym for *crt* gene activity regulator genes (Hodgson and Berry, 1998). The table below shows the old and revised names of the genes involved in *M. xanthus* carotenogenesis (Botella *et al.*, 1995).

Functionality	Old name	New Name
phytoene dehydrogenase	*carC*	*crtI*
biosynthetic cluster (orfs 1-6)	*carB*	*crtEBDC*
geranylgeranyl diphosphate synthase	*orf1*	*crtE*
carotene desaturase?	*orf2*	-
phytoene synthase	*orf3*	*crtB*
hydroxyneurosporene desaturase	*orf4*	*crtD*
neurosporene hydratase	*orf5*	*crtC*
carotene cyclisation?	*orf6*	-

Table 1-2 New and old designations for carotenogenesis genes of M. xanthus.

The majority of carotenogenic genes are found within the *crtEBDC* and *carA* operons (*carBA*). This cluster contains eleven open reading frames with potential translational coupling (where the efficiencies of translation of different genes within a polycistronic mRNA molecule are not independent, Rex *et al.*, 1994) between *orf2* and *orf3*, *orf4* and *orf5*, *orf5* and *orf6*, *orf6* and *orf7*, *orf7* and *orf8*, *orf8* and *orf9* and between *orf9* and *orf10*. All open reading frames except *orf11* appear to be preceded by ribosome binding sites and *orf3* and *orf9* each start with a GTG not ATG initiation codon. From sequence similarities, Botella *et al.* (1995) assigned some of the open reading frames with *crt* gene functions. Thus due to the similarity of its gene product with geranylgeranyl diphosphate synthase, *orf1* was renamed *crtE*. (See Table 1-2.)

The four successive dehydrogenations of phytoene are performed by a single polypeptide in related bacteria, but by two distinct desaturases in photosynthetic

organisms (Linden *et al.*, 1991 and 1994). There is evidence that in *M. xanthus* *orf2* of the *crtEBDC* (*carB*) operon encodes an enzyme capable of converting phytoene into phytofluene (Murillo, F. J., pers. comm.) as transformation of *E. coli* with a plasmid carrying *orf2*, enabled the strain to produce copious amounts of phytofluene. CrtI (CarC) is thought to catalyse the subsequent desaturations to form neurosporene and lycopene. Further evidence for this scenario was found when assessing the production of intermediates of the carotenogenic pathway (Martinez-Laborda, *et al.*, 1990), as only phytofluene is formed in a *crtI* (*carC*) mutant in the light, due to the action of a gene product presumably encoded within the *crtEBDC* (*carB*) cluster.

Two distinct regions of *orf9* show homology to ferrochetalase which introduces ferrous iron into PPIX to form haem. Genes *orf10* and *orf11* are 35% identical and contain putative HTH motifs at their N-terminal regions, with greatest homology to MerR of Tn*501* (Parkhill *et al.*, 1998). MerR forms a homodimer which binds to a region of dyad symmetry between the -10 and -35 elements of the *merTPAD* (mercury resistance operon) promoter. MerR acts as a repressor unless it has bound mercury in which case it introduces a sharp distortion in the promoter DNA and activates transcription (Parkhill *et al.*, 1998). MerR is also related to SoxR of *E. coli* which is a redox-sensing gene activator which expresses a regulon as a response to oxidative stress (Nunoshiba *et al.*, 1992). The other biosynthetic gene *crtI* (*carC*) is unlinked to the *crtEBDC* (*carB*) cluster and codes for phytoene dehydrogenase.

1.5. Regulation of carotenogenesis.

Induction of carotenogenesis by reactive oxygen species seems to be a commonplace phenomenon. Examples include the yeast *Phaffia rhodozyma* (Schroeder and Johnson, 1993). The organism lives within the bark of birch trees which produce reactive oxygen species as a defense mechanism. The reactive oxygen intermediates cause an increase in carotenoid production in cell cultures. This seems to be due to a selection of those cells which produce greater levels of carotenoids and an inverse relationship between sensitivity to superoxide radicals and carotenoid content was found. Resistance to the oxygen species was particularly increased in stationary phase. *Phaffia rhodozyma* possesses only manganese-containing superoxide dismutase, which is found in the mitochondria only. This has led to the inference that carotenoids may act as extra-mitochondrial protectors against oxidative stress. Subsequent studies on this organism have shown that oxygen species actively induce an increase in carotenoid levels within *Phaffia rhodozyma* cells (Schroeder and Johnson, 1995). This is due to two mechanisms. Firstly, singlet oxygen induces carotenoid synthesis. Additionally, the oxygen causes the oxidation and removal of an existing pool of carotenoids. This then relieves feedback inhibition of the biosynthetic enzymes and more carotenoids are produced with production greatest during stationary phase. A rapid turnover of the carotenoid pool allows a rapid response to light and relieves the necessity for regulatory genes.

1.6. Regulation in *M. xanthus*

Regulatory genes are found at the *carQRS* locus and *carA* which resides at the 3' end of the *carA* operon. These are the only true regulators of carotenogenesis identified to date. In addition, another two loci have been found to contain genes required for carotenoid production although these genes have pleiotropic effects and are therefore not strictly carotenogenesis genes. The first of these contains the *carD* gene which contains motifs resembling those found in HMGY(I) proteins in

eukaryotes. The second locus (*carE*) contains one of the genes for integration host factor (*ihf*) of *M. xanthus*. Over the last decade, the molecular basis for the action of the carotenogenic loci has been elucidated by a combination of biochemical and genetic means.

1.6.1. The history of research on carotenogenesis of *M. xanthus*.

1.6.1.1. Identification of the *carR*, *crtEBDC*, *carA* and *crtI* loci.

Many of the initial Car^c carotenogenic mutants were isolated in laboratories working with *M. xanthus* due to the obvious manifestation of the mutation, which causes a bright red colouration in the light and dark. Conversely, very few Car^- mutants have been isolated since nobody routinely incubates their strains in the light which is the condition required for observation of the Car^- mutant phenotype. The carotenogenic loci were originally mapped by the creation of Tn*5* insertions which cotransduced to known Car^c mutations, giving constitutive carotenogenesis. Five linked to *carR* and one linked to *carA* (Martinez-Laborda *et al.*, 1986). Another type of transposon insertion mutant was isolated (Balsalobre *et al.*, 1987) that prevented carotenoid accumulation in both light and dark conditions. This mutation mapped to a locus linked to *carA* and later termed *crtEBDC* (*carB*). Tn*5* *lac* carries a promoterless *lacZ* gene within Tn*5* and when integrated into the chromosome provide in-situ information on the transcriptional activity of the region around the transposon. Tn5 lac insertions at *crtEBDC* (*carB*) showed that the *crtEBDC* (*carB*) locus is transcribed in the light only, in wild-type, but both in the light and dark in constitutive *carA* and *carR* mutants. This suggested expression of *crtEBDC* (*carB*) occured from a light-inducible promoter which is regulated by the products of the *carA* and *carR* loci. Tn*5* *lac* insertions later also identified the *crtI* (*carC*) locus as a region absolutely required for carotenogenesis. The absolute requirement for *crtI* (*carC*) and *crtEBDC* (*carB*) implied that either these genes encode structural enzymes for carotenogenesis, or positive regulators of the structural genes.

1.6.1.2. Genetic dissection of the *carR* region.

A strain containing a transposon insertion at Ω1910 exhibited a dark yellow phenotype, although not as severe as a Car^c phenotype. D. A. Hodgson (pers. comm.) showed that the transposon insertion was a dominant effect and led to the postulation of a light-activated promoter upstream of an activator of carotenogenesis, which is activated by readthrough of an outward reading promoter from within the transposon. Transduction experiments showed that Ω1910 mapped to the *carR* region.

Screening *carR* mutants for spontaneous mutants which possessed a car^- phenotype, gave four strains which carried mutations epistatic over a *carR* lesion (Martinez-Laborda and Murillo, 1989). One lesion mapped to *crtEBDC* (*carB*) whereas the other three mutations were linked to *carR*. The three mutations at *carR* all abolished expression of Tn*5 lac* inserted at *crtEBDC* (*carB*). Thus linked to *carR* exists a region which is epistatic over constitutive mutations at *carR* for the production of carotenoids. A likely explanation would be that an element at *carR* has a positive role in expression of the *car* regulon, but its activity is repressed or inhibited by CarR in the dark.

Sequencing of the entire *carR* region revealed the presence of three translationally coupled genes: *carQ*, *carR* and *carS* (McGowan *et al.*, 1993). Similarity searches revealed that CarQ was predicted to be a sigma factor, CarR an integral transmembrane protein and CarS had no known homologues.

The regulation of the *carQRS* locus is highly dependent on the stoichiometry of CarR to CarQ (Gorham *et al.*, 1996). If *carQ* is present in greater copy number than *carR* then carotenogenesis becomes constitutive. This also results if the translational coupling between *carQ* and *carR* is relieved. If however *carR* is present in greater copy number than *carQ* then carotenogenesis remains light-induced.

It was also observed that CarR:β-galactosidase fusions expressed in *M. xanthus* disappear in the light and that a CarR-Protein A fusion in *E. coli* was membrane

located (Gorham *et al.*, 1996). The conclusion drawn was that CarR acts as an anti-sigma factor sequestering CarQ at the membrane in the dark, but releasing it in the light due to loss of active CarR. Deletion mutants showed that CarQ is a positive activator of the *carQRS* and *crtI* (*carC*) promoters and that CarS is responsible for activation of the *crtEBDC* (*carB*) locus.

1.6.1.3. Localisation of the Car structural genes.

The next step was to analyse the nature of the carotenoids produced by strains carrying different Car mutations (Martinez-Laborda and Murillo, 1989). It was found that *carR* strains produced the same carotenoids in the dark and in the light as the wild-type did in the light. The *carA* mutant however accumulated mainly phytoene in the dark, but wild-type carotenoids in the light. Since *crtEBDC* (*carB*) mutants are unable to produce phytoene and strains carrying a *crtI* (*carC*) lesion accumulate only phytoene, it was deduced that the enzymes required for phytoene production reside at *crtEBDC* (*carB*) and the gene encoding phytoene dehydrogenase is found at *crtI* (*carC*). This was subsequently confirmed by cloning and sequencing of the *crtI* (*carC*) gene (Fontes *et al.*, 1993). Expression studies using a *crtI::lacZ* transcriptional fusion showed that *crtI* (*carC*) expression was light induced and induction was maximal if the cells were in a stationary phase of growth (Fontes *et al.*, 1993). It was also noticed that in a strain carrying a *crtI* (*carC*) lesion in the light, more phytoene was produced than the total amount of carotenoids produced in a wild-type in the light. This was postulated to be due to feedback inhibition by end products (Martinez-Laborda and Murillo, 1989).

Subsequent experiments which introduced transposons into the *crtEBDC* (*carB*) region suggest the presence of multiple structural enzymes, and use of transcriptional fusions of promoterless *lacZ* to various sections of the *crtEBDC* (*carB*) cluster imply that the *crtEBDC* (*carB*) genes are transcribed from a single promoter which is *carQ* dependent and light inducible (Ruiz-Vazquez *et al.*, 1993). Sequencing of the complete *crtEBDC, carA* (*carBA*) region showed several interesting features (Botella *et al.*, 1995). The cluster contained eleven open

reading frames. Open reading frames one to six had good sequence similarity to genes for known carotenogenic enzymes and allowed a tentative assignment of function to these orfs. Genes *orf10* and *orf11* each contained a peptide motif which possessed homology to the DNA binding domains of the MerR family of transcriptional regulators while *orf7, orf8* and *orf9* had no homologues in the databases. The mutated *carA1* allele found in strain MR7 was found to differ from the wild-type in two places – a G deletion at the end of *orf9* and an A-T transversion at the start of *orf10*. It was interesting to note however that the *crtEBDC* (*carB*) promoter showed no homology whatsoever to the *carQRS* promoter, implying its expression is governed by a sigma factor other than CarQ. It is thought that there are actually two promoters within the *carBA* cluster. The initial promoter is upstream of an operon consisting of the first six open reading frames of the cluster, designated the *crtEBDC* operon and is probably transcribed by a vegetative promoter. The last five open reading frames of the cluster form a second operon (the *carA* operon) which appears to be transcribed from a Sig54-dependent promoter (Cervantes and Murillo, 1998).

1.6.2. Current understanding of the regulon.

The carotenogenic system involves the products of three unlinked genetic loci, which are coordinately expressed as a regulon, of which, the *carQRS* operon is the central regulatory locus. Also implicated in carotenogenesis are two genes, *carD* and *carE* which are not exclusively devoted to the carotenogenic regulon.

1.6.2.1. The *carQRS* operon:

The *carQRS* operon contains three translationally coupled genes under the control of a light-inducible promoter (McGowan *et al.*, 1993). CarQ activates expression of the *carQRS* operon and the *crtI* gene in a light-dependent fashion, with mutations in *carQ* yielding a Car^- phenotype. The CarQ gene product has significant sequence homology to a family of sigma-factors with proposed extra-

cytoplasmic functions (ECF-sigma factors) and has recently been proven to have *in vitro* sigma factor activity (Browning, 1997). Examples of this class of sigma-factor include AlgU, recently renamed AlgT, (alginate synthesis) in *Pseudomonas aeruginosa* (Martin *et al.*, 1993) and SigE (agarase expression) in *Streptomyces coelicolor* (Buttner, 1989). Comparison with σ^{70} of *E. coli* shows that CarQ lacks region 1 while region 3 is severely shortened. The 2.4 and 4.2 regions which are the proposed DNA-binding regions for the -10 and -35 sites respectively are still present (Dombroski *et al.*, 1992; Dombroski, 1997). This suggests that CarQ activates transcription from P^{carQRS} and P^{crtI} through alteration of the specificity of the cellular RNA polymerase from its housekeeping mode.

The second gene, *carR*, codes for a protein that is thought to act as an anti-CarQ factor. Its sequence corresponds to an integral membrane protein, predicted by hydropathy analysis to contain six trans-membrane helices and a large N-terminal cytoplasmic domain (McGowan *et. al.*, 1993). It has recently been proven to be present in the inner membrane when expressed in *Escherichia coli*, and present only in the dark and only when over-expressed, in *M. xanthus* (Browning, 1997). Mutations in *carR* give constitutive carotenogenesis (McGowan *et al.*, 1993) and constitutive expression of the *carQRS* promoter and the *crtI* promoter.

The final gene of the operon codes for a protein (CarS) that activates expression of the *crtEBDC* gene cluster. Deletion of *carS* results in a Car^- phenotype with a lack of expression of the *crtEBDC* cluster. The *carQRS* operon is still light-inducible however, which shows that CarS is a positive regulator of *crtEBDC*, but not *carQRS*. It seems that CarS is also a general inhibitor of induction, however, as deletion of *carS* gives three times greater expression from the *carQRS* promoter than in the wild-type. CarS exhibits no homology to any known protein sequences, and possesses no recognisable DNA-binding motifs. The nature of its activation of *crtEBDC* is obscure.

1.6.2.2. The *crtEBDC* cluster:

The *crtEBDC* cluster contains 11 open reading frames arranged sequentially over 12kb. The first six ORFs are structural genes for early and late stages in the biosynthesis of carotenoids (Botella *et al.*, 1995). The *crtB* gene itself is one of these structural genes, encoding a phytoene synthetase enzyme. Expression of the *crtEBDC* promoter is dependent on CarS (Balsalobre *et al.*, 1987; Martinez-Labordez and Murillo, 1989), and is therefore induced in the light.

The other five ORFs are of unknown function (Ruiz-Vasquez *et al.*, 1993), but include the genes which are mutated in *carA* strains. CarA represses the *crtEBDC* operon in the dark (Balsalobre *et al.*, 1987) and *carA* mutants are therefore Car^c due to loss of the *crtEBDC* repression. They accumulate large amounts of colourless phytoene in the dark with a small level of red pigments arising from desaturation of the accumulated precursor. This dehydrogenase activity may reside within *crtEBDC*. The *carA* lesion maps to the 3' end of the cluster and a *carA* mutant was found to have mutations in both *orf9* (a frame-shift affecting *orf9*, *orf10* and *orf11*) and *orf10* (an A/T transversion). Genes *orf10* and *orf11* contain putative helix-turn-helix motifs and this supports the idea that CarA acts as a transcriptional regulator.

1.6.2.3. The *crtI* gene:

From homologies to bacterial dehydrogenases, the *crtI* gene is suggested to code for the enzyme phytoene dehydrogenase (Balsalobre *et al.*, 1987). Homology was only moderate (30-40% overall identity) but was very strong in two domains, one at the C-terminal and the other at the N-terminal. It is thought to catalyse an essential step in the pathway for the production of photoactive carotenoids:- the conversion of colourless phytoene into coloured neurosporene and lycopene. Strains mutant in *crtI*, are Car^- and show accumulation of phytoene in the light. Induction of *crtI* is dependent on *carQ* and is therefore light-induced. Expression of *crtI* seems additionally to be activated by carbon limitation (Fontes *et al.*, 1993). This may be a consequence of the CarD requirement (see below).

Sequence analysis of the *crtI* promoter shows similarities to typical promoters of Gram negative bacteria. Homology of the -35 region is generally strong while that for the -10 region is much weaker. Interestingly, promoters which exhibit the greatest homology to the -10 region include two promoters from genes involved with multicellular development and a promoter from a vegetative gene all in *M. xanthus* (Fontes *et al.*, 1993; Figure 1-8). This reinforces a possible link between *crtI* expression and development, mediated by CarD.

crtI promoter	TCTTGTAACGTCCTGGCGGGTTCGCGCGTTCGCCAGG**T**	
vegA promoter	TAGACA	AAGGGT
tps promoter	TTGCAT	AATGCT
ops promoter	TTGCTC	CAAGCT

Figure 1-8 Alignment between the crtI, vegA, tps *and* ops *promoters. The transcriptional start site is shown in bold face.*

1.6.2.4. The *carD* gene:

A CarD mutant was initially isolated as a Car$^-$ Tn*5* insertion that was incapable of aggregation and fruiting body formation (Fru$^-$). This mutant also had blocked expression from *carQRS*, *crtEBDC* and *crtI* (Nicolas *et al.*, 1994). The CarD protein appears to modulate the activation of the *carQRS* and *crtI* operons by CarQ and also seems involved in expression of the A-factor and C-factor-dependent classes of developmentally activated genes of the *M. xanthus* fruiting body response.

The sequence of CarD is related to the eukaryotic HMGI(Y) family of transcriptional activators. Its C-terminal sequence contains four copies of a DNA-binding domain shared with these proteins (RGRP - Arg-Gly-Arg-Pro), which are known to enhance binding to the minor groove at A-T rich regions. These domains are adjacent to an acidic region which resembles a consensus serine phosphorylation site. For HMGI, phosphorylation has been shown to alter DNA-binding activity (Nicolas *et al.*, 1994). The *carQRS* promoter contains a tandem

repeat of TTTCC which is a similar sequence motif to the binding site of HMGI)Y) at the IFN-β promoter. Nicolas *et al.* (1996) showed that *in vitro,* CarD could bind specifically to this repeated motif within P^{carQRS}. At the IFN-β promoter HMGI(Y) mediates displacement of H1 histone by binding to nucleosomes. Binding to the IFN-β promoter changes bending within the promoter, allowing access by transcription factors. CarD also has a leucine zipper domain which may allow dimerisation or interaction with other DNA binding factors.

1.6.2.5. The *carE* locus:

This locus has recently been identified through a Tn*5* insertion which abolishes carotenoid production and gives pink colonies rather than red. Phenotypic instability in the mutant yields red colonies at a low frequency. The red and pink forms of the mutant have been shown by Southern blotting to differ in the orientation of the Tn*5* insert, but not in its location (Murillo, pers. comm.). Cloning and sequencing showed the locus encoded integration host factor (*ihf*). Subsequent Southern blotting showed that *M. xanthus* actually contains two copies of *ihf,* a gene involved with the control of gene expression through DNA bending. Its role in expression of the carotenogenic regulon is almost certainly a consequence of its general role in regulating DNA topology.

1.6.3. Current Model

The model as it presently stands is shown in Figure 1-9. In the dark grown cell of *M. xanthus* the transmembrane protein CarR binds to CarQ and holds it in an inactive state at the cell membrane. Light acting on the cell excites the photosensitiser protoporphyrin IX which in turn causes the excitation of molecular oxygen to the excited singlet state. This relatively stable species somehow mediates the inactivation of CarR. This could either be by direct damaging oxidation of the CarR protein or by the activation of a specific protease which

then removes CarR enzymatically. The loss of functionally active CarR releases CarQ from the membrane which is then free to activate transcription of the *carQRS* and *crtI* promoters. Expression of the *carQRS* operon causes the expression of CarQ, CarR and CarS. CarR is still removed by the action of light, so the net effect is the production of CarQ and CarS. CarQ activates transcription of *crtI* gene and CarS relieves the repression of the *crtEBDC* cluster by CarA. Expression of *crtI* and *crtEBDC* produces the carotenogenic enzymes which produce carotenoids. The carotenoids produced, by a process akin to feedback inhibition, quench the excited species responsible for transducing the light signal from light to the CarR/CarQ switch. Thus CarR is no longer inactivated and once more removes CarQ to the inner membrane. The down-regulation of the regulon therefore results.

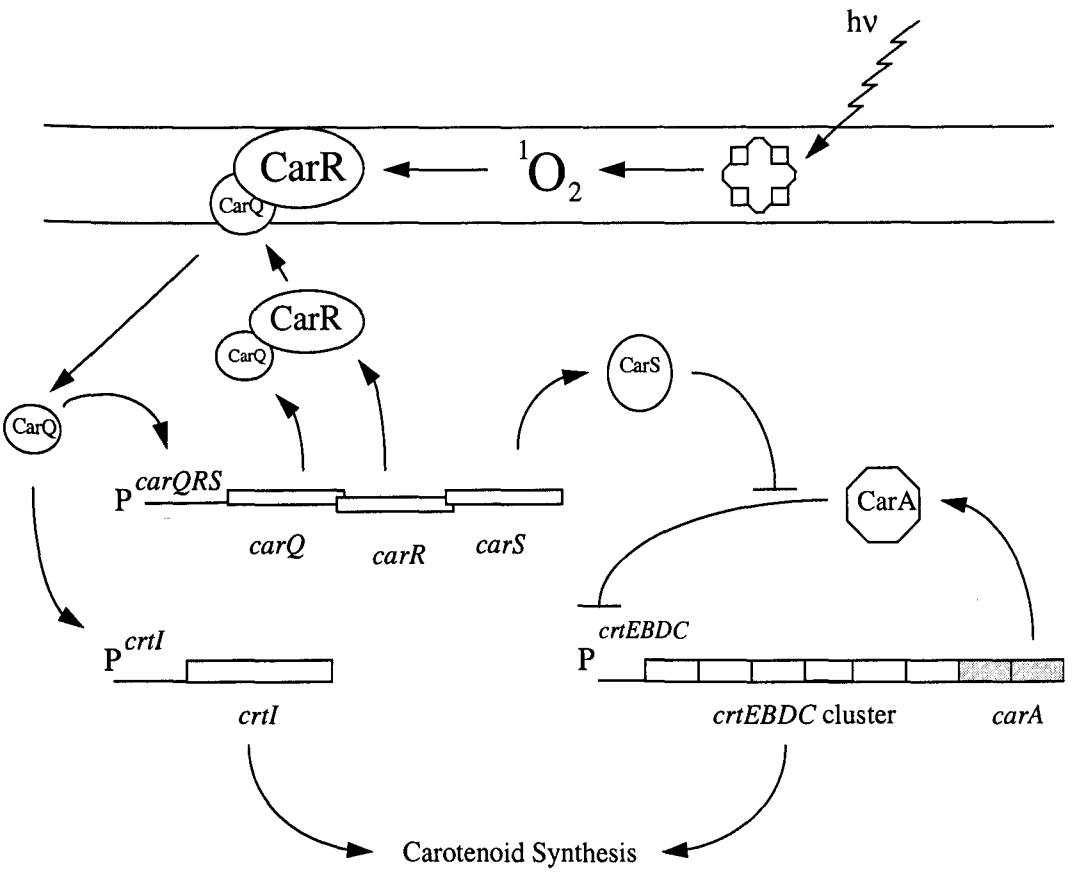


Figure 1-9 Regulation of carotenogenesis in M. xanthus.

Events subsequent to illumination are depicted.

1.7. Aims

Fontes *et al.* (1993) have assigned CarA with a positive regulatory role in expression of the *crtI* and *crtEBDC* operon due to a reduction in activity of these loci in a *carA* mutant. Whether this is a correct interpretation of the data is under doubt, as *carA* mutants are Car^c and the reduction in promoter activities may be due to the constitutive production of carotenoids. The initial goal of this project is to clarify the role of CarA within the carotenogenic regulon.

The crucial event in the response of *M. xanthus* to light, is expression from the *carQRS* promoter mediated by CarQ. Berry (1998), has constructed a set of site-directed mutants of the *carQRS* promoter. The mutant promoters will be be introduced into *M. xanthus* followed by assessment of *in vivo* activity. This may allow dissection of the interactions between the ECF sigma factor CarQ and its cognate promoter.

An expression protocol for the production of active CarQ has been devised by Browning (1997), which allows the use of a range of *in vitro* techniques for the study of CarQ-promoter interactions. The applicability of *in vitro* studies to the carotenogenic system of *M. xanthus* will be determined and discussed.

2. The roles of *carA*.

Sequence analysis of the *carA* region revealed 5 ORFs after *crtEBDC* that had no clear roles (*orfs* 7-11). It was observed that *orfs* 10 and 11 each contained a helix-turn-helix motif, implicating an involvement in DNA binding as expected for CarA. To date only a single mutant allele of *carA* has been isolated. The mutant allele contained two differences from the wild-type. The first was a deletion of a G at the end of *orf9* causing a frameshift, while the second was an A-T transversion at the start of *orf10* (Botella *et al.*, 1995).

Recent experiments using precise in-frame deletions of the ORFs at the *carA* region show that *orf10* alone causes the repression of *crtEBDC* in the dark (Cervantes and Murillo, 1998). Deleting *orf11* gave no discernable phenotype. GST fusions to the products of *orfs* 10 and 11 each bound weakly to the *crtEBDC* promoter. There is also some speculation on a possible involvement of vitamin B12 in the action of CarA as both *orf10* and *orf11* have been postulated to contain vitamin B12 binding domains, based on homology of the C-termini of these proteins to metal-cobalamin binding sites of methionine synthases (Cervantes and Murillo, 1998). The significance of these findings are obscure. Supplementing growth medium with vitamin B12 did enhance light-induced carotenoid production dose dependently, but a plausible model for the role of cobalamin co-factors in the function of CarA has not yet been constructed.

The precise molecular nature of CarA and how it functions still remains obscure.

It has been observed from the introduction of promoter probes (plasmids containing *car* promoters fused to a promoterless *lacZ* gene) into *carA* strains, that there is much lower expression of *carQRS* and *crtI* in a *carA* background than in the wild-type. Therefore *carA* has been assigned with a positive regulatory function on *crtI* and *carQRS* expression (Fontes *et al.*, 1993). Robson and Hodgson (personal communication) have also shown that the *carQRS* promoter is poorly activated by light in a mutant *carA* background.

An alternative explanation for this phenomenon other than a direct activating role for CarA is that the carotenoids produced constitutively in *carA* strains quench the high-energy intermediates in the light-induction process. This would effectively reduce the strength of the incident light stimulus and down-regulate the expression of the carotenogenic genes. The same argument is also applicable to the expression of *crtI* and CarA's exact activity is therefore uncertain. Is it a positive regulator of *crtI* and *carQRS* expression or not? This work aims to clarify its function in this respect.

2.1. Introduction.

In the wild-type *Myxococcus xanthus* cell in the light the most abundant carotenoids are the final carotenoids of the biosynthetic pathway. In a *carA* mutant in the dark the most abundant carotenoids are phytoene and phytofluene with small levels of final carotenoids sufficient to grant a Car^c phenotype. On illumination the carotenoids produced are those found in wild-type cells in the light. A *crtI* mutant accumulates only phytoene and phytofluene in both the light and the dark (Martinez-Laborda *et al.*, 1990). Phytoene and phytofluene are colourless and hence *crtI* mutants appear Car^-.

The nature of feedback control within the carotenogenic regulon has been mentioned in the literature, particularly with respect to the down-regulation of the regulon's activity by negative feedback on production of carotenoids (Hodgson, 1993). Other feedback mechanisms involving end-product inhibition of the carotenogenic enzymes have also been hinted at (Martinez-Laborda *et al.*, 1990) inspired by the situation in *Phycomyces blakesleeanus*. However, the effects of feedback due to constitutive production of carotenoids, or the lack of feedback regulation in a Car^- mutant are areas that until now appear to have been neglected. As well as assessing the role of CarA, the effect of Car^c and Car^- mutations on activity of various promoters within the regulon will be determined.

2.2. Does CarA activate P^{crtI} or P^{carQRS}?

Plasmids containing copies of the *crtI* (pMAR206) and *carQRS* (pDAH217) promoters fused upstream of a promoterless *lacZ* gene were mobilized into strains of *M. xanthus* by P1-mediated transduction (see section 7.2.9). On introduction into *M. xanthus* the plasmids are not replicated autonomously and the appropriate drug resistance is not maintained unless they integrate into the chromosome (Shimkets *et al.*, 1983). Integration occurs by homologous recombination across the promoter on the plasmid and the endogenous copy, creating a merodiploid with a tandemly duplicated promoter. One of the copies of the promoter is upstream of its normal transcriptional unit and the second copy is fused upstream of the promoterless *lacZ* gene. Thus as the wild-type gene is not disrupted in any way, the regulon functions exactly as if in the same strain without the promoter probe and the *lacZ* gene is transcribed exactly as the genes downstream of the endogenous promoter. Thus the *in-vivo* activity of the *crtI* and *carQRS* promoters can be determined accurately by assaying the production of β-galactosidase in a strain carrying either pMAR206 or pDAH217.

Using these transcriptional fusions as probes for the activity of the *crtI* and *carQRS* promoters, promoter activities were assessed in different mutant backgrounds. Figure 2-1 shows the activity of a wild-type strain (DK101) into which was integrated pDAH217 as a probe for the *carQRS* promoter, determined by β-galactosidase production. (N.B. In all β-galactosidase assays, at Time=0 a dark grown early exponential phase culture is split into two daughter cultures, one of which is illuminated and the other maintained in the dark. Samples are periodically removed for β-galactosidase and protein assays allowing a time course of β-galactosidase specific activity to be calculated (See section 7.3.1). β-galactosidase assays were performed at least in duplicate and results shown only if both curves were consistent.

After an initial lag period of ~2 hours induction occurs in the light with specific activity rising from 10 units/min/mg protein to 400 units/min/mg protein after 6 hours. The response is then down-regulated and this is due to the production of carotenoids which quench the signal from light and reduce the strength of the

stimulus affecting the CarQ/CarR switch. In the dark a background level of activity of ~10 units/min/mg protein is constant.

When plasmid pDAH217 was introduced into a strain carrying a *carA* lesion (DK717) and the transductant assayed for β-galactosidase expression a curve was obtained which was very different from DK101::pDAH217 (Figure 2-2).

Induction in the light still occurs with a similar lag time but maximal induction gives only 40 units/min/mg protein instead of 400 as found for the wild-type. Thus functional CarA is required for full induction of *carQRS* expression on illumination and this has also been found to be the case for *crtI* expression.

Plasmid pMAR206 was transduced into wild-type DK101 and the resulting integrants were assayed as before for β-galactosidase activity (Figure 2-3).

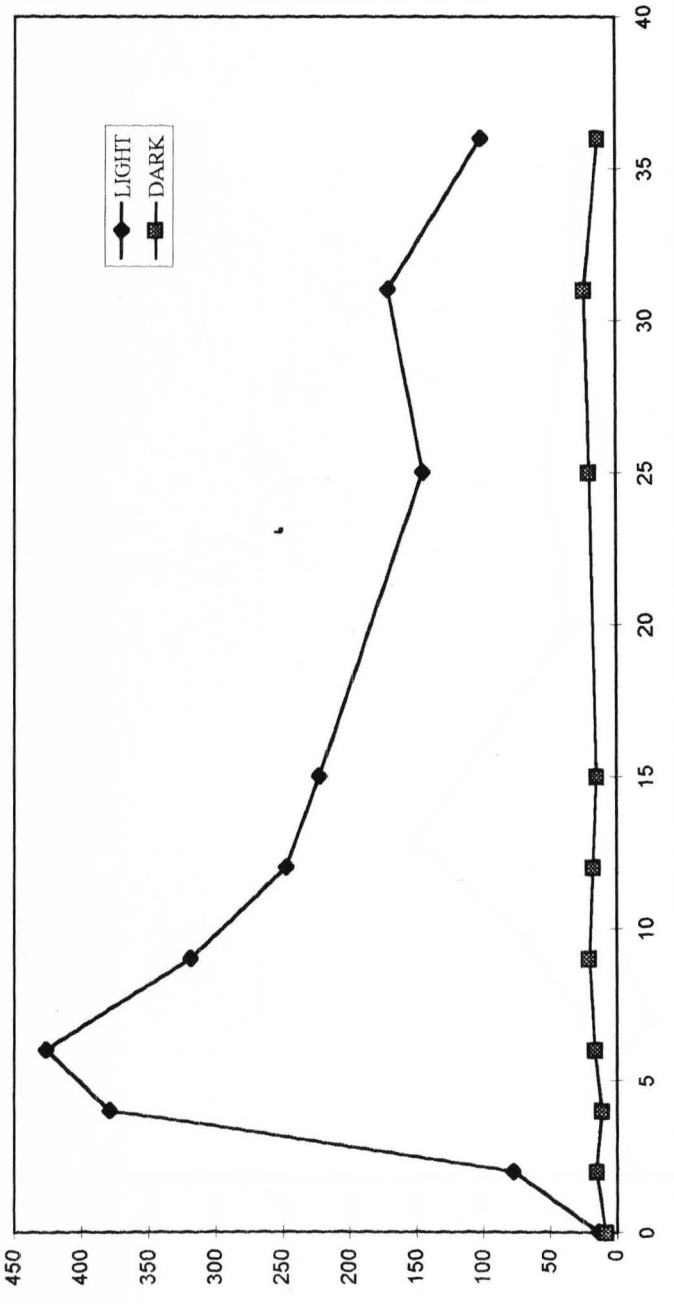


Figure 2-1 Activity of the carQRS *promoter in a wild-type strain (DK101::pDAH217).*

Y-axis is specific activity of β-galactosidase in units/min/mg protein. X-axis is time in hours

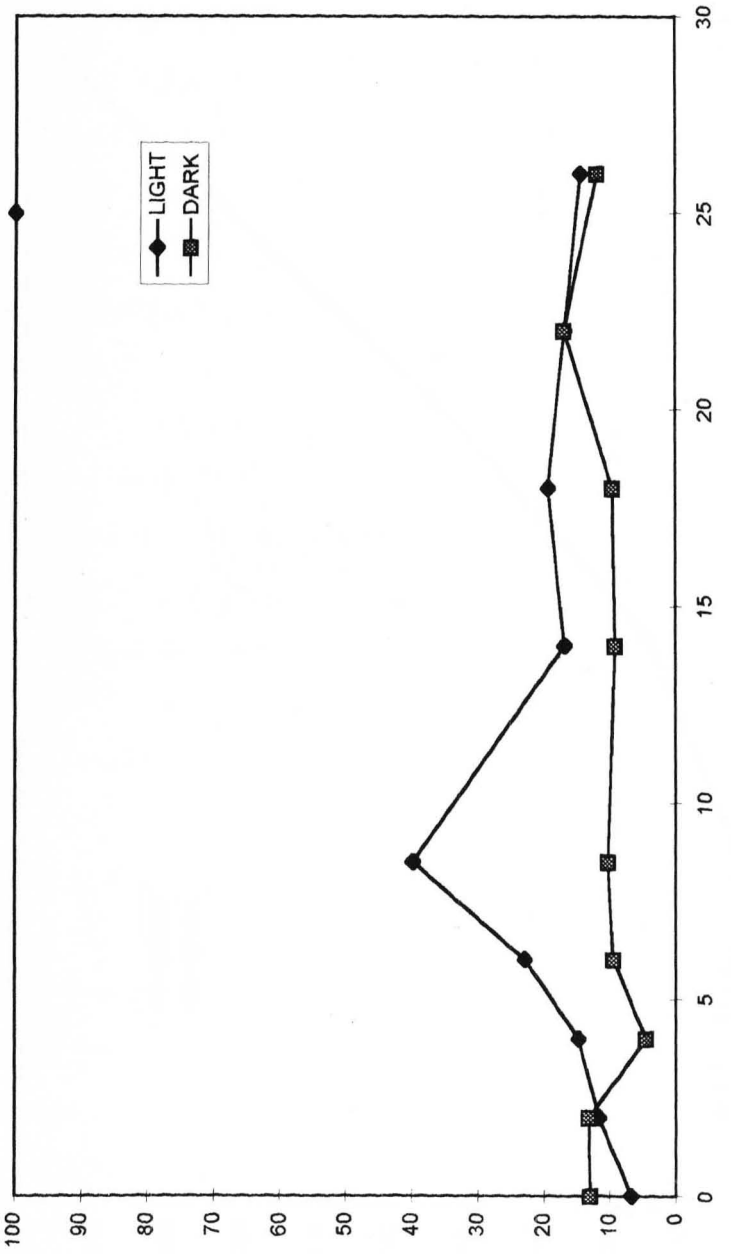


Figure 2-2 Activity of the carQRS *promoter in a* carA *strain (DK717::pDAH217)*

Y-axis is specific activity of β-galactosidase in units/min/mg protein. X-axis is time in hours

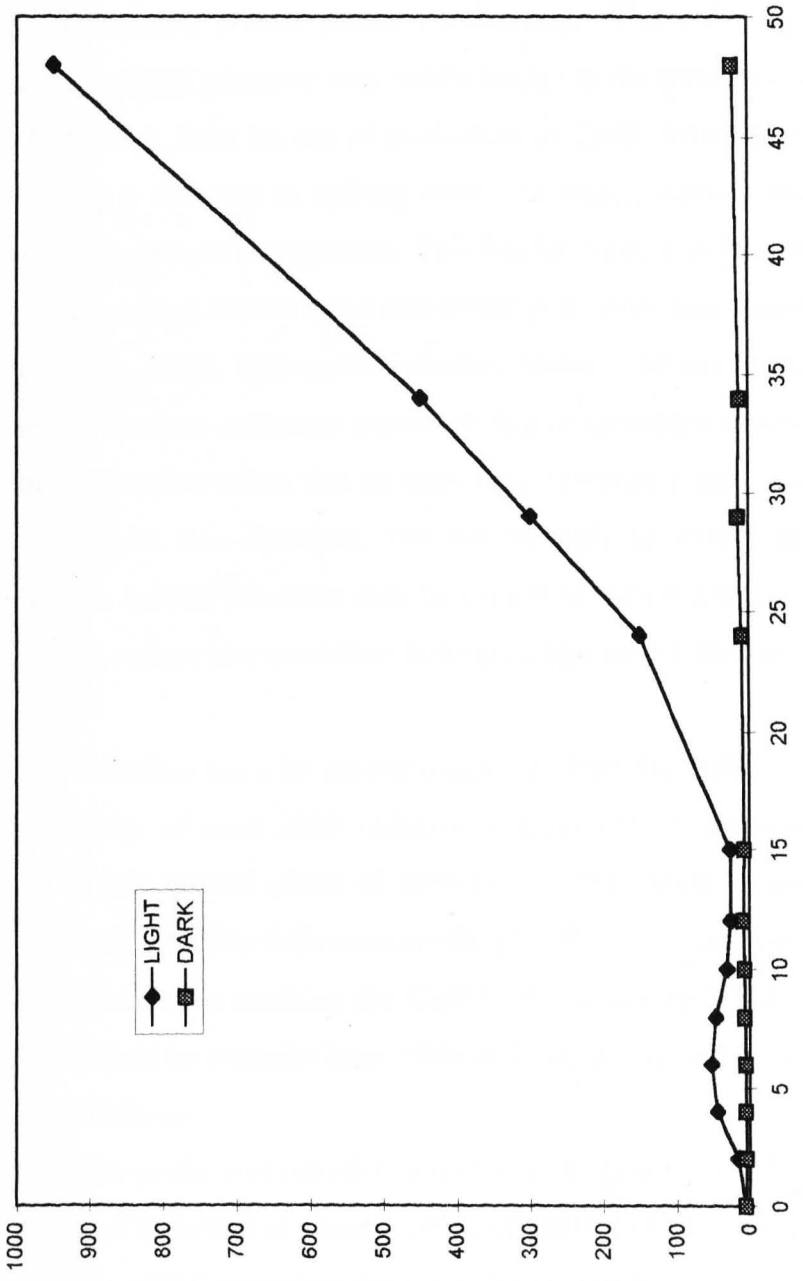


Figure 2-3 Activity of the crtI *promoter in a wild-type strain (DK101::pMAR206)*

Y-axis is specific activity of β-galactosidase in units/min/mg protein. X-axis is time in hours

After an initial lag period induction occurs in the light with a maximum activity at 6 hours with 50 units/min/mg protein (Phase I induction). The lower activity when compared to the *carQRS* promoter may purely be due to the greater distance separating the *crtI* promoter from the site of production of CarQ. After the initial induction there follows a decrease in activity until ~18 hours, when a second phase of induction occurs (Phase II induction). This has been previously reported and the second phase has been shown to be coincident with entry into stationary growth phase (McGowan, 1992). The second induction seems to be due to lack of carbon source from the medium and hence represents loss of catabolite repression. Consistent with this is the observation that on entry into stationary phase, activity of the dark-grown culture also increases, but not to such an extent as the illuminated culture. The second induction may be caused to start prematurely by resuspending cells from culture into a medium lacking carbon source such as MC7 (Fontes *et al.*, 1993).

The second phase of induction has a far greater magnitude than the first induction with a maximum activity of over 2000 units/min/mg protein. A theoretically problematic aspect of this second phase of induction is that cultures entering stationary phase are already replete with carotenoids. Therefore negative feedback should be reducing the stimulus reaching the CarR/CarQ switch such that light-dependent induction should be virtually impossible and yet, the second activation phase is strongly light induced.

When the same promoter probe was introduced into a *carA* mutant (DK717) and assayed, the first phase of induction is almost completely absent (Figure 2-4).

Thus both the *crtI* and *carQRS* promoters show a reduced activity in *carA* strains relative to the wild-type. This may be purely because *crtI* expression requires a product of the *carQRS* operon or alternatively because of a direct effect on P^{crtI}. Either CarA is an activator of these two promoters or the carotenoids produced constitutively by the *carA* mutant are causing feedback inhibition of the carotenogenic response. To determine which explanation was correct it was necessary to assess the activity of the two promoters in a strain carrying a *carA*

lesion which was also not producing carotenoids. Two experimental approaches were employed to create such a strain, one genetic and the other biochemical.

2.2.1. Genetic Approach

Creating a mutation in the *crtI* structural gene prevents the production of end-product carotenoids since the strain can no longer produce phytoene dehydrogenase. A *crtI* mutation should also have no further effect on other genes within the regulon as CrtI appears to have no regulatory role or downstream effects. Introduction of *crtI* lesions into various strains was achieved by the generalised transduction of a *crtI* allele containing a tetracycline resistance-confering Tn*5*-132 insertion from MR461 (Fontes *et al.*, 1993) into strains of interest. Transductants would be tetracycline resistant if a transducing particle introduced the *crtI* region from MR461 into the cell and a double cross-over event had replaced the endogenous copy of *crtI* with the allele containing the disruptive transposon insertion. When the *crtI* lesion was introduced into any strain, the transductants exhibited a Car^- phenotype.

As a control, the *crtI* lesion was first introduced into the carA mutant DK717 which carried no promoter probe. The resulting strain UWM501 gave background levels of β-galactosidase expression in both light and dark (Figure 2-5). It was necessary to perform this control as the Tn*5*-132 insertion within the *crtI* allele contains a *lacZ* gene and the background transcription level of this gene needed to be determined. The *lacZ* gene is in the incorrect orientation for readthrough transcription to occur from the *crtI* promoter and thus its expression is almost indetectable through β-galactosidase assays.

The disrupted *crtI* allele was then introduced into two different wild-type strains, DK101 and DK1050, each containing pDAH217 to create UWM502::pDAH217 and UWM503::pDAH217 respectively. When each strain was assayed for β-galactosidase activity the curves produced were virtually identical (Figure 2-6 and Figure 2-7).

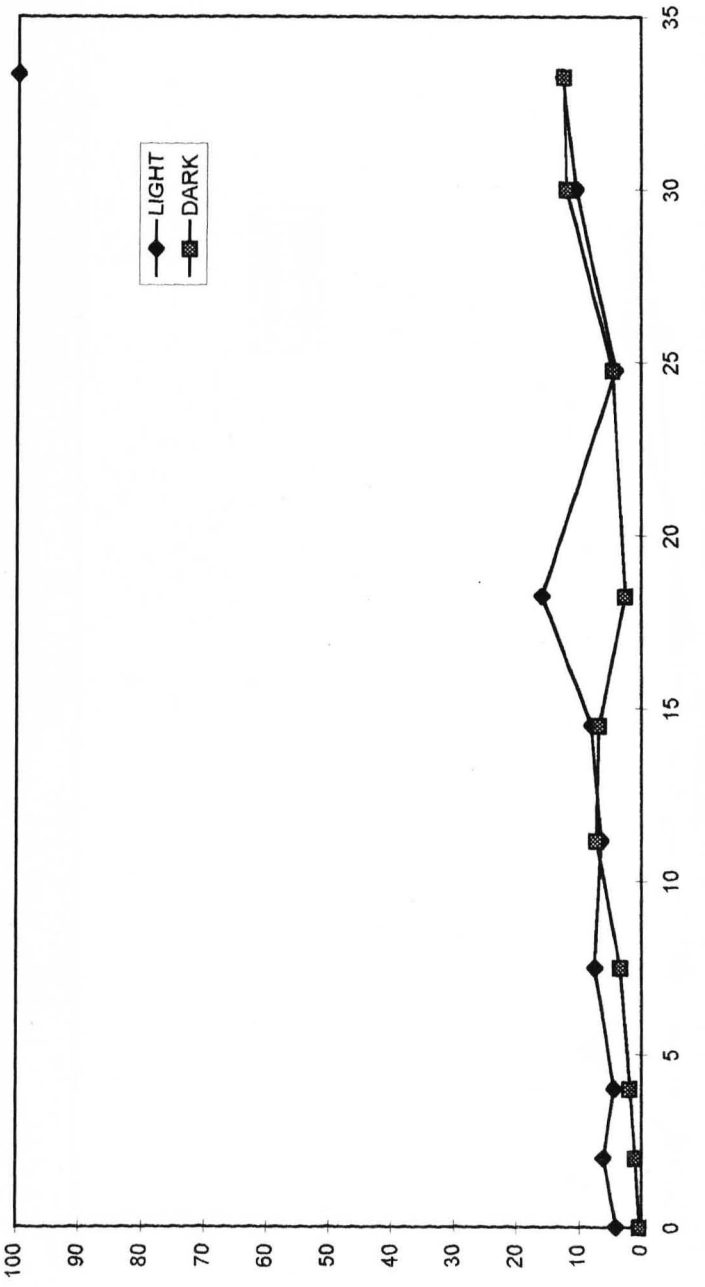


Figure 2-4 Activity of the crtI *promoter in a* carA *mutant (DK717::pMAR206).*

X-axis is specific activity of β-galactosidase in units/min/mg protein. Y-axis is time in hours

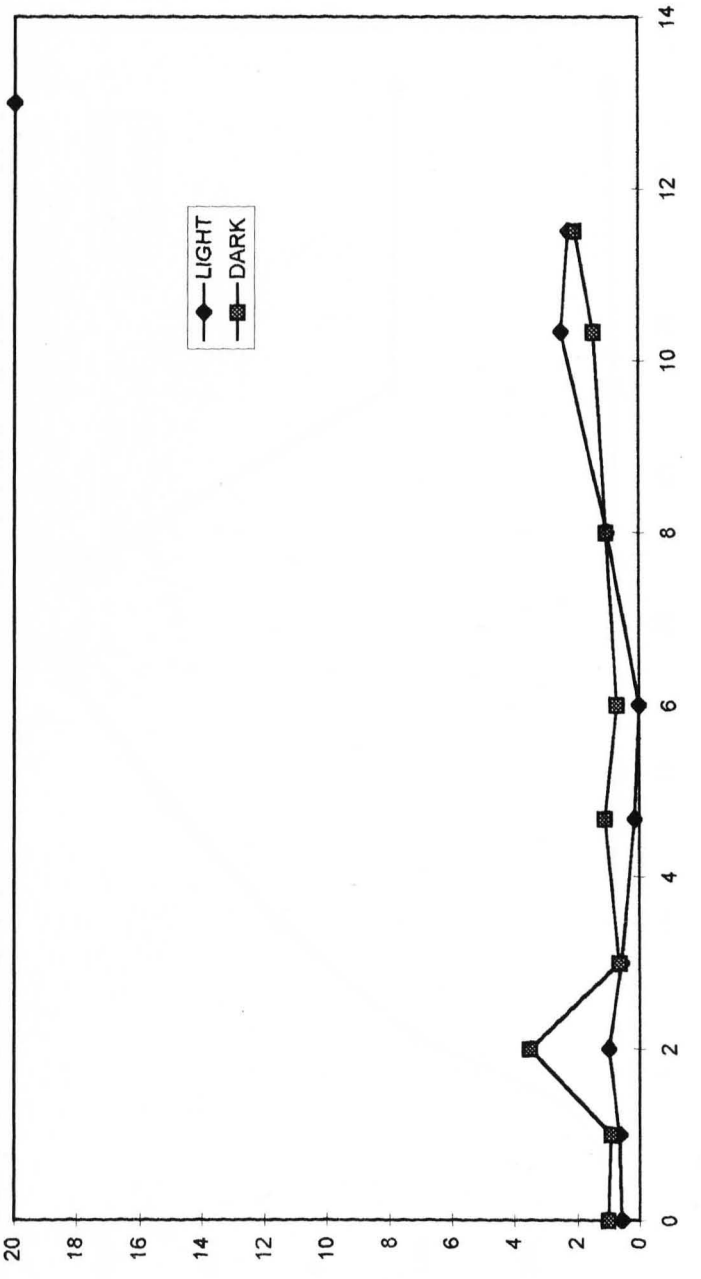


Figure 2-5 β-galactosidase expression of a crtI/carA *double mutant (UWM501).*

Y-axis is specific activity of β-galactosidase in units/min/mg protein. X-axis is time in hours

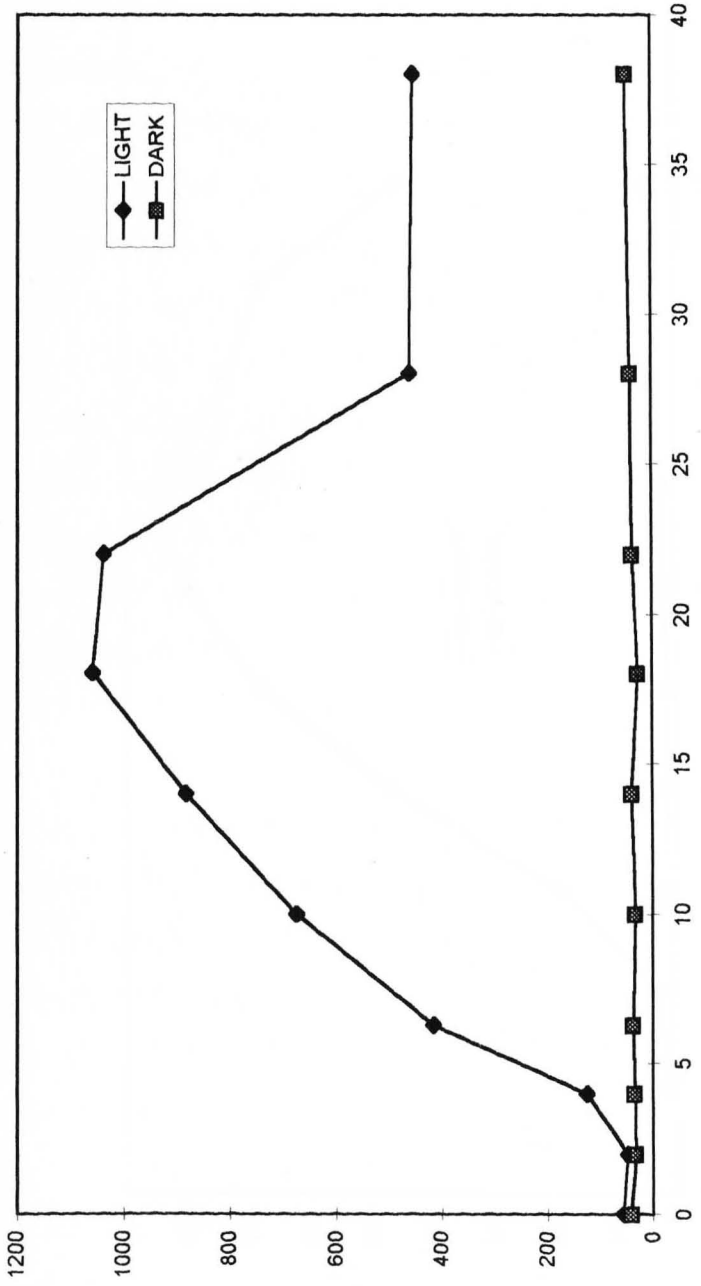


Figure 2-6 Activity of the carQRS *promoter in a* crtI mutant (UWM502::pDAH217).

Y-axis is specific activity of β-galactosidase in units/min/mg protein. X-axis is time in hours

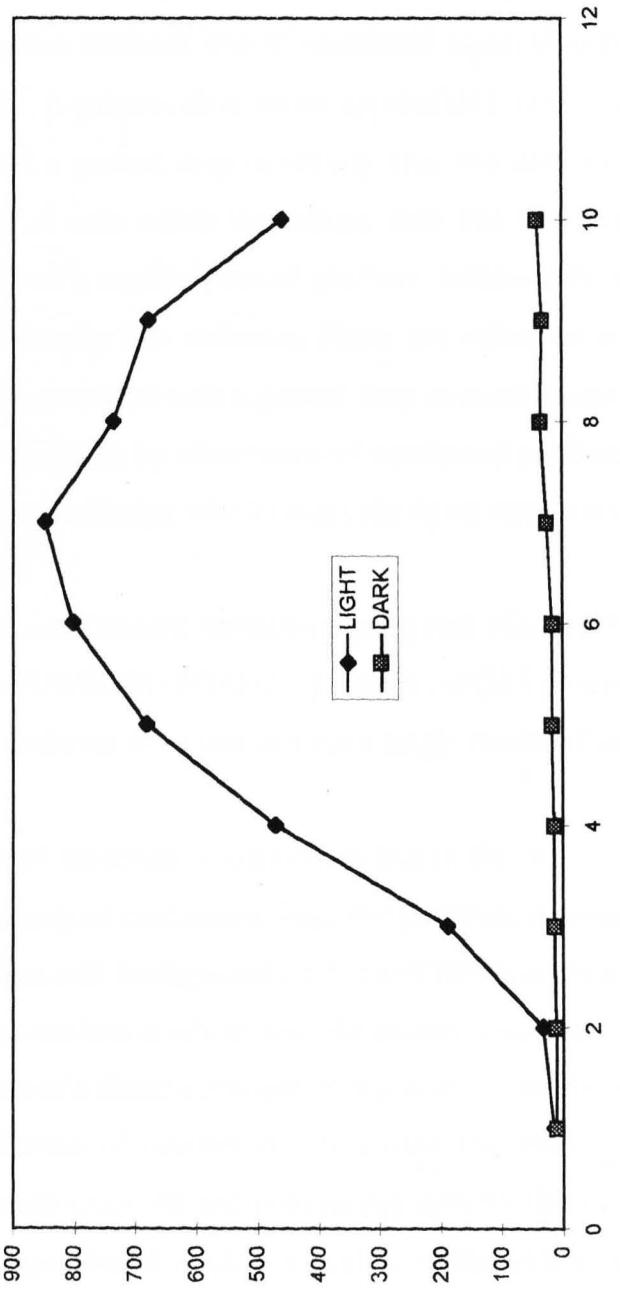


Figure 2-7 Activity of the carQRS *promoter in a* crtI mutant (UWM503::pDAH217).

Y-axis is specific activity of β-galactosidase in units/min/mg protein. X-axis is time in hours

In each case, induction occurs with exactly the same rate and timing as in the wild-type strain. However, negative feedback due to carotenoid accumulation is not seen and at maximal activity, β-galactosidase levels approaching 1,000 units are seen after ~15 hours. There is a gradual drop in activity after this time which may be due to the large number of cells within the culture, such that on average each cell in the population recieves a smaller dose of photons. Additionally, the drop in activity is consistent with entry into stationary phase and reduction in β-galactosidase expression may be associated with a general drop in macromolecule synthesis. The loss of negative feedback by elimination of carotenoid production proves that it is indeed carotenoid production which causes the down-regulation of *carQRS* activity in the wild-type.

Addition of the *crtI* lesion to a *carA* mutant containing integrated pDAH217 to form a *crtI/carA* double mutant (UWM501::pDAH217), shows *carQRS* promoter activity in that strain to be very different from that in a *carA* single mutant (Figure 2-8).

In fact, the time course and rate of induction is identical to that in the *crtI* mutant described above, with large amounts of expression. Thus the presence or absence of a *carA* lesion has no effect in a *crtI* background i.e. the *crtI* lesion is epistatic over a *carA* mutation. Therefore the low levels of *carQRS* promoter activity seen in the *carA* mutant is not because of a direct activation by CarA but instead is due purely to the constitutive production of carotenoids. In a *carA* background the constitutively produced carotenoids cause the end-product negative feedback loop to be established in the cell irrespective of irradiation and thus exposure to light gives much lower induction. Conversely, in cells carrying the *crtI* lesion the feedback loop can never become established and the *carQRS* promoter is induced strongly throughout illumination.

As an interesting aside, it was noted that in all strains created by introducing the mutant *crtI* allele, a 'stable' tan phenotype was observed (phase variation). In unilluminated cells of *M. xanthus* there is a phenotypically unstable pigmentation. Vegetative cells are typically yellow due to production of an uncharacterised pigment, rough and swarming, with a proportion (1%) appearing tan, smooth and

mucoid. The pigment has an absorption maximum at 379 nm and is also found in *Myxococcus virescens* (Burchard *et al.*, 1977). Interconversion between tan and yellow phenotypes occurs at low frequency: yellow cells switch to tan at a rate of around 10^{-2} to 10^{-3} per cell per generation and tan switch to yellow at a much higher frequency. Pigmentation appears to be related to 'stresses' experienced by the bacteria as selection with antibiotics and entry into stationary phase seems to cause the cell to become tan (or selection occurs against yellow cells). However, UV exposure, elevated temperatures and exposure to mitomycin causes an increase in the proportion of yellow cells (Laue and Gill, 1995). Research on phase variation has been hindered by the inability to isolate phase-locked mutants. The significance of the 'stable' tan phenotype of cells carrying a *crtI* lesion may be an artifact due to the stresses involved in creation of the cells, however, repeatedly streaking to single colonies over a period of several months in the absence of selection never yielded yellow colonies. It is possible that CrtI is involved in the production of the yellow pigment of *M. xanthus*, if so, then it may provide a phase-locked mutant for further study of phase variation.

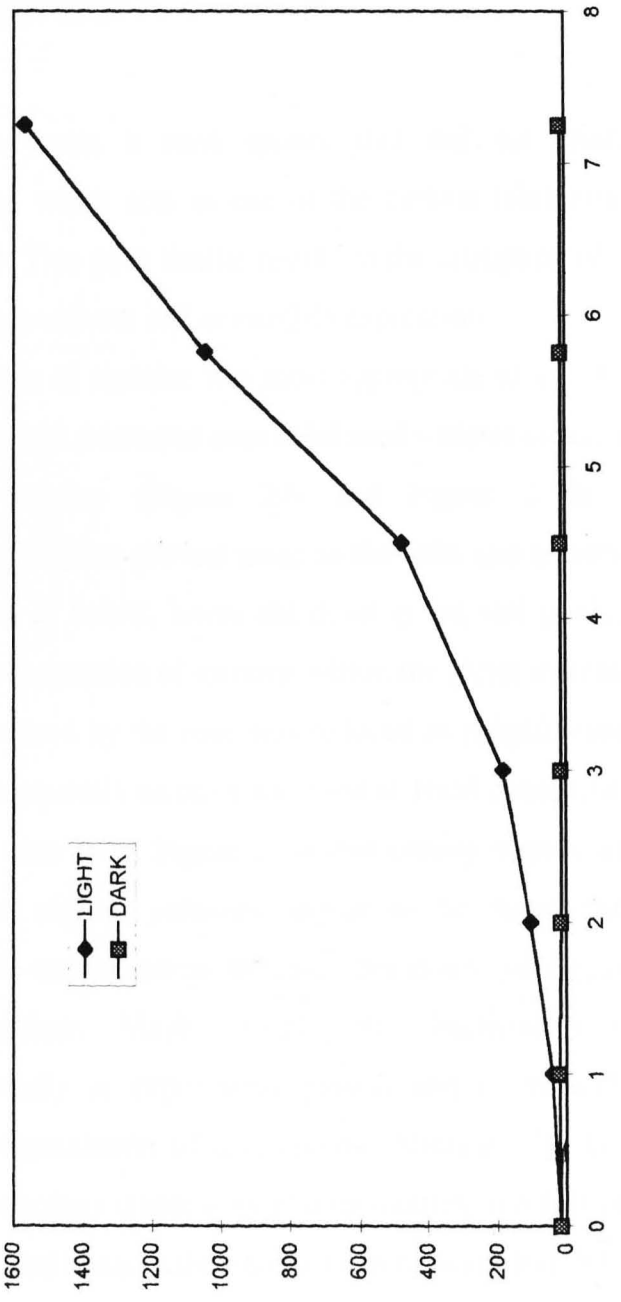


Figure 2-8 Activity of the carQRS *promoter in a* carA/crtIC *double mutant (UWM501::pDAH217).*

Y-axis is specific activity of β-galactosidase in units/min/mg protein. X-axis is time in hours

2.2.2. Biochemical Approach

The biochemical approach to create a *carA* mutant that did not produce carotenoids was to add nicotine which acts as one of the earliest inhibitors of carotenogenesis (Kleinig, 1974). This gave similar results to the disruption of *crtI* and confirmed that CarA does not activate *crtI* or *carQRS* expression.

To determine which concentration of nicotine was most appropriate to use in the experiments DK101 (wild-type) was overlayed onto solid media plates containing different concentrations of nicotine (Figure 2-9 and Figure 2-10). At concentrations above 1mM, the nicotine proved toxic to the cells and a lawn of cells did not develope. At 1mM or below, lawns did develop but still produced carotenoids. However, as the concentration of nicotine within the plates increased, the amount of carotenoids produced by the cells was reduced as judged visually by intensity of colouration. Consequently nicotine was used at 1mM concentration in all further work. It can be seen from Figure 2-10 that colony pigmentation varied between colonies. The smaller colonies appear to be more darkly pigmented regardless of nicotine concentration. Whether this is a physiologically relevant phenomenon is unclear. Maybe within the bacterial colony carotenogenesis proceeds primarily in exponential growth and is reduced in stationary phase due to the accumulation of carotenoids. Alternatively, as the colony ages, the cells within the colony disperse by gliding motility and thus older colonies are less densely pigmented because their constituent cells are less densely distributed. A final possibility is that the older, larger colonies have 'soaked up' most of the nicotine from the plates leaving the younger, smaller colonies in a nicotine-depleted envionment.

Nicotine appears to have two inhibitory effects on carotenogenesis in *Myxococcus fulvus*. At low concentrations (0.1-10μM) cyclisation was inhibited and accumulation of acyclic carotenoid glucoside esters resulted. At higher concentrations (20-100μM), hydroxylation of C-1' was also inhibited causing the accumulation solely of lycopene (Kleinig, 1974). As both cyclisation and C-1' hydroxylation are thought to involve protonation at C-2/C-2', nicotine is thought

to act by interfering with the reactions directly, not by inhibition of the enzymes responsible for catalysing the reactions (Liaaen-Jensen, 1963). Higher concentrations of nicotine proved lethal (>0.1mM). Nicotine, even at low concentrations, was observed to cause an increase in generation time and lower cell density at the end of log-phase growth (Kleinig and Reichenbach, 1973). These phenomena were not observed in this work using 1mM nicotine. Lycopene is a pale pink/red compound (Cunningham and Grant, 1998, p571) and the orange colour of *M. xanthus* in the presence of nicotine suggests that even at 1mM concentations *M. xanthus* is still capable of producing end-product carotenoids. However, the reduction in intensity of pigmentation implies a decrease in total carotenoid content in cultures containing nicotine. A reduction in the carotenoid content has also been observed in *carR* strains on the addition of nicotine (Murillo, F. J., pers.comm.), implying that inhibition of carotenogenesis by nicotine in *M. xanthus* is comparable to nicotine inhibition in *M. fulvus*.

When nicotine was added to wild-type cells carrying pDAH217, the induction of the *carQRS* promoter was seen to be increased two-fold over the same strain grown in the absence of nicotine (See Figure 2-1 and Figure 2-11 for comparison). Addition of nicotine, although enhancing promoter activity approximately two-fold, does not alter the strain's Car^+ phenotype. The enhancement of induction is not as severe as that seen on introduction of a *crtI* mutation, which concommitantly causes a more severe disruption of carotenoid production. Therefore, whereas insertion of a transposon into *crtI* abolishes carotenoid production, inhibition of CrtI by nicotine only partially hinders production but sufficiently to see an enhanced induction of the *carQRS* promoter. A possible inference is that there may be another ORF at *crtI* essential for carotenoid production which is disrupted by transposon insertion but whose product is unaffected by nicotine.

Enhancement of induction by the presence of nicotine is also seen for the *crtI* promoter (see Figure 2-3 and Figure 2-12 for comparison).

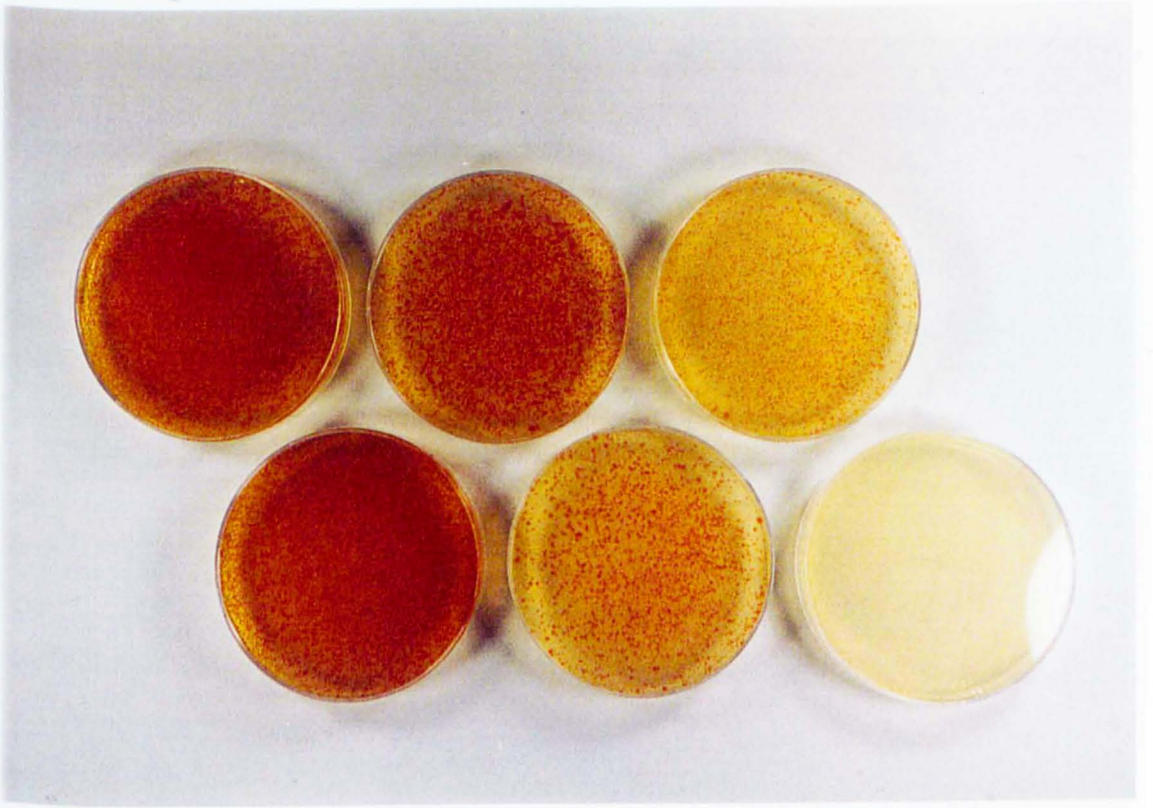


Figure 2-9 Effect of varying concentrations of nicotine on carotenogenesis of DK101 illuminated with white light.

Concentrations of nicotine within plates from left to right are 0μM, 0.1μM, 1μM, 10μM, 0.1mM and 1mM.

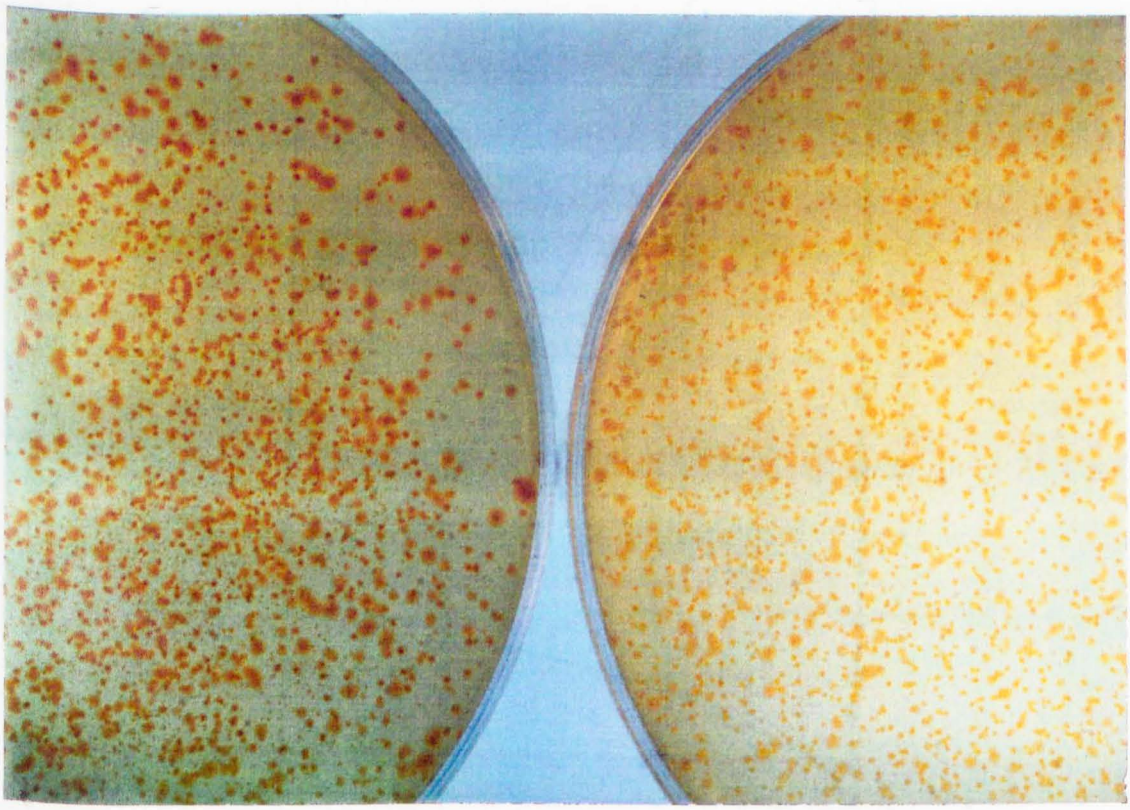


Figure 2-10 Colony pigmentation on addition of nicotine.

Colonies of M. xanthus *exposed to white light. [nicotine] in plate to left is 0.1mM, [nicotine] in plate to right is 1mM.*

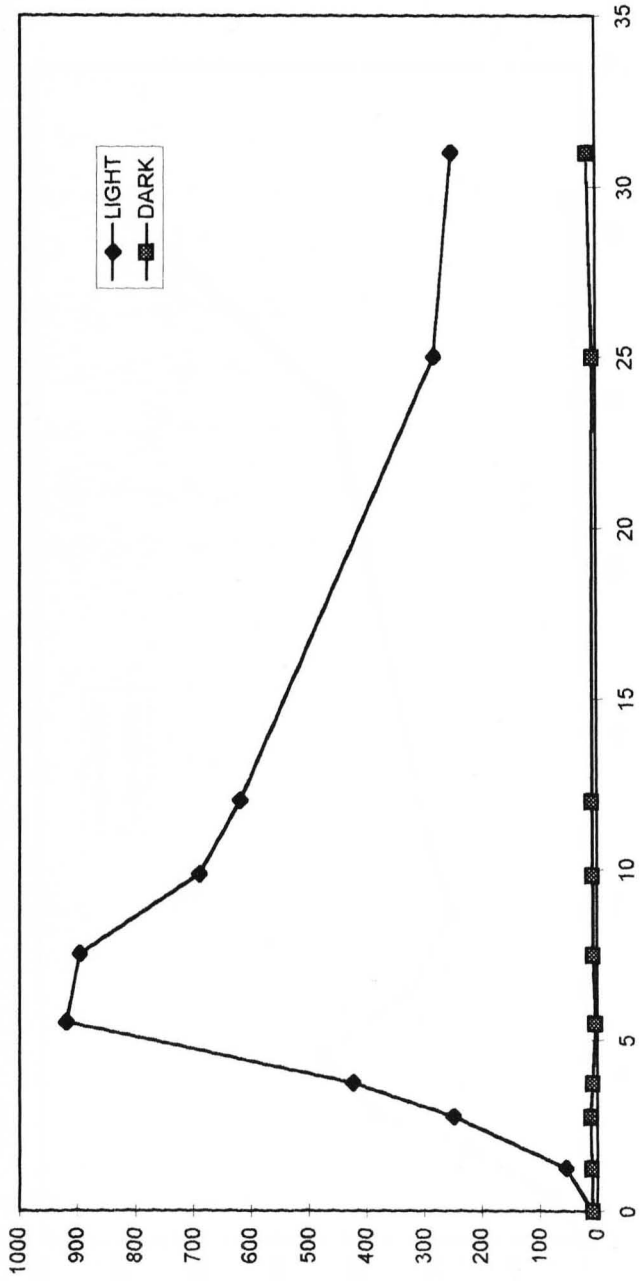


Figure 2-11 Activity of the carQRS *promoter in the presence of 1mM nicotine in the wild-type (DK101::pDAH217).*

Y-axis is specific activity of β-galactosidase in units/min/mg protein. X-axis is time in hours

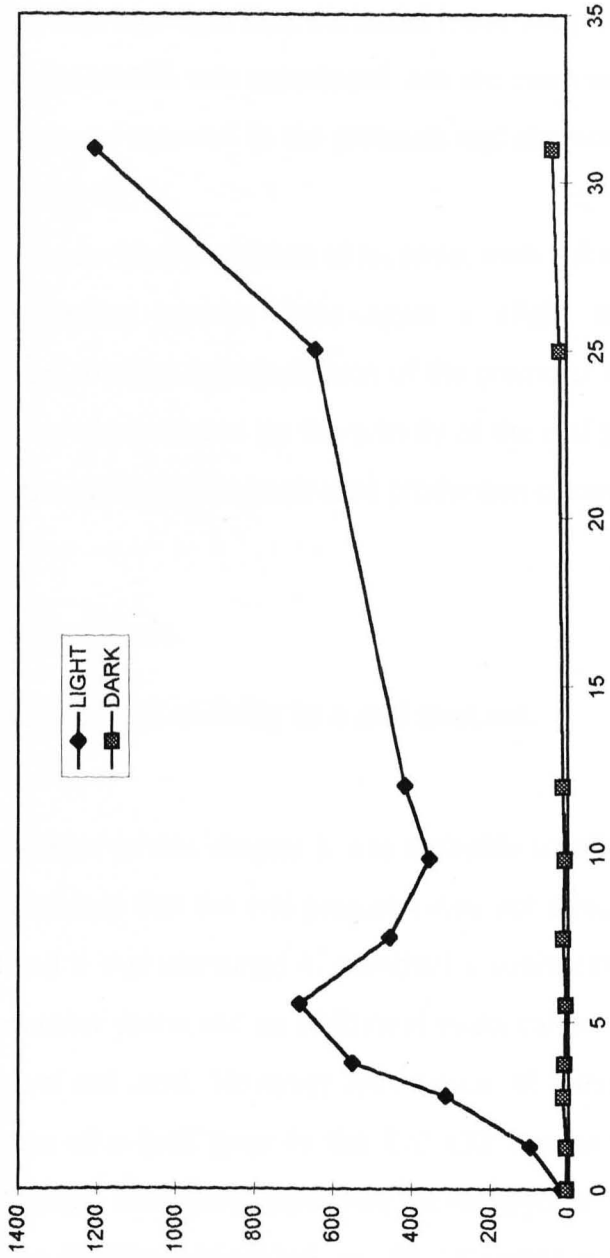


Figure 2-12 Activity of the crtI *promoter in a wild-type strain exposed to 1mM nicotine (DK101::pMAR206).*

Y-axis is specific activity of β-galactosidase in units/min/mg protein. X-axis is time in hours

Nicotine enhances the initial induction but doesn't affect the secondary induction seen on carbon limitation. Thus it would seem that the second induction is less dependent on illumination with blue light than the initial induction phase.

The *crtI* promoter probe pMAR206 was introduced into the *carA* strain DK717 and the activity of the promoter assessed in the presence and absence of nicotine (see Figure 2-13 and Figure 2-4).

Promoter activity is enhanced with the addition of nicotine, with induction up to a maximum of 70 units/min/mg protein. Once again a slight inhibition of carotenogenesis is sufficient to allow light-induction of the promoter and it can be concluded that CarA has no direct impact on the activity of the *crtI* promoter but rather exerts its effect indirectly through constitutive production of carotenoids.

2.3. Outstanding Problems.

2.3.1. Assessment of *crtI* activity in a *crtI* mutant.

To complete the work detailed in this chapter it was desirable to provide genetic as well as biochemical evidence that the *crtI* promoter was not directly activated by CarA. Towards this end it was attempted to construct a strain carrying both a *crtI* lesion and a *crtI* promoter probe and an additional strain carrying pMAR206 with mutations in both *crtI* and *carA*. However construction of these strains was hampered by the presence of a *lacZ* gene in the Tn*5*-132 insertion within the selectable *crtI* allele. Repeated attempts to transduce the tetracycline resistant *crtI* allele from MR461 into DK101::pMAR206 or DK717::pMAR206 by Mx8 mediated generalised transduction yielded no transductants. An alternative strategy was to introduce pMAR206 into *crtI* mutants UWM501 or MR461 by P1-mediated specialised transduction but this method also failed to produce transductants regardless of multiplicity of infection. Presumably the presence of the large *lacZ* gene in the chromosome and on the piece of DNA being transduced into the cell promotes homologous recombination across *lacZ* rather than across *crtI*. Most of the possible outcomes of recombination across *lacZ* either break the

chromosome and are fatal, or do not result in the chromosome acquiring the appropriate selectable marker and recombinants are therefore removed by drug selection.

A different approach was to generate a kanamycin resistant plasmid containing regions of the *crtI* gene flanking a promoterless copy of *lacZ* and a tetracycline resistance determinant. This plasmid could be electroporated into *M. xanthus* and selecting for tetracycline resistant, kanamycin sensitive electroporants would select for cells which had integrated the *lacZ* fusion by double cross-over across the flanking regions of homology to *crtI*. This would create a strain which had its endogenous copy of *crtI* replaced by an allele containing a *lacZ* transcriptional fusion and tetracycline resistance cassette. Introduction of the plasmid into wild-type DK101 and *carA* mutant DK717 would allow assessment genetically of the effect of *crtI* and/or *carA* mutations on the *crtI* promoter.

An alternative approach was to use primers containing engineered restriction sites and regions of homolgy to internal regions of *crtI,* to PCR a internal fragment of *crtI* with flanking *Bam*HI and *Eco*RI sites. Upon restriction and ligation into *Bam*HI/*Eco*RI-cut pDAH274, a plasmid would be created which carried around 1.2Kb of *crtI,* lacking the conserved regions found at the 3' and 5' ends of phytoene dehydrogenases, upstream of a promoterless *lacZ* gene. Introduction of this plasmid into DK101 by P1-mediated transduction would generate a merodiploid with two incomplete copies of *crtI*, with the copy of *crtI* lacking its C-terminal coding region transcriptionally fused to *lacZ*.

Lack of time prevented completion of these strategies, however, future studies will continue the attempts to create a *crtI* mutant containing a *crtI* promoter probe.

If the current notion of how CrtI is involved in carotenogenesis is correct, the *crtI* promoter would be expected to exhibit enhanced induction in a *crtI* mutant with a lack of down-regulation by feedback mechanisms. These phenotypes would be expected to be irrespective of a *carA* lesion.

2.3.2. Effect of nicotine on *carQRS* expression in a *carA* background.

When β-galactosidase expression from a *carA* mutant carrying pDAH217 was assayed in the presence of nicotine, no effect could be seen due to the addition of nicotine (Figure 2-14).

The results obtained from DK717::pMAR206 showed that 1mM nicotine was sufficient to see an elevated rate of induction on illumination and it was expected that the same would be true when looking at the *carQRS* promoter. This experiment was repeated several times using different clones of DK717::pDAH217 and in no case did an enhanced induction arise due to incubation with nicotine. The reason for this is not clear. If *crtI* expression is enhanced by nicotine, then it would be expected to be due to a reduction in quenching of high energy signalling molecules such as singlet oxygen, resulting in greater levels of active CarQ. This is seen to enhance *crtI* induction but seems to have no effect on *carQRS* induction. An explanation for this might involve the difference between the two promoters, with different ancilliary factors being required by the two promoters (This phenomenon may also be manifested in the light induction of the *crtI* promoter in a *carR* mutant {see section 3.2}).

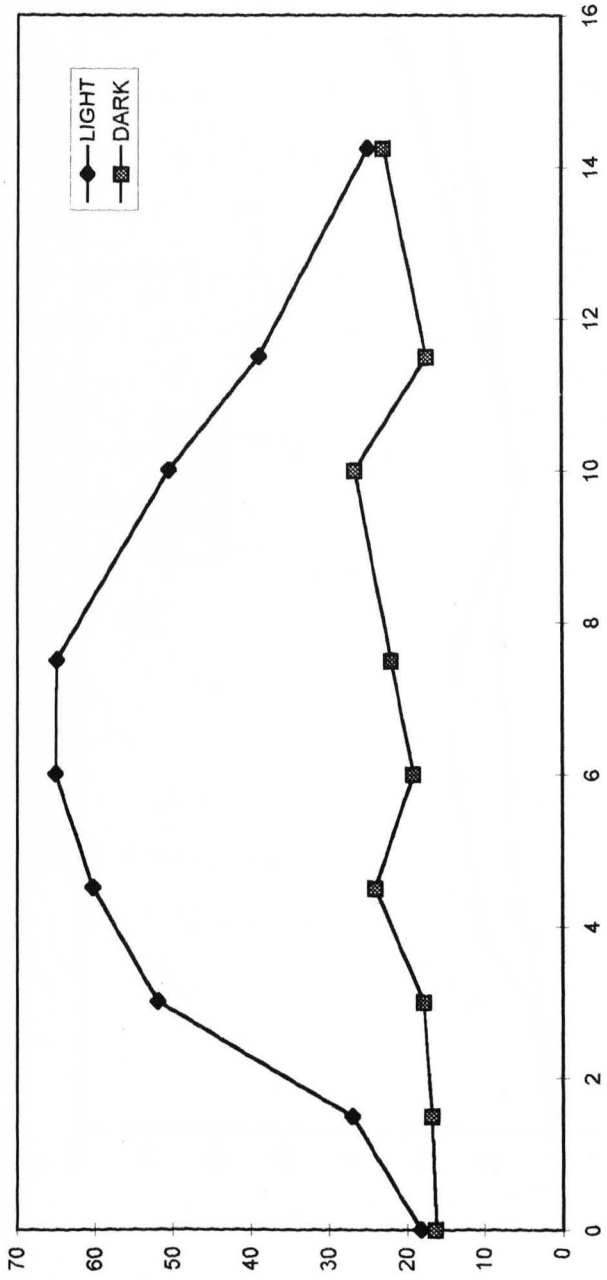


Figure 2-13 Activity of the crtI *promoter in a* carA *background in the presence of 1mM nicotine (DK717::pMAR206).*

Y-axis is specific activity of β-galactosidase in units/min/mg protein. X-axis is time in hours

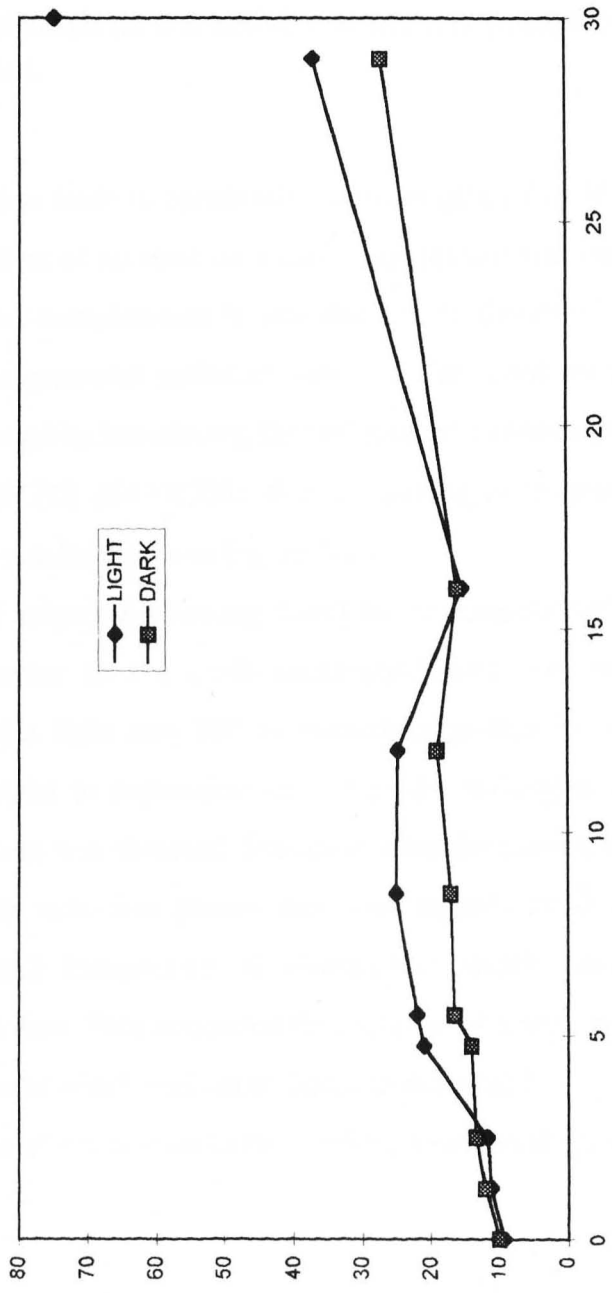


Figure 2-14 Activity of the carQRS *promoter in a* carA *mutant with 1mM nicotine (DK717::pDAH217).*

Y-axis is specific activity of β-galactosidase in units/min/mg protein. X-axis is time in hours.

2.3.3. Effect of nicotine on the activity of the *crtI* promoter in a *carR* mutant.

The two loci whose mutation leads to constitutive carotenogenesis in *M. xanthus* are *carA* and *carR*. The effect of nicotine on a Car^c *carA* mutant was determined as described above and for completeness it was decided to determine whether nicotine had any effect on promoter activities within a Car^c *carR* mutant. The effect of nicotine was assessed by introducing the *crtI* promoter probe into a strain carrying a *carR* lesion (DK718::pMAR206). β-galactosidase expression by this strain was assayed and the results shown in Figure 2-15.

Consistant with the *carR* mutation allowing CarQ to be constitutively active, activity of the *crtI* promoter in the *carR* background was very high (400 units/min/mg protein in the light and 100 units/min/mg protein in the dark). Additionally, when compared to expression of *crtI* in the wild-type, a similar pattern of bi-phasic induction was observed. However, what the current model did not explain was why both induction phases were still light-induced when the availability of active CarQ irrespective of illumination should mean equal activities in both light and dark. This suggested the presence of a separate level of control over the *crtI* promoter which was not included in the model.

Nicotine had no effect on *crtI* expression in this *carR* background (Figure 2-16).

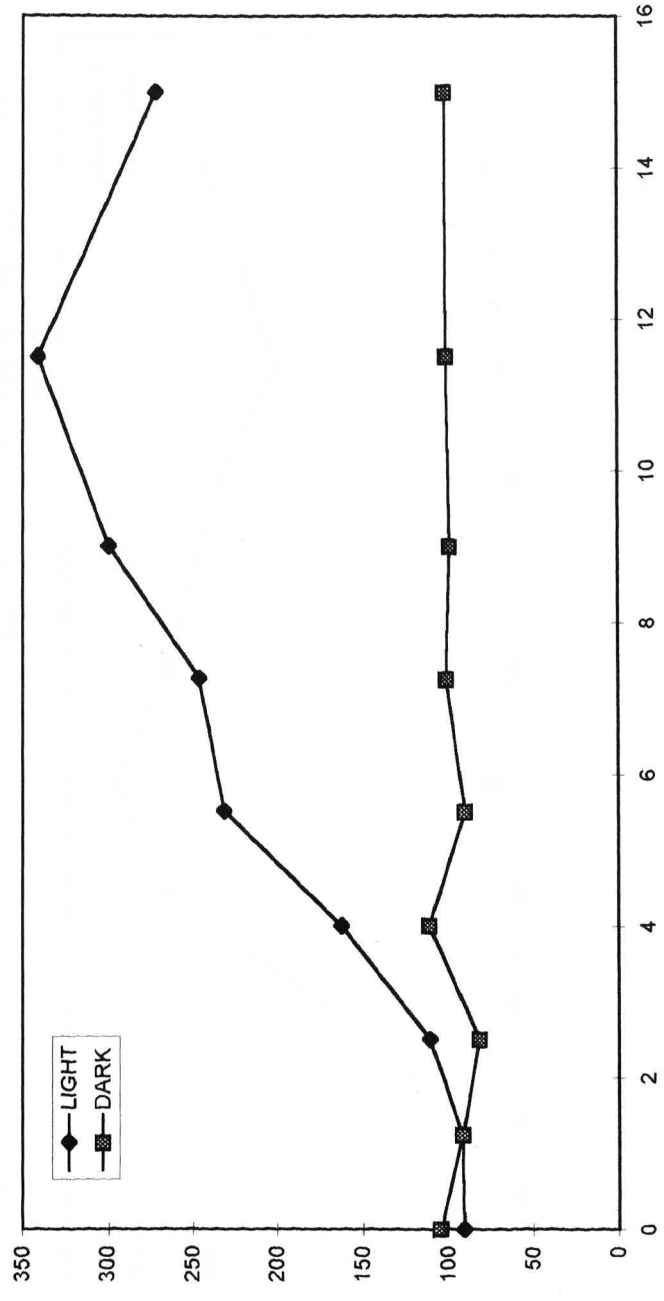


Figure 2-15 Activity of the crtI *promoter in a* carR *mutant (DK718::pMAR206)*

Y-axis is specific activity of β-galactosidase in units/min/mg protein. X-axis is time in hours.

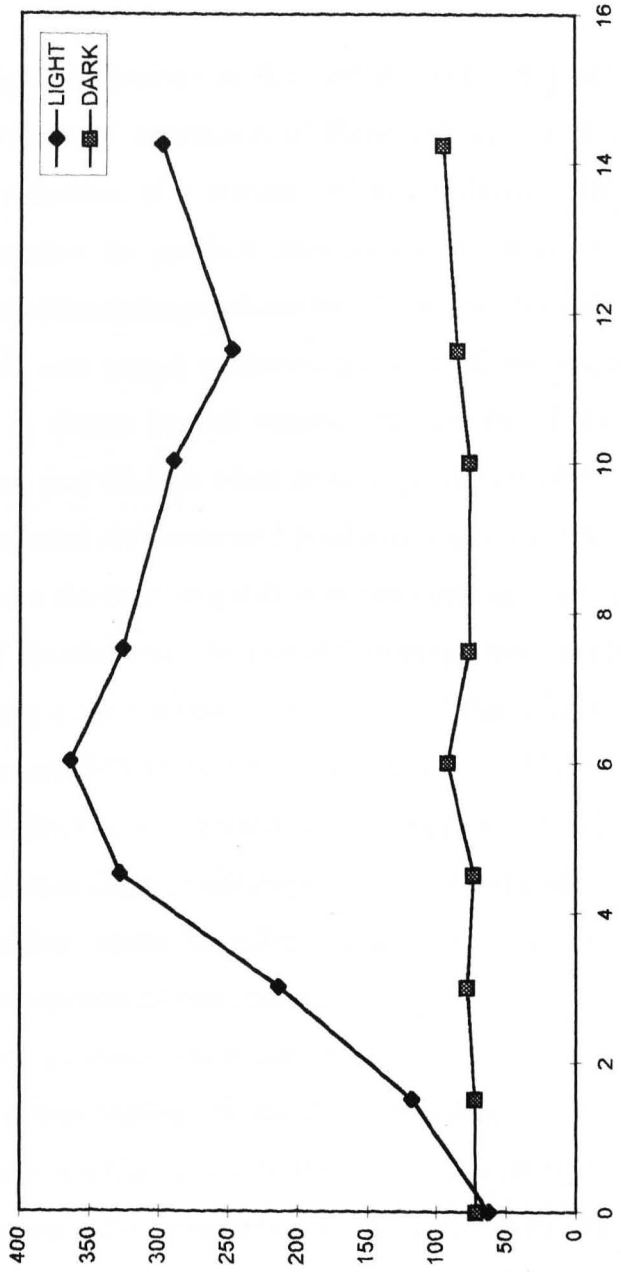


Figure 2-16 Activity of the crtI *promoter in a* carR *mutant (DK718::pMAR206) in the presence of 1mM nicotine.*

Y-axis is specific activity of β-galactosidase in units/min/mg protein. X-axis is time in hours.

2.4. Conclusions.

The use of *lacZ* transcriptional fusions to the *carQRS* and *crtI* promoters has enabled *in vivo* determination of expression of these two loci in a variety of mutant backgrounds. Introduction of a mutant *crtI* allele allowed strains to be constructed which are unable to produce carotenoids. In wild-type strains, carotenoid production upon illumination reduces the strength of the light stimulus being transduced to CarR and causes a down-regulation of the regulon. This negative feedback loop is absent in *crtI* strains and activity of the *carQRS* promoter is seen to increase past the time when down-regulation is observed in the wild-type. Thus *crtI* is essential for carotenoid production and it is the formation of carotenoids which causes the down-regulation of promoter activity seen in the wild-type after 6 hours of illumination. The *carQRS* promoter has a much reduced actvity in strains containing a *carA* lesion. Transduction of the *crtI1* allele into a *carA* mutant carrying the *carQRS* promoter probe created UWM501::DAH217, which demonstrated no difference in promoter activity from a *crtI* single mutant. It was concluded therefore that CarA is not required for normal expression of the *carQRS* promoter, but rather exerts its effect on *carQRS* activity indirectly, through the constitutive production of carotenoids.

Difficulties in the creation of a strain which carried both a *crtI* mutation and a *crtI* promoter probe caused a biochemical approach to be taken to determine the activity of the *crtI* promoter in a Car$^-$ strain. Nicotine has been reported to cause a reduction in the production of carotenoids by blocking steps early in the biosynthetic pathway. Illuminating cells in the presence of 1mM nicotine caused a slight decrease in colouration, while using greater concentrations of nicotine had lethal effects. Addition of nicotine to wild-type cells carrying pDAH217 or pMAR206 showed that nicotine delayed the onset of feedback regulation consistent with a reduction in the amount of carotenoids produced. Mutants with *carA* lesions show much reduced *crtI* promoter activities in comparison to the wild-type. Addition of nicotine to *carA* cells carrying pMAR206 caused an increase in *crtI* promoter activity consistent with CarA's effect on P^{crtI} being

indirect, through constitutive carotenoid production. Thus it would seem that CarA does not affect the *crtI* promoter directly, but can influence activity of the carotenogenic regulon as a whole through its repression of *crtEBDC* expression and subsequent carotenoid production.

It is clear that feedback by the production of carotenoids has important implication for the activities of promoters within the carotenogenic regulon and the effects of any mutations on carotenoid production must be taken into account when assessing any direct effects on promoter activity. Thus *carA* mutants affect promoter activity indirectly because they cause the negative feedback loop of carotenoid production to be established prior to illumination. Strains with *crtI* lesions however lack the feedback loop and therefore promoter activities are high because negative feedback can never be established.

However, as well as clarifying the role of CarA within the carotenogenic system, the work described in this chapter has uncovered some interesting and as yet unexplained phenomena, particularly concerning *crtI* gene expression. These phenomena will be discussed in the next chapter.

3. The roles of *crtI*.

Experimental work has shown that there are significant gaps in our understanding of the *crtI* locus and that its expression is not as easy to explain as that of the *carQRS* operon.

To allow for the assessment of the action of nicotine on a Car^c strain other than a *carA* mutant, *crtI* promoter activity assays were performed for a *CarR* mutant in which the sigma factor CarQ is constitutively free and able to activate its cognate promoters. In this strain the *crtI* promoter shows two phases of promoter activity, unexpectedly, both were subject to light induction, with the second phase consistent with entry into stationary phase. Additionally, the *crtI* promoter's induction on illumination is enhanced by addition of nicotine as predicted, but this is not true for the *carQRS* promoter. Functional differences between the two promoters were also hinted at by *in vitro* transcriptional run-off assays which proved it was possible to obtain a transcript from the *carQRS* promoter but not the *crtI* promoter using purified CarQ (Browning, 1997).

Another interesting phenomenon is that disruptive *crtI* mutations cause a Car^- phenotype instead of a Car^+ phenotype. This is unexpected as it is known that expression of *carBA* on its own (for instance in a *carA* mutant in the dark) is sufficient for the production of coloured carotenoids. Therefore it was expected that a *crtI* mutant would still have a Car^+ phenotype, not the Car^- phenotype that is observed in all strains carrying a *crtI* mutation. Additionally, the presence of nicotine which inhibits CrtI does not grant a Car^- phenotype.

All these features of CrtI and the *crtI* promoter suggest that its behaviour is not as simple as the model implies. It has become necessary to address these unexplained aspects of *crtI* activity.

3.1. Introduction.

The known gene at *crtI* encodes a protein with significant similarity to phytoene dehydrogenases from a diverse set of organisms including *Neurospora crassa*, *Rhodobacter capsulatus* and two species of *Erwinia* (Fontes *et al.*, 1993). CrtI contains a noncovalently bound FAD co-factor and is homologous to both hydroxyneurosporine desaturase (CrtD) and protoporphyrinogen oxidase (which catalyses the oxidation of protoporphyrinogen to protoporphyrin IX) of *M. xanthus* (Dailey and Dailey, 1998). Disrupting *crtI* causes the strain to accumulate phytoene, the substrate for phytoene dehydrogenase and trace amounts of phytofluene (Martinez-Laborda *et al.*, 1990). Expression of *crtI* requires CarQ directly, and CarD (indirectly or directly) and is maximal in conditions of carbon-limitation and illumination (Gorham *et al.*, 1996, Nicolas *et al.*, 1994 and Fontes *et al.*, 1993). These were the facts known about *crtI* three years ago and the model explains all these characteristics except for the activation by carbon-limitation.

3.2. Light Induction of *crtI* in a *carR* mutant.

The *crtI* promoter is light-induced in a *carR* background (DK718::pMAR206), see Figure 2-15 and Fontes *et al.* (1993). The simplest explanation perhaps being that the *carR5* allele carries an incompletely disrupting mutation and that the carR protein produced is still able to bind carQ at a reduced affinity, with enough free CarQ unbound to CarR to allow constitutive expression of the regulon. To investigate the possibility that the light-induction of *crtI* was due to a peculiarity of the particular clone assayed or the *carR* allele used, DK718::pMAR206 was recreated and the *crtI* promoter probe was also introduced into a strain carrying a different mutant *carR* allele (DK406 carries the *carR4* allele while DK718 carries the *carR5* allele). Both strains were assayed for β-galactosidase activity. In each case, light-induction was apparent in both phases of induction. Thus the

phenomenon is not unique to a particular *carR* allele or clonal isolate (Figure 3-1 and Figure 3-2).

With a biphasic induction it was difficult to conclude that each phase was being truly light-induced. It was possible that the initial activation was due to a carry over of stationary phase cells from the original inoculum and that this sub-population of cells was exhibiting a phase II activation, not a true initial induction. Conversely, phase II may be due purely to loss of catabolite repression and is biologically irrelevant in the context of induction of carotenoid production. In order to avoid confusion due to these possibility the β-galactosidase assays were repeated but after 3 hours the cultures were centrifuged and the pelleted cells were resuspended in a buffer lacking carbon source. This sent each culture into stationary phase and phase II induction of *crtI* expression. In this way it proved possible to assess a single phase of induction uncomplicated by the presence of a second phase. When a media replacement β-galactosidase assay was performed on a wild-type strain containing pMAR206 it was clear that carbon limitation stimulated expression of *crtI* and that light also induced expression irrespective of illumination (Figure 3-3).

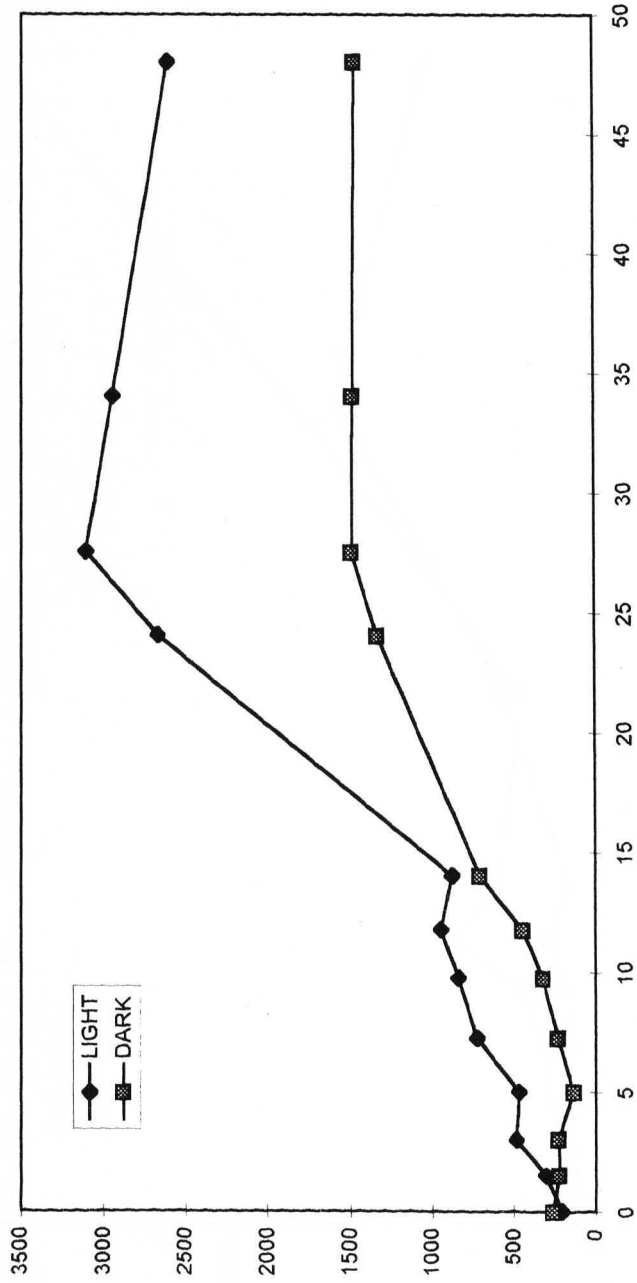


Figure 3-1 Activity of the crtI *promoter in the* carR5 *mutant (DK718::pMAR206).*

Y-axis is specific activity of β-galactosidase in units/min/mg protein. X-axis is time in hours.

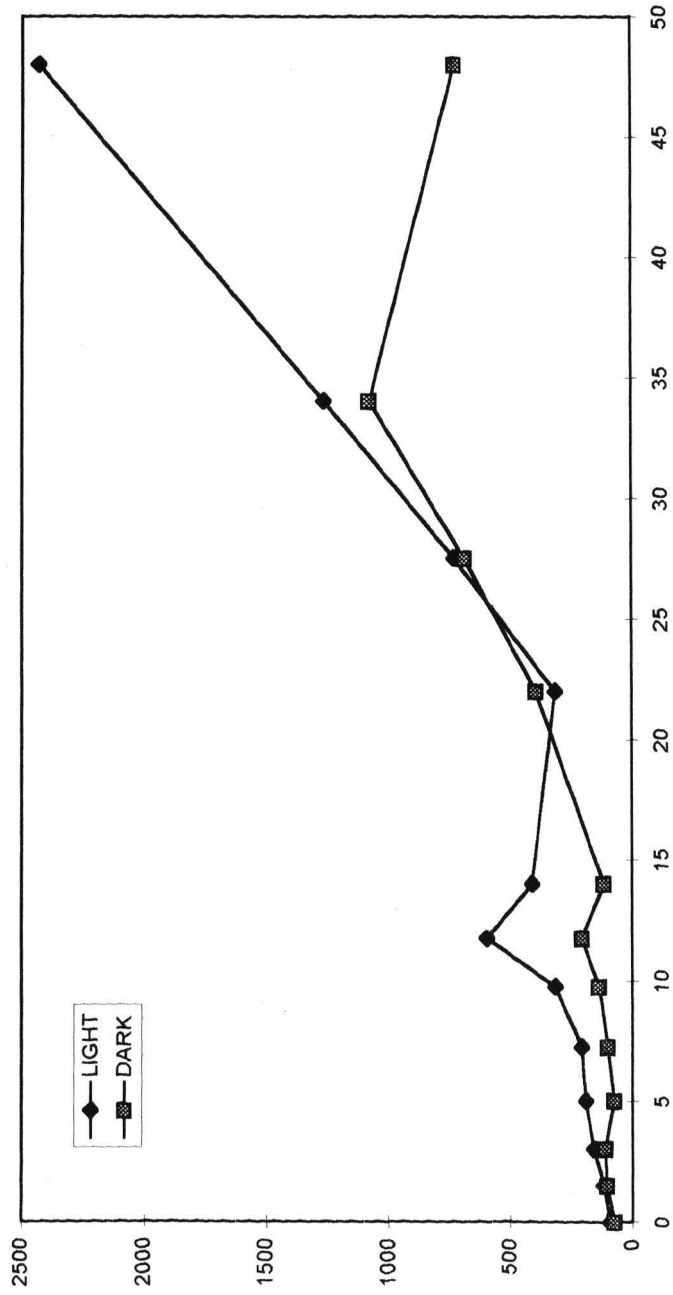


Figure 3-2 Activity of the crtI *promoter in the* carR4 *mutant (DK406::pMAR206).*

Y-axis is specific activity of β-galactosidase in units/min/mg protein. X-axis is time in hours.

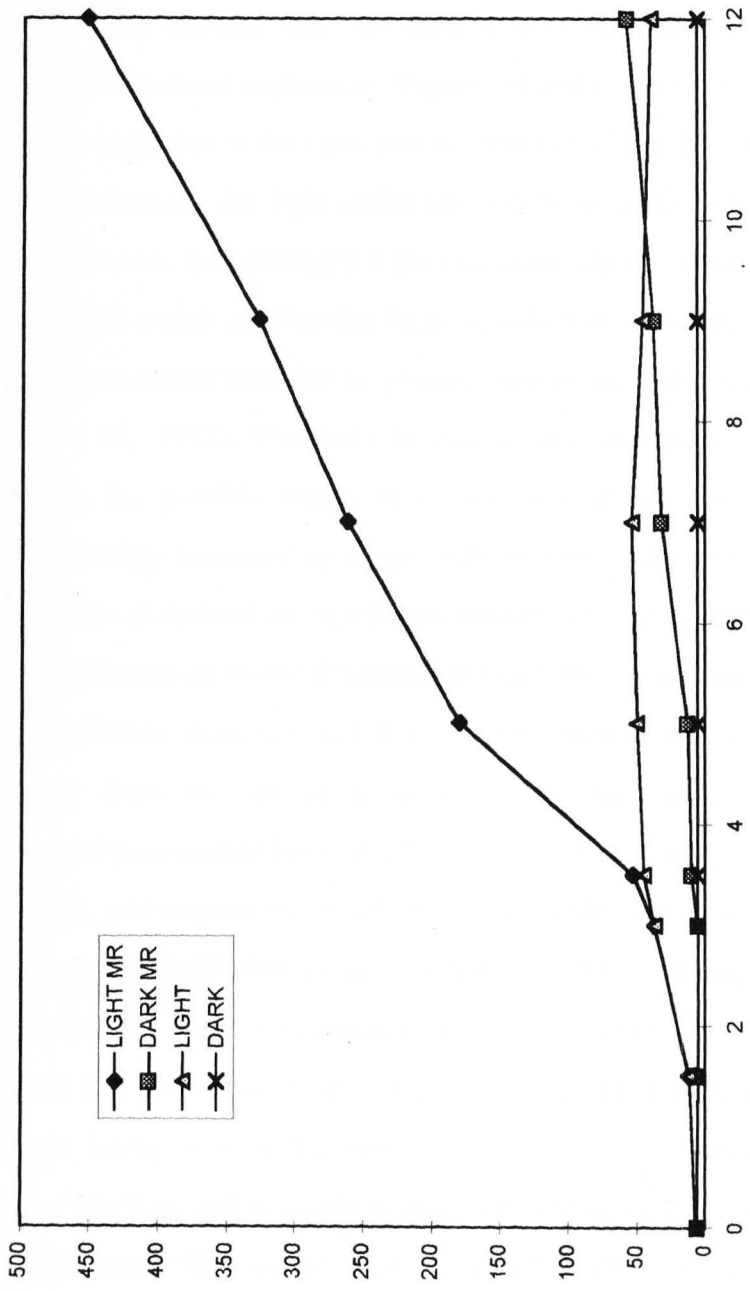


Figure 3-3 Activity of the crtI *promoter in a wild-type strain (DK101::pMAR206) with carbon limitation after 3 hours.*

Y-axis is specific activity of β-galactosidase in units/min/mg protein. X-axis is time in hours. The suffix MR denotes that those cultures were subjected to media replacement after 3 hours.

Assessing β-galactosidase expression of the two *carR* backgrounds using media replacement after 3 hours showed that for each strain, starvation enhanced induction and light also stimulated expression (Figure 3-4 and Figure 3-5).

The occurrence of light-induction in the constitutive presence of CarQ implies that for *crtI* there is a mechanism for light-induction which is additional to that mediated by CarQ. However, the additional light induction causes an increase in activity of only ~two-fold which implies the light induction is more likely to be due a physiological effect rather than due to genetic factors such as a regulatory gene product (Hood *et al.*, 1992). This may be due to any number of possible explanations, including the possible action of an activator of *crtI* transcription which has its activity directly increased by a light-induced conformational change. Alternatively, the increase in thermal energy within the cell upon illumination may allow a faster rate of diffusion of molecules inside the cell. In this scenario, CarQ may have its rate of diffusion increased and therefore its concentration at regions within the cell away from its site of production may be increased upon illumination. This would presumably have no effect at the site of CarQ production (i.e. the *carQRS* operon, but at other loci (such as *crtI)* the effective concentration of CarQ will be increased manifesting as an increased activity in the light. This wildly speculative possibility could also explain why nicotine enhances induction of *crtI* but not *carQRS* in a *carA* mutant (see section 2.3.2). If the induction of *crtI* in the light in a *carA* background in the presence of nicotine is mainly due to increased diffusion of CarQ to *crtI*, it might explain why the same increase is not seen at the *carQRS* promoter. This supposition assumes that the slight inhibition of carotenoid production seen on the addition of nicotine is significant enough to allow greater amounts of thermal energy from light into the cell. However, carotenoids which absorb light are excited to a singlet state which emits the excitation energy as light. Thus absence of carotenoids is actually more likely to decrease rather than increase thermal energy within the cell.

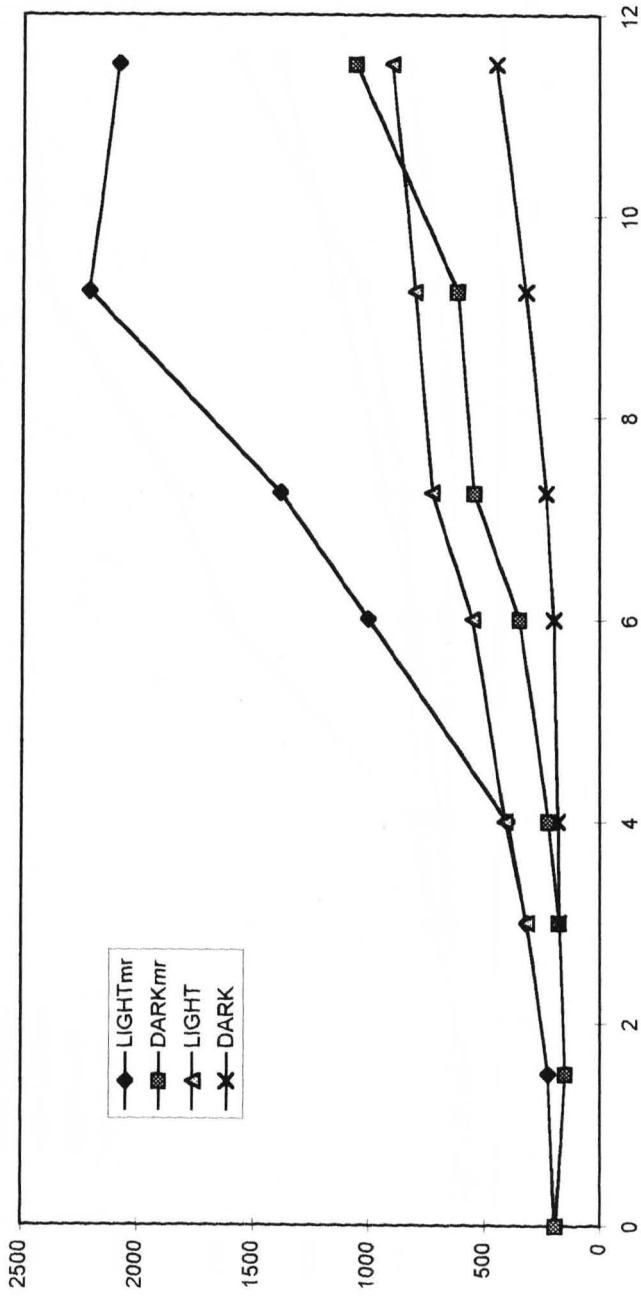


Figure 3-4 Activity of the crtI *promoter in a* carR *strain (DK718::pMAR206) with carbon limitation after 3 hours.*

Y-axis is specific activity of β-galactosidase in units/min/mg protein. X-axis is time in hours. The suffix MR denotes that those cultures were subjected to media replacement after 3 hours.

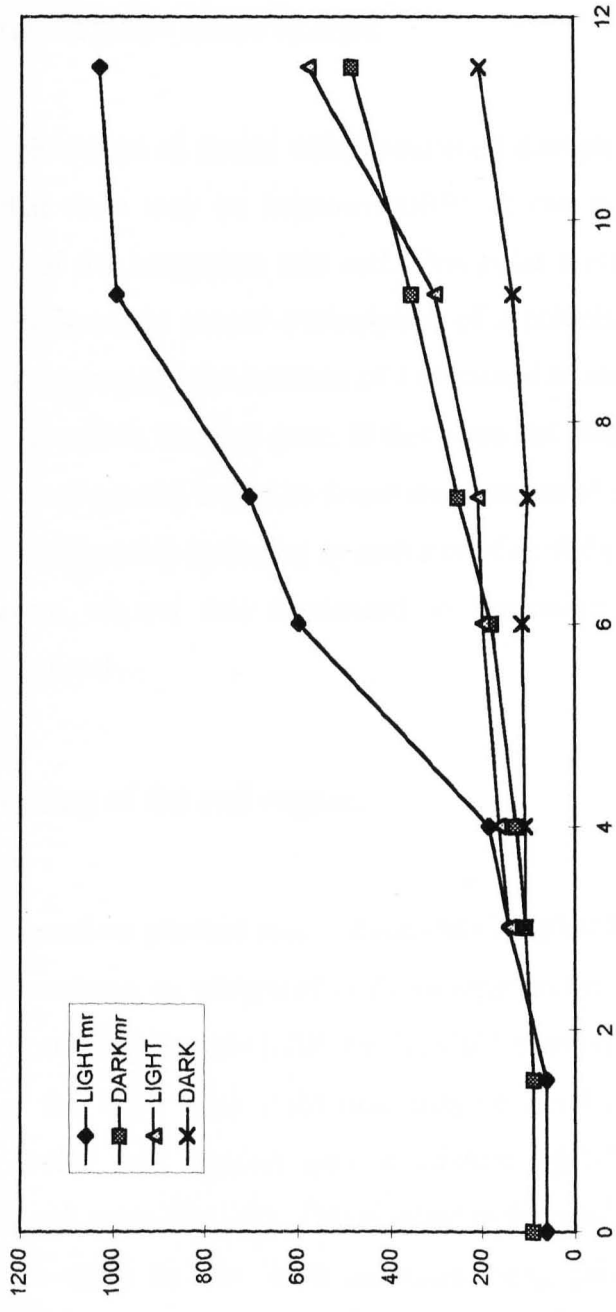


Figure 3-5 Activity of the crtI *promoter in a* carR *strain (DK406::pMAR206) with carbon limitation after 3 hours.*

Y-axis is specific activity of β-galactosidase in units/min/mg protein. X-axis is time in hours. The suffix MR denotes that those cultures were subjected to media replacement after 3 hours.

3.3. Are additional genes found at *crtI*?

The unexpected Car$^-$ phenotype of strains which contained disruptive transposon insertions suggests that there may be unknown ORFs at *crtI* which are also disrupted by insertion of the transposon into *crtI*. This polar effect would be a consequence if the *crtI* promoter caused transcription of a polycistronic mRNA containing more than one gene and the insertion of Tn*5* caused transcriptional stop signals to be introduced within the first gene. If this were the case, it would be expected that an additional gene(s) would be found downstream of *crtI* reading in the same direction of *crtI*, possibly including an activator of *carBA* expression.
The region downstream of *crtI* was sequenced to determine whether any additional genes were present.

3.3.1. Sequencing of the *crtI* region.

The *crtI* region was cloned by plasmid rescue from DK101::pMAR206 which is the wild-type strain containing an integrated *crtI* promoter probe. Chromosomal DNA was extracted from DK101::pMAR206 by the CsCl method of McGowan (1992). Incubation of the DNA with restriction enzyme *Eco*RV followed by inactivation of the *Eco*RV and ligation gave a mixture which was used to transform *Escherichia coli* strain MC1061. This resulted in the creation of plasmid pDEW100 which was ~15kb in size. PCR reactions using primers 3 and 4 (7.4.6.2) confirmed it carried the *crtI* promoter but restriction mapping suggested it had undergone substantial rearrangements of the *crtI* region. Therefore plasmid pMAR202 which contains ~20 kb of the *crtI* region was chosen as the template for sequencing (Fontes *et al.* 1993). This plasmid was transformed into DH5α, a strain lacking a host methylation modification system. Sequencing was performed using the method of Sanger *et al.* (1977) using dye-terminators. Sequencing reactions involved 30 cycles with a 15 second melting step at 96^0C followed by a 15 second annealling step at 50^0C and a final extension step of 3 minutes at 60^0C.

Reaction products were visualised by gel electrophoresis (Alta Bioscience, Birmingham UK). For an example of sequence data, see section 8.2.

3.3.2. DNA Sequence of the *crtI* region.

Sequencing was performed on both strands of the *crtI* region. The DNA sequence marked below in Figure 3-10 follows on from the sequence of the phytoene dehydrogenase gene (*crtI*) as determined by Fontes *et al.* (1993) such that position 1905 from Fontes *et al.* (1993) is the first base described with the initial GTGA containing a GTG triplet overlapping the TGA stop codon of the *crtI* gene. Analysis of the novel sequence was initially performed using FRAME (Bibb *et al.*, 1984), BLAST 2.0 and PSI-BLAST (Altschul *et al.*, 1997). FRAME analysis (Figure 3-9) suggested that there was an open reading frame of 852bp within the novel sequence, that conformed to expected bias in codon base usage (Bibb *et al.*, 1984) for a high-GC content organism. Thus for this putative ORF, the third codon base was a G or C in 90.2% of codons. GC percentage in codon positions one and two were 79.3 and 52.3 respectively with an overall GC content of 73.9% which agrees well with values described for myxobacterial DNA (Mandel and Leadbetter, 1965). This putative open reading frame is designated (gene of *u*nknown *f*unction B) *gufB*. The presence of an open reading frame divergent with the *crtI* gene was also revealed by FRAME. This ORF has been designated *gufC* (gene of *u*nknown *f*unction C) and is shown in Figure 3-11. The first 164bp of the sequence of *gufC* are known, comprising 55 codon triplets and GC content in each of the three codon positions corresponds to that expected for a gene of *M. xanthus*. The first codon position is 78.1% GC, the second is 45.5% GC and the third codon position is 92.7% GC with an overall GC percentage of 72.6% over the entire 164bp. Attempts to identify possible ribosome binding sites were made by comparisons to sequence complementary to the 3' sequence of the myxobacterial 16S rRNA (3'-OH-UCUUUCCUCCACUA...-5', Oyaizu and Woese, 1985). No nucleotide stretches with significant similarity to this sequence were found upstream of *gufC* or *gufB*.

```
GTG AGC GCC GCG CCA TCC AGT GGC GCG CTC ACC CGG
M   S   A   A   P   S   S   G   A   L   T   R

CAG GCG GGG AAG GGG CGC CCG CAT CCG CCG GGC GGG
Q   A   G   K   G   R   P   H   P   P   G   G

GCG GGG GCG AAG CTG GCA TAG TTG CGA GGC ATG
A   G   A   K   L   A   ***             M
```

*Figure 3-6 Putative peptide gene (*olpA*) upstream of* gufB.

Initial GTGA is a GTG initiation codon overlying the TGA stop codon of crtI. *The stop codon of the peptide is shown with ***. The final ATG is the initiation codon of* gufB. *It is assumed that the first codon encodes formyl-methionine.*

It is possible that *crtI* is translationally coupled to a peptide which serves to deposit ribosomes upstream of the 5'-end of *gufB* (Rex *et al.*, 1994). This could imply that *crtI* and *gufB* are transcribed as a single mRNA species from the *crtI* promoter and would thus form an operon. In which case, it might be reasonable to assume that *crtI* and *gufB* are involved in the same physiological function.

The overall GC percentage of the *olpA* peptide coding region is 81.1% with GC content in the three coding positions of 80%, 76.7% and 86.7% respectively. This conforms to the expected relative codon usage bias of *M. xanthus*, given that the overall region has such a high GC content. There are four residues in the coding sequence for *olpA* that are coded for by rare codons (Wright and Bibb, 1992) i.e. Pro(5), Ser(7), His(20) and Ala(30). Conventional thinking suggested that the presence of rare codons at the start of a gene may have regulatory consequences with synonomous codon usage correlating diectly with levels of expression

(Shields and Sharp, 1987). It is probable that the tRNAs for infrequently used codons have relatively low concentrations in the cell. A possible regulatory consequence would be the lower rate of translation of an mRNA containing an abundance of infrequently used codons (Konigsberg and Godson, 1983). A typical model might be that on translation of *olpA* the ribosomes 'stall', allowing increased levels of translation of *gufB*.

The region surrounding *olpA* was assessed for the possibility of secondary structure formation within a transcribed RNA product using MFOLD (Zucker, 1989). MFOLD predicted 20 different models of possible region of secondary structure within the region, however certain hairpin loops were present in the majority of predicted folds. Appendix 8.3 shows representative examples of MFOLD results displayed as squiggle plots. Predicted hairpin loops of interest are shown in Figure 3-7. Of the twenty squiggle plots, no two plots suggested an identical stem-loop around the *crtI/olpA* junction, while fifteen of the squiggle plots consistently predicted a large stem-loop within the *olpA/gufB* junction area. This hairpin lies at such a position that the initation codon of *gufA* lies at the bottom of the stem. If this structure is capable of forming *in vivo* then a possible mechanism for translational coupling between *olpA* and *gufB* is clear. Translation initiation may only be possible if ribosomes pass along the mRNA of *olpA* disrupting the stem-loop structure. This will cause the section of mRNA containing the *gufB* translational start site to revert to single stranded nucleic acid and allow initiation of *guB* translation to occur. Site directed mutagenesis of the hairpin would allow this phenomenon to be confirmed or denied.

The lack of obvious secondary structure of mRNA at the *crtI/olpA* junction does not preclude the possibility of translational coupling between these two genes as the overlapping start and stop codons of the genes may imply a model of translational coupling where after translation of *crtI*, the ribosome 'shifts' back a nucleotide to re-initiate translation of *olpA*.

Met

CGTATGTGCCGCCGCTGGAGCCGGAGTCGCCCGTTCAAGCTCGCGCGGGGTGAGC

GCCGCGCCATCCAGTGGCGCGCTCACCCGGCAGGCGGGGAAGGGGCGCCCGCATC

Met

CGCCGGGCGGGGCGGGGGCGAAGCTGGCATAGTTGCGAGGCATGCCCTTCTTCAT

CCCATTCGCGGTGGGTGGCCTGGTGCTGACGGCACTGGGCCTGGGTGTGAGGAAG

Figure 3-7 Predicted hairpin loops around olpA.

Bases which form hairpin loops are underlined. The GTG and ATG codons denoted 'Met' are the initiating codons for olpA *and* gufB *respectively.*

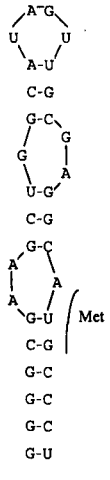


Figure 3-8 Predicted hairpin loop at the start of gufB.

The AUG initiation codon for gufB *is on the downward strand of the stem.*

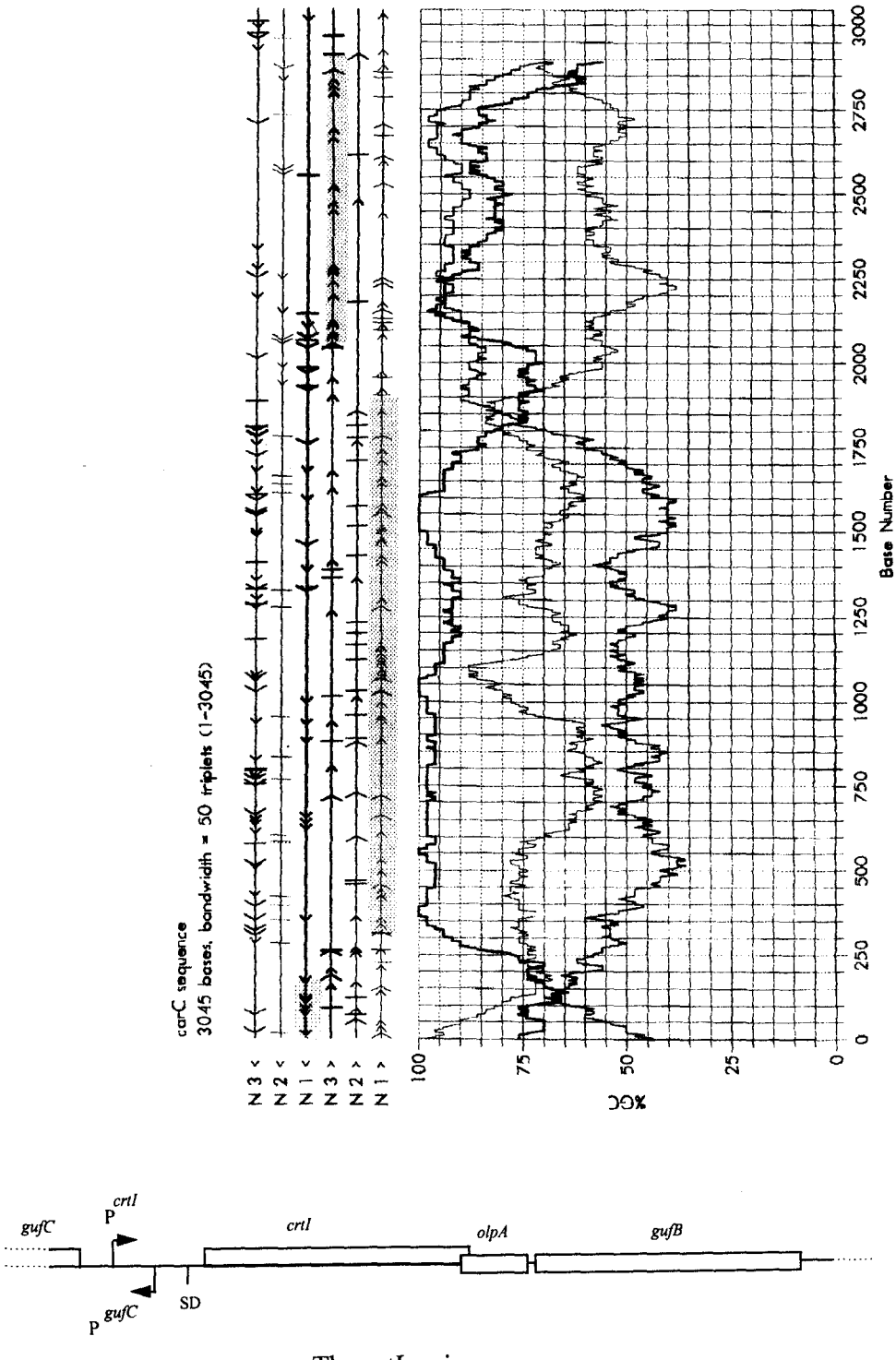
In *Bacillus subtilis,* genes for biologically active oligopeptides are found downstream of genes of related function (Grossman, 1995). The initial steps in commitment to sporulation of *Bacillus subtilis* involve a phospho-relay system terminating in phosphorylated SpoOA. One of the phosphorylated intermediates in the relay is SpoOF which is dephosphorylated by the phosphatase RapA. Immediately downstream of *rapA* is *phrA* which encodes a small peptide. PhrA is secreted and processed to a smaller peptide which is imported by the SpoOK oligopeptide permease. The peptide then inhibits the action of RapA and allows sporulation to procede (Perego and Hoch, 1996). A similar system is involved in the development of genetic competence in *B. subtilis*. ComX is a small peptide pheromone which is exported form the cell, processed to a smaller peptide and imported by SpoOK. The peptide then inhibits a cellular phosphatase which is responsible for preventing accumulation of ComA~P and the subsequent development of competence. The gene for ComX is immediately downstream of *comQ* which appears involved in the export of ComX from the cell (Soloman and Grossman, 1996).

It is possible that OlpA has a role in the cell other than to translationally couple *crtI* and *gufB* expression. It may regulate an aspect of carotenoid synthesis or it may control a separate aspect of cell physiology. It seems unlikely that it's presence between *crtI* and *gufB* is purely to couple the two genes, as this could be easily achieved by the cell without recourse to an intervening peptide coding region, however the possibility cannot be discounted.

Translation of the DNA sequence of *gufB* gave a theoretical protein product of ~28 kDa whose sequence is shown in Figure 3-10. BLAST 2.0 analysis failed to reveal the presence of homologues in the database, however analysis with the iterative PSI-BLAST algorithm revealed GufB possessed similarity to several proteins, with the best similarity to members of the PE family identified recently in *Mycobacterium tuberculosis* (Cole *et al.,* 1998){17% identity plus 10% similarity over 80.3% of GufB, with an E-value of 6e-36}. For all PE family homologues, similarity was found to be towards the C-terminal half of GufB. One of the PE family homologues has been shown to be a lipase/esterase of *M.*

tuberculosis (Cole *et al.*, 1998). Other homologues include human butyrophilin (Banghart *et al.*, 1998) {30% identity plus 11% similarity over 20.7% of GufB, with an E-value of 7e-13}, peroxidase/peroxidase-like proteins (29% identity plus 5% similarity over 20.4% of GufB with an E-value of 3e-5) and Gcr-1 of *Chlamydomonas rheinhardtii* {22% identity plus 9% similarity over 34% of GufB with an E-value of 4e-4} (Wakarchuk *et al.*, 1992). The *gcr-1* gene contains many GC-rich direct repeats but is of unknown function.

Analysis of *gufC* by BLAST 2.0 also failed to reveal any homologues in the database, which was to be expected due to the small length of the known sequence of *gufC*. However, searching with PSI-BLAST, iterating until convergence, gave three homologues with homology spanning the entire known sequence of GufC, including a gene from *Caenorhabditis elegans* (cDNA yk46f1.3) which is a homologue of the vanadate resistance protein GOG5/VRG4 from *Saccharomyces cerevisiae* (21% identity plus 26% similarity over 69% of GufC, with an E-value of 9e-12)(Poster and Dean, 1996). The other homologues were an Sqv-7-like protein from man involved in the glycosylation pathway (21% identity plus 24% similarity over 75% of GufC, with an E-value of 2e-11) and a human homologue of a protein encoded in cosmid C52E12 of *C. elegans* required for vulval invagination (29% identity plus 26% similarity over 75% of GufC, with an E-value of 3e-10).



The *crtI* region

Figure 3-9 FRAME Analysis of the crtI *region*

Putative open reading frames are marked in grey. The ORF between positions ~250 and ~1850 is the crtI *gene. The ORF between 2100 and 2850 is* gufB. *The presence of an additonal gene reading from ~100 in the opposite orientation to* crtI *is shown and designated* gufC. *The* crtI *region is shown below, with areas of interest marked (see later).*

```
GTGAGCGCCG CGCCATCCAG TGGCGCGCTC ACCCGGCAGG CGGGGAAGGG GCGCCCGCAT 1965
 ***
CCGCCGGGCG GGGCGGGGGC GAAGCTGGCA TAGTTGCGAG GCATGCCCTT CTTCATCCCA 2025
                                          M  P  F   F  I  P

TTCGCGGTGG GTGGCCTGGT GCTGACGGCA CTGGGCCTGG GTGTGAGGAA GGTGCTGACG 2085
F  A  V  G   G  L  V   L  T  A   L  G  L  G   V  R  K   V  L  T

GAGACGGGCG TCACCACGCC TGGGGATGCC CGGCTGGCCG AGGCGCGTGA GCGGCACCGC 2145
E  T  G  V   T  T  P   G  D  A   R  L  A  E   A  R  E   R  H  R

GGGGCGGTGG CCGCGCTCCG GGCGGACCGC CTCCAGGTTC GCGATGGCGT GGCCACCCAT 2205
G  A  V  A   A  L  R   A  D  R   L  Q  V  R   D  G  V   A  T  H

GGCGCGCTCC AGGCGCGGGT GCACGTGGAG GTGCTGGTGC CCTTCGGCGC GCTGTTGGAG 2265
G  A  L  Q   A  R  V   H  V  E   V  L  V  P   F  G  A   L  L  E

CGGCTGGAGC GCTGGGGGCA CGTCCAGGAA GCCGAGCTGC TCGAACCCGA GGCGCTGGAG 2325
R  L  E  R   W  G  H   V  Q  E   A  E  L  L   E  P  E   A  L  E

GCGCTGCGCG CGCTCCTGCG CGAATCCCCG TCCCGCGCGA CGCGGCGAAA CTGGCCCCTG 2385
A  L  R  A   L  L  R   E  S  P   S  R  A  T   R  R  N   W  P  L

CTGGGCGCGG GTGCCGAGGT GCCCGCCGCG CTGGAGTCCG TGCTGGCGTG GCTGGACCGG 2445
L  G  A  G   A  E  V   P  A  A   L  E  S  V   L  A  W   L  D  R

GGCTGGCTGG ACGAGGATTC GCCGCCGGTG GTGCTGGATG GGACGCCGCT GTACGCAGTC 2505
G  W  L  D   E  D  S   P  P  V   V  L  D  G   T  P  L   Y  A  V

ATCTCGGCCC GCGCCATCCT GGCGGGGAGC GCTCCGGAGG AGGGTGCGCG GGCGCTCGAT 2565
I  S  A  R   A  I  L   A  G  S   A  P  E  E   G  A  R   A  L  D

GAGGCCTCCG CGCTGCTGGC TCGGACGACG GCGTTCCTGG GGGCGCTGCG CGTGCGGCTG 2625
E  A  S  A   L  L  A   R  T  T   A  F  L  G   A  L  R   V  R  L

ACGGCGCTGG AGCAGCGGGT GGCGGGGCTG CATGGACGGG CCTCGGCGCA GCTCGCCTAC 2685
T  A  L  E   Q  R  V   A  G  L   H  G  R  A   S  A  Q   L  A  Y

CTGGACGCGG CCAGCTTCGA GGCGGGCGGC GAGGAGCCCC GGGAGCGGCT GACGCGGCTG 2745
L  D  A  A   S  F  E   A  G  G   E  E  P  R   E  R  L   T  R  L

GCGGTGCTCG TGGGCCAACT CGCCGTGCTG CTGCGCACGC CCGTGCTGAA CTCGGAGGGG 2805
A  V  L  V   G  Q  L   A  V  L   L  R  T  P   V  L  N   S  E  G

CGTCTGACGC CCATGCTCGC CGCGCGTGCG GAGGACGACG CGCCTTCCAA CGGTTGATGC 2865
R  L  T  P   M  L  A   A  R  A   E  D  D  A   P  S  N   G  ***

AGCGGCACCG TGGCCCGCGC TCGAGCGCGG GCGAGCCGCC GCATCAGACG TAGCGTTCAG 2925
```

Figure 3-10 DNA sequence of the region downstream of crtI.

The translation product of gufB *is shown in single letter code below the corresponding DNA sequence. Asterisks denote stop codons. The stop codon beginning at position 1906 is the stop codon for* crtI *as revealed by Fontes* et al., *1993.*

```
CCGCACGCTG CCTCCCTCGC TTGCCATGCC CTACCGGATG CCAAGGCCCT CCGGTACCTG
                          Met

GGTGAACGGA ATCCCACCCA CTCATTCACC TGGCGAACGC GCGAACCCGC CAGGACGTTA
                                           +1            -10

CAAGAGCACC GTCCGGCACT GGCCCCCGGG CGAGCGCCAT TCCAGCCTTC CTTGGCCACC
           -35

GTGCCACCGC ATTCCTCCTC CCACCTGCGC CGCAACGCCG CCGCCTACTG CGTGGTCCTG
M  P  P  H   S  S  S   H  L  R   R  N  A  A   A  Y  C   V  V  L

GTGGGGCTGC TCCTCACCGC CGTGTCGGCG ACCTATGTGC AGCAGAGCAT CCACGAGCGC
V  G  L  L   L  T  A   V  S  A   T  Y  V  Q   Q  S  I   H  E  R

CGCCTGCACC GCTTCGACGG CGCGGTCCAT GACGGGGTGC TGGGC
R  L  H  R   F  D  G   A  V  H   D  G  V  L   G
```

Figure 3-11 DNA sequence of gufC *and its relationship to* crtI.

Primary structure of GufC is denoted by single letter amino acid code underneath the coding DNA sequence. Met denotes the initiating codon of crtI. *Bases in bold show the -10 and -35 consensus hexamers of the* crtI *promoter and the +1 position. It is assumed that the initiation codon GTG encodes a formyl-methionine residue.*

3.3.3. Clues to the structure and functions of GufB and GufC.

Due to the upsurge in predictive analyses of genomes and proteins made necessary by entire genome sequencing projects, there are an abundance of public-domain world-wide web servers which provide predictive algorithms for the analysis of primary DNA and protein structures. The sequences of GufB and GufC were extensively analysed using these resources. It should be noted that at best these analyses give highly speculative predictions, but may suggest directions which future experimental work may prove fruitful.

3.3.3.1. GufB.

The homologue which showed the greatest similarity to GufB is a member of the PE PGRS family (Pro-Glu, with **p**olymorphic **G**C-rich **r**epetitive **s**equences), whose existance was discovered during the sequencing of the entire genome of *Mycobacterium tuberculosis* (Cole *et al.*, 1998). There are 99 PE family proteins in the *M. tuberculosis* genome including 61 members of the PGRS sub-family. The genes for the PE PGRS proteins are clustered around the *M. tuberculosis* genome and are based around multiple copies of repetitive polymorphic repeats or PGRSs. PE refers to a conserved Pro-Glu (PE) motif near the N-terminal of the majority of these proteins. All PE members have a conserved N-terminal region of ~110 residues which is believed to form a globular domain (Cole *et al.*, 1998). The C-terminal regions of the proteins can be variable in length, with some proteins lacking any extension to the globular domain and others having a 100-1400bp C-terminal domain. Only one member of the PE PGRS family has had a function ascribed to it, being an esterase/lipase (Cole *et al.*, 1998). The abundance of a highly conserved family of proteins in a potent pathogen has led to the suggestion that the PE proteins serve to provide a source of antigenic variation. Polymorphism seems to occur in the PE PGRS proteins as a result of genetic instability in the tandem repeats within the PGRS moiety (Cole *et al.*, 1998 and Robertson and Meyer, 1992). Homology between GufB and the PE proteins extends from around residue 165 to the end (residue 284) of GufB and over the

first 110 residues of the PE proteins, i.e. the globular N-terminal domain common to all PE proteins. GufB does not share the PE motif (Figure 3-12).

It is possible that the conserved domain is a general protein folding domain, yet its absence from the database except for the PE proteins suggests this is not the case. Alternatively, the globular domain may provide a generic module for membrane association, but once again the lack of homologs outside of *M. tuberculosis* argues against this possibility. A role for GufB in immunological evasion is also unlikely as GufB lacks homology to the PGRS domain of the PE PGRS proteins and *M. xanthus* is non-pathogenic. The significance of the PE globular domain in GufB is therefore obscure.

```
GufB    165   AVISARAILAGSAPEEGARALDEASALLARTTAFLGALRVRLTALEQRVAGL
              A     A  A    + G+       +A  A TT+ L A    ++A    +
wag22   6     AVPETIAAAATDLADLGSTIAGANAAAAANTTSLLAAGADEISAAIAALFGA
Z96071  6     AAPEIVVAAATDLAGIGSAISAANAAAAAPTTAVLAAGADEVSAAIAALFSG
Z96800  6     AQPEMIAAAAGELASIRSAINAANAAAAAQTTGVMSAAADEVSTAVAALFSS

GufB    217   HGRASAQLAYLDAASFEAGGEEPR--ERLTRLAVLVGQLAVLLRTPVLNSEG
              H +A         AAS +A     +  + LT                P+L +
wag22   57    HGRA------YQAASAEAAAFHGRFVQALTTGGGAYAAAEAAAVTPLLNSIN
Z96071  57    HAQA------YQALSAQAAAFHQQFVQTLAGGAGAYAAAEAQVEQQLLAAIN
Z96800  57    HAQA------YQAASAQAAAFHAQVVRTLTVDAGAYASAEAANAGPMLAAVN

GufB    266   RLTPMLAARAEDDAPSNG
                   L  R       +NG
wag22   101   APVLAATGRPLIGNGANG
Z96071  101   APTQALLGRPLIGNGADG
Z96800  101   APAQALLGRPLIGNGANG
```

Figure 3-12 Homology between PE PGRS members wag22, Z96071, Z96800 and GufB.

'+' denotes similar residues. For this purpose, the following amino acids are considered to be similar: (I,L,M,V), (H,K,R), (D,E,N,Q), (A,G), (F,Y,W) and (S,T). The helix-turn-helix motif of region 4.2 is underlined. The PE motifs of the PE PGRS family members are underlined as is the potential leucine zipper of GufB (see below).

In order to determine whether GufB contained any transmembrane (TM) helices, its sequence was passed through the TMpred algorithm on the ISREC server

(www.ch.embnet.org/) (Rost *et al.*, 1995). Two models were suggested, a favoured model involving three TM helices and a less favoured model involving two. The first, much preferred, model suggested helices at residues 3-19 (in-out), 123-144 (out-in) and 246-262 (in-out) {Helices 1, 2 and 3}, while the second predicted helices at residues 3-21 (out-in) and 246-262 (in-out) {Helices 1 and 3}. Charged residues at the termini of TM helices tend to be distributed with positive charges on the inside and negatively charged residues on the outside. It is thought that as the cytoplasm is negatively charged relative to the extra-cytoplasmic space, distribution of charges serves to lock the TM helix in position (von Heijne, 1989). There are regions of positively charged residues at position 22 corresponding to the C-terminal of helix 1, around position 120 at the N-terminal end of helix 2, and at the N-terminal of helix 3. Negative charges are found at the C-termini of helices 2 and 3 (Figure 3-13). This suggests that the two-helix model may be a valid alternative, with helix 1 going out-in and helix 3 being in-out.

```
                HELIX 1
1     MPFFIPFAVGGLVLTALGLGVRKVLTETGVTTPGDARLAEARERHRGAVAALRA
                              ++    -        - +   - +-+++       +

54    DRLQVRDGVATHGALQARVHVEVLVPFGALLERLERWGHVQEAELLEPEALEAL
      -+    +-     +      + + -           -+ -+   +   - -   - -   -

                                HELIX 2
108   RALLRESPSRATRRNWPLLGAGAEVPAALESVLAWLDRGWLDEDSPPVVLDGTP
      +    +-    +  ++           -      -       -+    ---       -

162   LYAVISARAILAGSAPEEGARALDEASALLARTTAFLGALRVRLTALEQRVAGL
            +          --   +   --       +          + +     - +

                                                HELIX 3
216   HGRASAQLAYLDAASFEAGGEEPRERLTRLAVLVGQLAVLLRTPVLNSEGRLTP
      + +          -     -    -- +-+   +               +       - +

270   MLAARAEDDAPSNG
          + ---
```

Figure 3-13 Distribution of charged residues and predicted TM helices within GufB.

Glu and Asp residues are asumed to be negatively charged, while His, Arg and Lys residues are positive. Potential TM helices are underlined.

The sequence of GufB was also searched to determine whether any protein motifs could be found within its structure using 'ScanProsite', again on the ISREC server (at www.expasy.ch/) (Bairoch *et al.*, 1997). In this way the presence of a possible leucine zipper motif was discovered at residues 185-206.

185 LDEASALLARTTAFLGALRVRL 206

The leucine zipper is a protein dimerisation motif originally described for eukaryotic DNA binding proteins including Fos and Jun. In this motif, there is a leucine residue every seventh position such that if the region of the zipper is in an α-helical conformation (with ~3.5 residues/helical turn), the leucines line up on one face of the helix, with one leucine every two helical turns. Two proteins, each with a leucine zipper may then align themselves such that the leucine side-chains on their respective helices interdigitate (Landschulz *et al.*, 1988). Although mainly found in eukaryotes, leucine zippers have also been described previously in prokaryotes including a putative example in CarD (Nicolas *et al.*, 1996). The significance of the leucine zipper is unclear. It may function as a dimerisation domain if that region of GufB is α-helical. None of GufB's homologues have leucine zippers, so its presence in GufB may instead be a coincidental result of the high (17%) proportion of leucine residues within the protein which is probably due to the codon position bias of the organism.

The predictive algorithm of Altschul *et al.* (1990) confidently suggests that GufB is mainly α-helical (67%), with very little β-sheet (7%), three transmembrane helices of around 20 residues and includes helices which seem particularly long (up to 40 residues). The algorithms used to compare proteins according to organisation of secondary structure, are particularly inappropriate to study membrane proteins as the TM helices are modelled as if they were general α-helices. When these programs were used regardless, a hit of probability (0.7) was found to the cytoplasmic tetracycline repressor (Gatz and Quail, 1988), which may be purely a consequence of the high α-helical nature shared by GufB and TetR. Searching the PROSITE database of protein domains with the sequence of GufB

(Altschul *et al.*, 1990) scored a weak hit on a gene within *flaA* of *Bacillus subtilis* (E-value of 7e-2). The similarity was found to *orf6* of the *flaA* locus which is believed to be anchored at or within the membrane (Albertini *et al.*, 1991). Interestingly, this domain is homologous to *frzCD* of *M. xanthus* which is a homologue of the MCPs involved in chemotaxis. MCPs (methyl-accepting chemotaxis proteins) are membrane-spanning chemoreceptors involved in transduction of a chemotactic signal to the cell's motility system.

The predictors of secondary structures were very certain that GufB is predominantly α-helical and contains three TM helices. Further structural predictions about GufB seem premature at best.

The role of GufB is also still uncertain, however certain features of its sequence suggest speculative possibilities. The postulated presence of a translated leader sequence which causes ribosomes to be taken from the end of *crtI* to the beginning of *gufB* suggests that the two genes are both involved in carotenogenesis. One of the homologous PE proteins is an esterase/lipase. It is conceivable that GufB is also an esterase/lipase, possibly even involved in carotenogenesis, as the gene product required for the addition of a glucoside ester to carotenoids (i.e. the conversion of 7,8 dihydro, 3-4' dehydro-rhodopin to 7,8 dihydro, 3-4' dehydro-rhodopin glucoside ester) has not yet been identified. The TM helices of GufB imply a localisation of GufB in the membrane, consistant with the site of action of carotenogenic enzymes. If this scenario is correct it will prove necessary to determine conclusively whether *crtI* and *gufB* are co-transcribed or are expressed independently. The upstream region of *gufB* certainly does not appear to contain any obvious promoter sequences (although my opinions on assigning promoters on the basis of similarity to a consensus will be aired later), so a single operon model is perhaps most likely. If *crtI* and *gufB* are part of a single operon, it is possible that there are other genes downstream of *gufB* which are also in the operon. Future work will extend sequencing downstream of *gufB* to determine whether further genes are present.

3.3.3.2. GufC.

It is hard to find any constancy of function between the three homologues of GufC yet identified. Of the three homologues, the only one which may suggest a role for GufC in *M. xanthus* is the gene in *C. elegans* which is homologous to the vanadate resistance protein GOG5/VRG4 of *Saccharomyces cerevisiae.* VRG4 is responsible for correct glycosylation and trafficking of proteins through the Golgi apparatus and vanadate resistance seems to occur due to underglycosylation of secreted invertase. Whether any parrallels can be drawn between the function of VRG4 and GufC is unlikely and would be speculative in the extreme.

Searching GufC for protein motifs gave no hits except for a single transmembrane helix running between residues 13 and 33, with a particularly favoured orientation running in-out, with a negatively charged Glu residue at the C-terminal of the helix and positive residues at the N-terminal. The paucity of infomation gleaned about GufC is certainly a consequence of the small amount of its sequence that is known, although the ability to actually find homologues of such a small region of known sequence is testament to the power of the iterative PSI-BLAST program.

3.3.3.3. Consequences to *crtI* due to *gufC*.

Both *crtI* and *carQRS* have genes of unknown function reading divergently from them. For *carQRS* it has been postulated that correct expression of *gufA* is required for normal transcription (McGowan, 1992). Whether the same might be true for *crtI* and *gufC* is unknown. The transcriptional start site of *crtI* is only 80bp from the initiating codon of *gufC* and therefore the promoters of the two genes would be very likely to overlap. This may have regulatory consequences for expression of the two genes. Scenarios can be envisioned where expression of one gene precludes expression of the other, or as is the case for *carQRS*/*gufA*, expression of one gene may be dependent on expression of the other. The effect of transcriptional activators/repressors of either gene could also have profound

effects on activity of the other gene. It must be noted that the site-directed mutations introduced into the *crtI* promoter by Martinez-Argudo *et al.* (1988) may be exerting effects on the expression of *crtI* indirectly, through a direct effect on expression from the *gufC* promoter. Which explanation is correct would be difficult if not impossible to determine, but the two alternative possibilities must not be overlooked.

In short, although the situation for *crtI/gufC* is analagous to that of *carQRS/gufA*, the inter-relation of the promoters of *crtI* and *gufA* may reflect why expression of *crtI* is so different from that of *carQRS*.

It is possible that the function of *gufA* and *gufC* is related. GufA and GufC both have TM helices within their first 30-40 residues and it is conceivable that they have a similar role in the cell. However, any role may be as basic as to exert an effect on expression of the carotenogenic loci *crtI* and *carQRS*. It should be pointed out that this is speculation and should be taken with several large pinches of salt.

3.3.4. Why does a *crtI* mutation render strains Car⁻?

In all strains created and isolated, which contain a *crtI1* allele, there is a Car⁻ phenotype. Initial thoughts suggested that as expression of *crtEBDC* alone is sufficient to cause production of carotenoids (i.e. in a *carS* constitutive mutant or in a *carA* mutant), then a *crtI* mutant in the light would also be able to create coloured carotenoids. This was seen not to be the case and suggested that the transposon insertion within *crtI* in the *crtI1* allele was also causing the disruption of unidentified gene or genes involved in carotenogenesis, down-stream of *crtI* - i.e. a polar effect. The postulated downstream gene(s) could potentially include carotenogenic enzymes, or a positive regulator of *crtEBDC* expression. Sequencing and subsequent analysis of the region downstream of *crtI* identified a gene (*gufB*) which may be translationally coupled to *crtI* and thus possibly involved with carotenogenesis. GufB shares homology to the PE protein family of *M. tuberculosis* which includes an esterase/lipase. It is possible that GufB possesses an esterase activity involved in carotenogenesis. If this is true, then the

initial postulation that carotenogenic genes would be found downstream of *crtI* would appear to be vindicated.

However, a more likely explanation for the Car$^-$ phenotype of a *crtI* mutant arises due to the nature of the carotenogenic pathway of *M. xanthus*. CrtI catalyses the conversion of phytoene into lycopene, a conversion which is absolutely required for the production of coloured carotenoids. Thus by preventing the production of CrtI, the *crtI1* allele is causing a block in the carotenogenic biosynthetic pathway. This hypothesis is strengthened by the data available about the amount of each intermediate in the carotenoid pathway in various mutant strains (Martinez-Laborda *et al.*, 1990). It is clear that in any strain carrying a *crtI* lesion there is an accumulation of phytoene, with traces of pytofluene, but with no detectable levels of any intermediates further down the pathway. There is therefore a strong case for the Car$^-$ phenotype of a *crtI* mutant being due solely to removal of the enzyme catalysing an essential step in the pathway of carotenogenesis. The constitutive production of carotenoids in a *carA* mutant in the dark will involve a low-level basal expression of *crtI* which is independent of expression of the *carQRS* operon.

3.4. Conclusions

The *crtI* gene encodes phytoene dehydrogenase which catalyses the conversion of phytoene into lycopene, an essential step in the pathway of carotenoid production. Thus *crtI* mutants are Car$^-$ and appear tan. Expression of *crtI* is biphasic and both phases are activated by both CarQ, with phase II also seemingly activated by carbon limitation. Additionally, there is a two-fold induction of *crtI* expression in the light which is independent of CarQ.

There are several features of *crtI* which are as yet unexplained. Why is expression of *crtI* more complicated than that of *carQRS*? Why is *crtI* catabolite-repressed? Why can *in vitro* run-offs transcripts be obtained from P^{carQRS} but not for P^{crtI}? All these queries suggest that there are many aspects of the expression of *crtI* that are not yet understood. Ongoing work in our lab and that of F. J. Murillo will hopefully shed more light on these areas of uncertainty.

Analysis of the DNA sequence of the *crtI* region obtained by Fontes *et al.* (1993) and novel DNA sequence downstream of *crtI* obtained in this study, indicates the presence of two additional genes at this locus. Translationally coupled to the end of *crtI* appears to be encoded a leader peptide (OlpA) which terminates immediately pior to *gufB* and implies a level of translational coupling between the two genes. Gene *gufB* starts around 100bp downstream of the end of *crtI* and encodes a membrane protein which is homologous to a family of proteins of unknown function in *Mycobacterium tuberculosis*. It is likely that *gufB* is transcribed with *crtI* as a single operon and is also involved with carotenogenesis. The second novel gene, *gufC* is divergent from *crtI* and has homologues of diverse function. Any potential role of *gufB* or *gufC* in carotenogenesis is unclear, although it is clear from the close proximity of the promoters for *crtI* and *gufC* that expression of these two genes must be inter-related.

Future work will entail complete sequencing of *gufC* and analysis of its coding sequence. Sequencing further downstream of *gufB* will determine whether any more genes fall in the putative *crtI* operon. Creation of disrupting *lacZ* transcriptional fusions of both *gufB* and *gufC* will potentially allow

characterisation of the roles and expression of these two genes. Other techniques which may be used to map transcription of these genes include primer extension analysis to identify transcriptional start sites (and therefore promoter sequences) and RT-PCR (reverse transcriptase PCR) and northern blotting to determine whether the genes are transcribed separately or as polycistronic mRNAs.

4. Molecular Analysis of the *carQRS* Promoter.

4.1. Introduction

Promoters are the prime cis-acting genetic elements controlling differential gene expression. Each gene is transcribed from a promoter and in most cases, an increase in transcription of a gene leads to a proportional increase in gene product production. Therefore the recognition of a promoter by the cellular transcriptional machinery and the resulting production of mRNA is of paramount importance to the cell. The levels of transcription of a particular gene are determined by availability of sigma factor, the nature of the gene's promoter element and how it interacts with the transcriptional complex and other factors/genetic elements.
Determination of which promoters might be expressed is under the control of the sigma factor within the initiating RNA polymerase complex. Each sigma factor mediates the recognition of a sub-set of promoters by RNA polymerase, dependent on the sequence of the promoter. For promoters recognised by sigma factors of the σ^{70} family, the most important aspects of the promoter required for recognition are hexamers at -35 and -10, and the spacing between these two hexamers. However, due to experimental difficulties, little is known of the exact nature of interactions between sigma factors and their cognate promoters (for review see deHaseth *et al.*, 1998).

4.1.1. Initiation of transcription.

There are thought to be around 5,000 genes and 2,000 promoters in the genome of *E. coli*. In a genome of approximately 4Mb of DNA, it is extremely important that RNA polymerase can differentiate between promoters and the far larger non-promoter DNA. RNA polymerase within the cell exists in two main forms, the core enzyme which is responsible for elongation during RNA synthesis and has a sub-unit composition of $\alpha_2\beta\beta'$, and the initiating holoenzyme which is core

enzyme with associated sigma factor ($\alpha_2\beta\beta'\sigma$). Specificity for promoters is mediated by the σ (or specificity) factor of the RNA polymerase holoenzyme.

The initial events in transcription by RNA polymerase are DNA binding and promoter selection. There are many models as to how this is achieved by the holoenzyme including: on/off binding to DNA, with complexes at non-promoter DNA thermodynamically unstable relative to binding at promoters; non-specific binding to DNA followed by a one-dimensional walk to promoter sites and several other variants (von Hippel and Berg, 1989).

When the holoenzyme binds to a promoter it forms a closed complex (RP_C) with the DNA of the promoter in a 'closed' duplex state. There then occurs a significant isomerisation as the closed complex changes conformation to an open complex (RP_O). This isomerisation is commensurate with melting of the DNA between positions -9 and +3, as shown by susceptibility to cross-linking agents which can attack only single-stranded DNA (deHaseth and Helmann, 1995). The isomerisation is also accompanied by a substantial change in the protection of promoter bases from DNase I cleavage. In the closed complex protected bases extend from around -5 to approximately -55, whereas in the open complex protection extends from -55 to around +20 (Schickor *et al.*, 1990). The closed to open conversion involves two, poorly characterised, intermediates (Craig *et al.*, 1995). In the second of these intermediates (RP_{C2}), the two promoter hexamers are 'over-rotated' by 68^0 relative to each other and this torsional strain produced by the RNA polymerase seems to induce opening of the DNA duplex (deHaseth and Helmann, 1995). The melting of duplex DNA is the rate limiting step at some promoters and is enhanced by negative supercoiling which reduces the energy requirement associated with duplex melting.

RNA production now commences within the β sub-unit of the open complex. While the transcription complex is in the initiating mode, transcription is not guaranteed and transcripts up to around 9 residues can be aborted without the polymerase having moved off the promoter (Krummel and Chamberlin, 1989). It seems that it is escape from the initiating mode into the elongation mode that is the rate-determining step in transcription, a process that on average has a half-life of around 60s. The sigma factor is still attached to the RNA polymerase complex

while it is in initiating mode and when the nascent transcript has reached a size of 9-16 residues, the complex enters the elongation mode and the sigma factor is released.

As the complex enters the elongation mode there are once again significant changes to the structure of the complex. The complex becomes more compact and stable and there is a brief extension of the protected region of DNA to position +24. The region of melted DNA in the transcription bubble also increases in size, growing from 10 to 18 nucleotides (Yager and von Hippel, 1991). The sigma factor is then released from the complex leaving the core polymerase to continue polymerisation and leaving the sigma factor free to interact with a further RNA polymerase core enzyme and mediate initiation of another round of transcription.

4.1.2. Promoters of *M. xanthus*.

Knowledge of promoters in *M. xanthus* is limited with very few identified. In 1995 the transcriptional start sites for only 8 genes were known (Keseler and Kaiser, 1995), and not many more have been identified since, making determination of consensus sequences difficult. Another complication is that there are multiple sigma factors in *M. xanthus* and for the majority of promoters, it is unknown which sigma factor mediates their transcription. Until it is known which sigma factor is responsible for expression from which particular promoters, it will prove impossible to determine a consensus promoter sequence for the housekeeping and alternative *M. xanthus* sigma factors. A brief summary of the sequences and proposed regions of interest of known *M. xanthus* promoters is now given. It has been attempted to group them tentatively by cognate sigma factor (speculatively in most cases).

σ^{70}-like and constitutively expressed promoters.

pkn5	AAATCCTACGGACCACGTACAGTCAGGGTCAACGTTGGCAAC**G**
pkn6	CAACGTTGACCCTGACGTGACGTGGTCCGTAGGATTTCGACA**T**
4403	GGTTGATTCATGAATAAGCCGTTTTTGATGTACACCCGTTTT**A**CC
frzZ	TTTCGATGGTCCGCGCTGCACCAAAGGTGTAGGGTTCTACCG**C**G**A**
relA	CCGTTCACAGTTCGGGCAGACCACGGCGAGGGAGGACA**G**CA
sasS	GGTCATGGGACAGCGCCGTGAGGCGCCACATATAAAGTACGC**A**CA
vegA	TTCCTTTAGACAAAACCATTTTTGGAAGGTAAGGGTATGGGC**A**
crtEBDC	CCTTGACAAGCTCTGGACGCAAACGCTACCTCTAGGAA**A**
carD	GTGGCGTCCGACGTGGCTCGGTGCTAAGCCCAGGCCT**G**TT
rpoE1	TTGTTCCCGTGCATACATTTGGCGGGTAAGCGACAACTCATTC**C**

Putative and proven SigECF-dependent promoters.

CarQ-dependent:

carQRS	CGAGCGCCGGAAACACTTTCGCAGGTGGCCCGTAGAGGAGTC**G**
crtI	GGTGCTCTTGTAACGTCCTGGCGGGTTCGCGCGTTCGCCAGG**T**

SigE-dependent:

rpoE1	GCTAGGGAATATGTTTCCCTGGCCGGCGTCGTGTATTCCCGA
rpoE2	ACTCGGGAAGGCATGTTTCCTGGCTCAATATCTTGCAATTCACG

unidentified promoters.

4400	GGCCGGAGGCGCGAGGTGCATGCCGGCGCTACAACACCCCCG**G**TC
gufA	GCATCCCGTCCTGCCCTGGAGGGCGAGGCCGTTAGTGTTGGG**C**

σ^{54}-like promoters.

carA	TTGGGCGAAGCGCTCTTGCTTGCGGCCTTCCTGCGCT
4521	GTCGAGCACGCG TCTTGCTTTGGCTCACGGC**T**CTTC
mbhA	GAATGGCACGCCATCT GCTTCGCGGCTGCGC**G**GAGC

Figure 4-1 The -35 and -10 regions of promoters from M. xanthus.

Areas of interest are underlined and transcriptional start sites (if known) are in bold. These are regions identified by the respective authors, and obvious regions of homology in other promoters. Promoters are classed according to which sigma factor they are probably recognised by. This is speculatively assigned according either to evidence provided by the authors or by obvious homology. Promoter sequences from: pkn5/pkn6 *Zhang* et al. *(1996): 4521 Keseler and Kaiser (1995): 4403 Fisseha* et al. *(1996):* vegA *Komano* et al. *(1987):* aphII *Biran and Kroos (1997):* mbh *Romeo and Zusman (1991):* relA *Harris* et al. *(1998):* sasS *Yang and Kaplan (1997): 4400 Brandner and Kroos (1998):* rpoE1 *Ward* et al. *(1998):* rpoE2 *Ward M. (pers.comm.):* frzZ *: Trudeau* et al. *(1996).*

Cursory examination presents some immediate observations. There are no very obvious homologies at the -35 and -10 regions detectable by direct sequence comparison. Additionally, if a comparison is made to the *E. coli* consensus promoter sequence for Sig70 (the housekeeping sigma factor of *E. coli*), (TTGACA...TATAAT) similar motifs can be seen in the *M. xanthus* promoters but not in the same positions and not consistantly. For instance, the *relA*, *pkn6* and *crtEBDC* promoters have very similar hexamers at -35 to *E. coli* but one base pair displaced from each other. Additionally of these three promoters, *pkn6* has a hexamer around -10 which is very similar to that of *E. coli*, however, neither *relA* of *crtEBDC* have any apparent similarity to the -10 hexamer of *E. coli*. Yet homology at -10 can be found among *M. xanthus* promoters (e.g. promoters of *pkn5* and *frzZ*). This makes it very difficult to define a consensus for any of the *M. xanthus* sigma factors and is probably a consequence of the multiple sigma factors present in *M. xanthus*. The presence of hexamers which have a great degree of similarity to the consensus promoter sequence of *E. coli* suggests that there is at least one sigma factor in *M. xanthus* which recognises a similar consensus to that recognised by Sig70 of *E. coli*. The gene encoding Sig70 of *M. xanthus* (*rpoD*) has been cloned, sequenced and the gene product shown to be an active sigma factor *in vitro* (Inouye, 1990; Biran and Kroos, 1997; Davis *et al.*, 1995).

Generally, there are lots of problems associated with the identification of promoter sequences. In most cases a promoter sequence can only be identified through experimental procedures such as primer extensions or S1 nuclease studies which determine transcriptional start sites. These results can be ambiguous and are certainly not unequivocal. A primer extension analysis which gives a slightly displaced transcriptional start site, can cause significant problems when attempting to make comparisons with other promoters. An ambiguity in the length of the spacer region between the -10 and -35 sites compounds uncertainty in the exact position of the various promoter elements and thus in assignments of -10 and -35 regions. Difficulties in precisely defining promoter sequences may explain why it is so hard to find a consensus for SigA-dependent promoters.

Another difficulty arises from the nature of the search for consensus itself. It is impossible to define a promoter as being Sig70-dependent until it has been proven to be so experimentally. This is also true for all promoters of any given sigma factor. Assignment of promoter sequences purely through homology to known promoter sequences thought to be recognised by the same sigma factor, is a recipe for disaster, yet the temptation to do so is omnipresent. Ward *et al.* (1998) have recently discovered what appears to be two copies of *rpoE* in *M. xanthus*. The gene for RpoE1 was cloned during attempts to use the yeast two-hybrid system to screen *M. xanthus* for genes whose products interacted with the chemotaxis protein FrzZ. They assumed from the sequence of RpoE1 that it was an ECF-sigma factor (SigECF) and postulated that *rpoE1* would be under the control of an RpoE-dependent promoter. The DNA sequence upstream of *rpoE1* was searched for regions with homology to known promoter consensus sequences and a region of homology to an SigECF-dependent promoter consensus and a region similar to a Sig70-dependent promoter were found. Subsequent expression studies showed however that *rpoE1* is exclusively transcribed from the Sig70-like promoter and not from the SigECF-dependent promoter, in all growth conditions assessed. The validity of the proposed SigECF-dependent promoter for *rpoE1* is therefore in doubt and emphasises the folly in assigning promoter sequences from sequence data only. In effect, the attempt to find an SigECF-dependent promoter upstream of *rpoE1* was biased towards successfully finding one. When found, experimental work showed that the 'promoter' was not transcribed *in vivo* and therefore not a promoter, yet the conclusion arrived at by the authors is that the conditions under which the SigECF-dependent promoter is active have not yet been determined.

The need to derive a consensus promoter sequence for a sigma factor is far from clear. Even were it possible to unambiguously determine the consensus promoter sequence, it is impossible to deduce anything about the activity of any promoter without experimental work, except perhaps to identify which sigma factor is most likely to cause its transcription. Even this scenario is unlikely, as assignment of a specific sigma factor to a particular promoter, requires the promoter to be similar to the consensus promoter, a situation which is not necessarily going to be the case.

The folly of assigning promoter activity to regions of DNA by sequence homology alone is probably going to be compounded however, during the upcoming project to completely sequence the *M. xanthus* genome. While a noble goal, current understanding is too limited to make this a viable option for promoter identification.

In some cases, with sufficient genetic evidence it is possible to assign promoters unambiguously. The *carQRS* and *crtI* promoters are two good examples of this. It is known that both promoters are completely dependent on CarQ for their expression, thus they presumably share the same sigma factor and may therefore have similar promoter sequences. This can be seen to be the case on examination with identical hexamers at -35, 67% identity at -10 and with identical spacing between the two recognition hexamers, but only if the proposed start site of transcription of *crtI* is displaced by 12bp. The identification of these regions of DNA as promoter elements is unambiguous, however, even with such clear evidence, there is still uncertainty as to the exact nature of the sub-elements within both of these promoters. For instance, the expression of *crtI* is very different from that of *carQRS*. Run-off transcription assays show that CarQ causes transcription of the *carQRS* promoter but the same was not true for the *crtI* promoter (Browning, 1997). Additionally, the *crtI* promoter is regulated by carbon availability which is not the case for the *carQRS* promoter. The repressive effect of carbon source availability may be specific to *crtI* or due to a non-specific consequence of nutritional state such as the degree of DNA supercoiling which has been shown to be the case in response to osmotic and anaerobic stress in enteric bacteria (Bhriain *et al.*, 1989). It is still possible that CarQ is not the sigma factor required for *crtI* expression, and that the requirement of CarQ for *crtI* expression is due to CarQ mediating the production of an additional factor at an as yet unidentified locus, which then goes on to activate *crtI* transcription. Alternatively, *crtI* expression may require an activating protein which was not supplied during the run-off assays.

4.1.3. ECF sigma factor-dependent promoters.

The original description of the ECF sub-family of sigma factors (section 5.2.3) (SigECFs) described a clear similarity of sequence for SigECF-dependent promoters. It should be noted that the similarity profile was derived from a large number of different species (Lonetto *et al.*, 1994).

SigECF-dependent promoters have a highly conserved -35 hexamer but with significantly less identity at the -10 hexamer. If SigECF-dependent promoters had a conserved -10 region as well as a conserved -35 region, it would prove impossible for multiple SigECFs to be maintained with independent functions within the same species. The reason for this difference in conservation between the -35 and -10 hexamers is unclear, particularly concerning the great degree of similarity at the -35 hexamer. It is conceivable that the ECF sigma factors will prove to be degenerate in their recognition of SigECF-dependent promoters due to the similarities between different SigECF-dependent promoters even within the same species.

4.1.4. Promoters of the *M. xanthus* carotenogenic regulon.

At the moment there is considerable confusion in the literature about the nature of the *crtI* promoter. Currently, evidence for the start site of the *crtI* promoter comes solely from primer extension analysis (Fontes *et al.* 1993). Evidence for the start site of the *carQRS* promoter comes from both primer extension experiments (McGowan *et al.* 1993) and from transcriptional run-off assays (Browning 1997) and thus is more certain than that of *crtI*. Fontes *et al.*, (1993) and Martinez-Argudo *et al.* (1998) describe a start site that results in -10 and -35 sequences that

show only poor similarity to those of the *carQRS* promoter. If the start site is shifted upstream by 12bp relative to that described by the above authors then there appears to be a striking degree of similarity to the *carQRS* promoter (Figure 4-2). The *crtI* promoter with the 12bp displacement will henceforth be refered to as the *crtI* (+12) promoter)

No other promoters in *M. xanthus* appear to resemble the *carQRS* or *crtI* promoters, including that of *crtEBDC* (which is to be expected as *crtEBDC* expression is not CarQ dependent). The promoter of *crtEBDC* seems to share homology to those of several vegetatively expressed genes of *M. xanthus* (Figure 4-1), particularly , suggesting it may be under the control of the housekeeping sigma factor, SigA.

Martinez-Argudo *et al.* (1998) have performed a site-directed mutation approach to analyse the *crtI* promoter. They created specific mutations within the -10 and -35 regions and analysed the results *in vivo,* through β-galactosidase fusions and by colony colouration. However, if the true start site of *crtI* is displaced by 12bp as described above to bring P^{crtI} into line with P^{carQRS}, then the regions being mutated will actually be around the transcriptional start site and at the area around position -23. Due to the similarity of the displaced *crtI* promoter to that of *carQRS* around -35 conforming to the consensus for SigECF-dependent pomoters, a triple mutation was engineered that changed the GAC at position -45 to -43 to CCG. The triple mutation had no significant on P^{crtI} activity and this result does argue against the 12 bp displacement of P^{crtI}.

Point mutations generated between -33 and -29 almost abolished promoter activity. This is the region of similarity between P^{carQRS} and P^{crtI} as shown by Fontes *et al.* (1993). If the 12 bp displacement of P^{crtI} described above is correct then the mutated residues actually lie between -21 and -17. Mutation of the T at -9, the G at -10 or the C at -11 all greatly reduced promoter activity. Note once again that if the 12 bp displacemant of P^{crtI} is correct, the mutated residues actually lie from +1 to +3. Martinez-Argudo *et al.* (1998) also determined the minimum stretch of the *crtI* promoter required for activity. Full expression

required a stretch of DNA extending from somewhere between -54 and -25, to somewhere between +57 and +120 (the initiation codon of *crtI* lies at +61). The reason for the requirement of such a large amount of downstream DNA for promoter activity is far from clear.

carQRS	CGAGCGCCGGAAACACTTTCGCAGGTGGCCCGTAGAGGAGTCG
crtI (+12)	CCAGTGCCGGACGGTGCTCTTGTAACGTCCTGGCGGGTTCGCG

carQRS	CGAGCGCCGGAAACACTTTCGCAGGT GGCCCGTAGAGGAGTCG
crtI (+0)	GGTGCTCTTGTAACGTCCTGGCGGGTTCGCGCGTTCGCCAGGT

Figure 4-2 Homology between the carQRS *and* crtI *promoters.*

The two possible promoters of crtI *are shown, each alongside the* carQRS *promoter. Conserved bases are underlined.* crtI *(+12) denotes the promoter of* crtI *as described in Fontes* et al. *(1993) with the start site displaced 12 bases upstream. -10 and -35 nucleotides are in bold.*

To date, several areas of importance within the *carQRS* promoter have been identified. McGowan (1992) has shown that the minimum stretch of the *carQRS* promoter required for activity extends from between positions -136 and -145 (including the promoter for *gufA*). The *carQRS* promoter is thus very large and implies that additional factors are required for expression of *carQRS.*

Determination of the transcriptional start site of *carQRS* allowed assignment of the -35 and -10 recognition regions, an assignment which has since been strengthened by the similarity of these regions to those of other SigECF-dependent promoters. The discovery that *carQRS* expression is CarD-dependent (Nicolas *et al.* (1994) prompted Berry (1998) to search for possible regions of homology to the binding sites for HMGI(Y) proteins. He found a tandem repeat of TTTCC centred on positions -75 and -65 of the *carQRS* promoter. Subsequent work has shown that a CarD-GST fusion will bind to this area of the *carQRS*

(Cayuela and Murillo, 1998). Thus there are several regions of interest within the *carQRS* promoter: the CarD binding sites, the -35 and -10 recognition motifs and the *gufA* promoter.

4.2. Site-directed mutations within P^{carQRS}

Berry (1998) has constructed a set of sixteen site-directed mutations spread over the length of the *carQRS* promoter, with each mutant promoter fused upstream of a promoterless *lacZ* gene (Figure 4-3). Mutant sequences are designated Alt1 to Alt16 and were primarily engineered at sites such that the mutation introduced a *Kpn*I site into the sequence. The mutant promoter constructs also contain the P1 *inc* region and a kanamycin resistance determinant. Constructs containing the sixteen mutant promoters are designated pAEB601 to pAEB616 (see Figure 4-3) and are refered to as the pAEB6XX constructs, such that the Alt4 promoter is found on pAEB604. Construct pAEB600 carries the wild-type *carQRS* promoter. Introduction of these constructs into the Mx8 attB site of wild-type *M. xanthus* enables assessment of the activity of the promoter mutations *in vivo*.

4.2.1. Integration at the Mx8 AttB site.

Mx8 is a general transducing phage of *M. xanthus* which is capable of a lysogenic lifestyle and was first isolated as a lysogen (Martin *et al.*, 1978). Lysogeny arises by a specialised integration system encoded within the phage. The Mx8 genome contains an *attP* site which is a region of 29bp overlying an inverted repeat of 11bp (Tojo *et al.*, 1996). The *attP* sequence is identical to a site on the *M. xanthus* chromosome, the $attB^{Mx8}$ site, 3.5Mb from the *carQRS* locus (Chen *et al.*, 1991). Also found at the *attP* region of the phage genome is the *intP* gene which encodes an integrase protein and includes the *attP* site within the coding region of the gene. Recombination of the phage genome into the chromosome by homologous recombination across the *att* sites causes the *intP* gene to be truncated (denoted *intX*) and the gene product of the truncated gene (IntX) has been proven to be

defective in integrase/resolvase activity (Tojo *et al.*, 1996). Therefore integration into $attB^{Mx8}$ prevents excision events and the lysogen is stable.

It has been shown that plasmids carrying the Mx8 *attP* region, when introduced into *M. xanthus* integrate stably at the $attB^{Mx8}$ site (Orndorff *et al.*, 1983). The population of strains produced by integration are in the majority due to single insertions but a minority do integrate in multiple copies (Youderian, P. and Gill, R. pers. comm.). Another consideration is the presence of two $attB^{Mx8}$ hairpin sites in the *M. xanthus* chromosome at which recombination with *attP* can occur. Nothing is known about the relative frequency of insertions into each of the two sites or whether there is any difference in expression levels between the two sites.

Introduction of the mutant promoter *lacZ* fusions into the $attB^{Mx8}$ site has obvious advantages, with no disruption of any of the genes in the carotenogenic regulon, allowing activity of the mutant promoters to be assessed in a completely wild-type context. However, is has been shown that expression from within the $attB^{Mx8}$ locus can give different levels of expression of integrated genes than when expression of the same genes is assessed in their natural location in the *M. xanthus* chromosome (Gill, pers. comm.; McGowan, 1992).

4.2.2. Introduction of mutant promoter constructs into *M. xanthus*.

In this study, introduction of constructs containing the mutated *carQRS* promoters into DK101 was surprisingly difficult. P1-mediated transduction of plasmids containing substantial regions of *M. xanthus* DNA typically give rise to large quantities of transductants (~100 per plate) due to homologous recombination. Transduction of the pAEB6XX constructs which contain the mutated promoters (~150bp) and the Mx8 *attP* region (including *int* gene) gave transductants only rarely. This was also true when electroporation was attempted as the method for the introduction of plasmids.

With the presence of *attP* and the Mx8 *intP* gene on the introduced plasmid, integration should be favoured across the *att* sites over integration at the 160bp

carQRS promoter. With promoter probe plasmids such as pMAR206, the promoter fragment on the plasmid is much larger (~2.6kb) than that found on the pAEB6XX constructs, making homologous recombination into the chromosome much more probable for pMAR206, than for the AEB6XX plasmids. The small size of the promoter fragment in the pAEB6XX constructs means that should integration at the $attB^{Mx8}$ site be ineffective, the plasmid is unlikely to integrate into the chromosome at all. Therefore the inefficient integration of pAEB6XX plasmids into DK101 is most probably due to a defective copy of the *intP* gene on these plasmids.

4.2.3. Checking the pAEB6XX constructs.

Plasmids carrying the mutated promoters were checked by PCR using PCR primers 1 and 2 (see section 7.4.6.2.1). Primer 1 anneals to a site within P^{carQRS} while primer 2 anneals in the opposite orientation within the *lacZ* gene. Products should only be obtained by PCR if the plasmid carries both *lacZ* and P^{carQRS} in the correct orientation. The 480bp product obtained from the PCR reaction was then incubated with restriction enzyme *Kpn*I. Since the mutated promoters each contained a novel *Kpn*I site due to the presence of the Alt mutation, cleavage of the PCR product confirmed that the copy of P^{carQRS} within the plasmid still contained the engineered mutation. The integrated plasmids were also checked by PCR to confirm their presence within the chromosome of the recipient strain using primers 1 and 2 and *Kpn*I digestion as above. All strains were screened in this fashion before β-galactosidase assays were performed (see Figure 4-4 for example of the PCR/*Kpn*I screen).

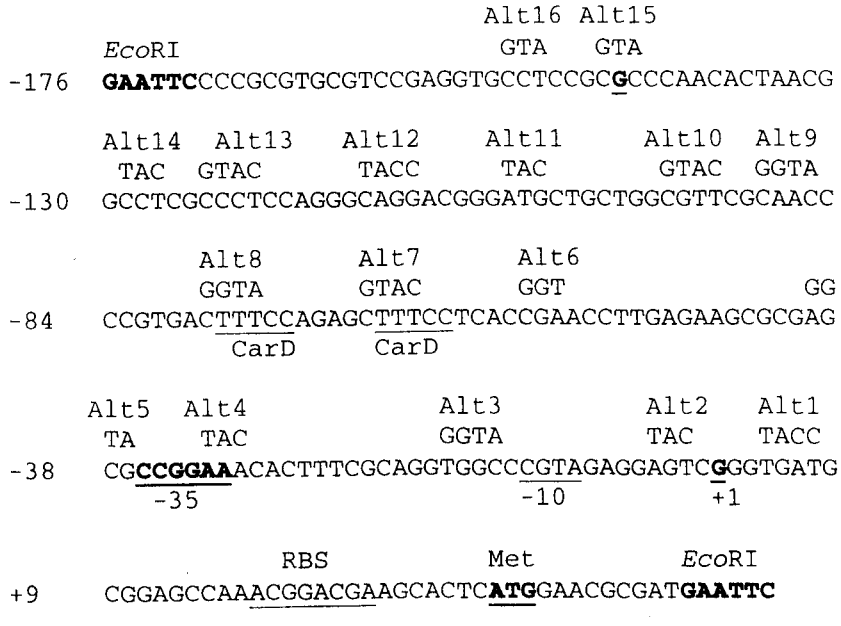


Figure 4-3 Site-Directed mutations within the carQRS *promoter.*

The DNA sequence shown is the Eco*RI fragment of pAEB120 containing the* carQRS *promoter sequence. Sequence is of plasmid origin up to position -164. Changes in the sequence engineered into the promoter by site-directed mutagenesis are shown above the wild-type sequence with the label for the plasmid carrying the mutant promoter. All the mutations introduce a novel* Kpn*I restriction site (GGTACC). Plasmid names correspond to the plasmid which carries that particular mutant promoter.For example, pAEB602 carries the mutant promoter Alt2 which has the sequence GGTACCG instead of the wild-type sequence GGAGTCG approaching position +1. The base underlined at -144 is the transcriptional start site of the* gufA *gene. The proposed CarD binding site is also underlined centered at positions -74 and -64.*

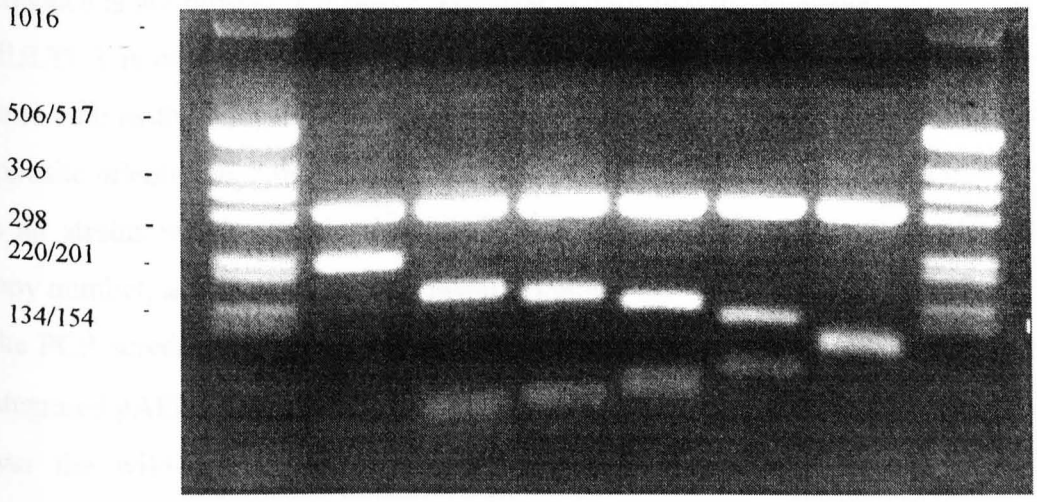


Figure 4-4 Screening for the presence of site-directed mutations within P^{carQRS}.

Lane 1: Molecular Markers. Lanes 2-7: PCR products derived from DK101::pAEB600, DK101::pAEB601, DK101::pAEB602, DK101::pAEB603, DK101::pAEB604 and DK101::pAEB606 using primers 1 and 2, restricted with KpnI. *The presence of a* KpnI *site between the* carQRS *promoter and the* lacZ *gene gives a 300bp and a 220bp fragment on restriction of PCR product with* KpnI. *Presence of an engineered Kpn*I *site within* P^{carQRS} *eliminates the 220bp fragment and gives rise to two fragments which together comprise the full 220bp.*

The possibility that the more than one copy of the plasmid had integrated to form multiple insertions was also checked by PCR. Primers MULTI 1 and MULTI 2 are complementary to regions flanking the Mx8 *attP* site within the Mx8 *attP* region. Amplification between these primers should only occur in strains containing an intact copy of the *attP* site, i.e. only in the case of an *E. coli* strain carrying one of the pAEB6XX plasmids and in *M. xanthus* strains which contain tandem multiple insertions of the pAEB6XX plasmid in the $attB^{Mx8}$ site. Since absence of an amplification product does not mean that the target template sequence is not present, a positive control was built into the PCR reaction. Primer MULTI 3 is complementary to a site within the Mx8 *attP* region which is the same side as the annealing site of MULTI 1 relative to the *attP* site, and in the opposite orientation. Amplification between MULTI 1 and MULTI 3 should occur in all strains which contain integrated pAEB6XX plasmids, regardless of their copy number, acting as an internal positive control for the PCR reaction.

The PCR screen for multiple insertions were performed on all strains containing integrated pAEB6XX plasmids that gave greater levels of β-galactosidase activity than the wild-type construct DK101::pAEB600 (i.e. DK101::pAEB606 and DK101::pAEB607). The PCR reactions contained primers MULTI 1, MULTI 2 and MULTI 3 in the ratio 2:1:1(see section 7.4.6.2.2). Amplification between primers MULTI 1 and MULTI 2 gave the expected product of 694bp. This product was only seen when PCRs were performed on unintegrated plasmids carrying the intact AttP region. PCR reactions performed on DK101 containing integrated pAEB606 and pAEB607 and reactions performed directly on the unintegrated plasmids, all gave the expected product of 212bp due to amplification between primers 1 and 3 (Figure 4-6). This indicated that both DK101::pAEB606 and DK101::pAEB607 carried integrated plasmids in single copy number and not as tandem multiple insertions.

PCR would ideally have also been used to characterise the nature of the event which integrated the plasmids into the chromosome, ie. into which of the two *att* sites the plasmids integrated and in which orientation. Without this control, the orientation of the Alt promoters and their chromosomal context remain variables which are unaccounted for when comparing Alt promoter activities.

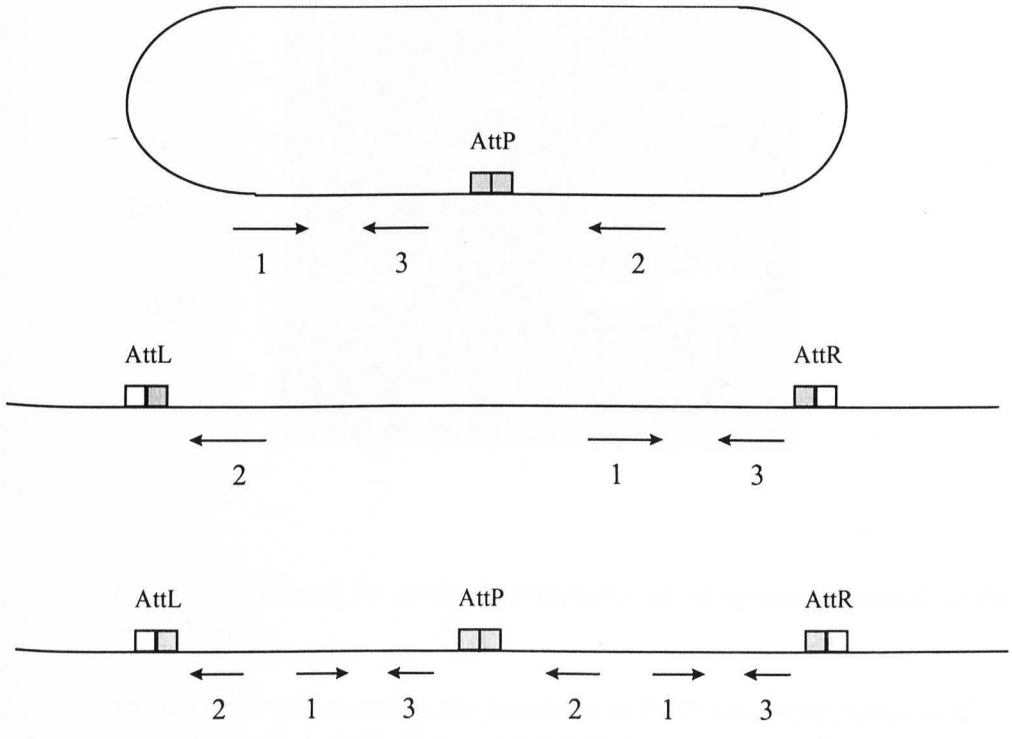


Figure 4-5 The attP *region and annealing sites for primers (MULTI) 1, 2 and 3.*

The products expected from PCR reactions involving primers (MULTI) 1, 2 and 3 differ between (top) plasmid containing attP*, (middle) single integrated plasmid into* $attB^{Mx8}$ *and (bottom) multiply integrated plasmids at* $attB^{Mx8}$*. Filled boxes represent* attP *site-derived sequences and open boxes derive from* $attB^{Mx8}$*. Not to scale. All three scenarios give a product between primers 1 and 3. Product is only formed between primers 1 and 2 in the case of plasmid carrying* attP *top) and in the case of a multiply inserted plasmid (bottom).*

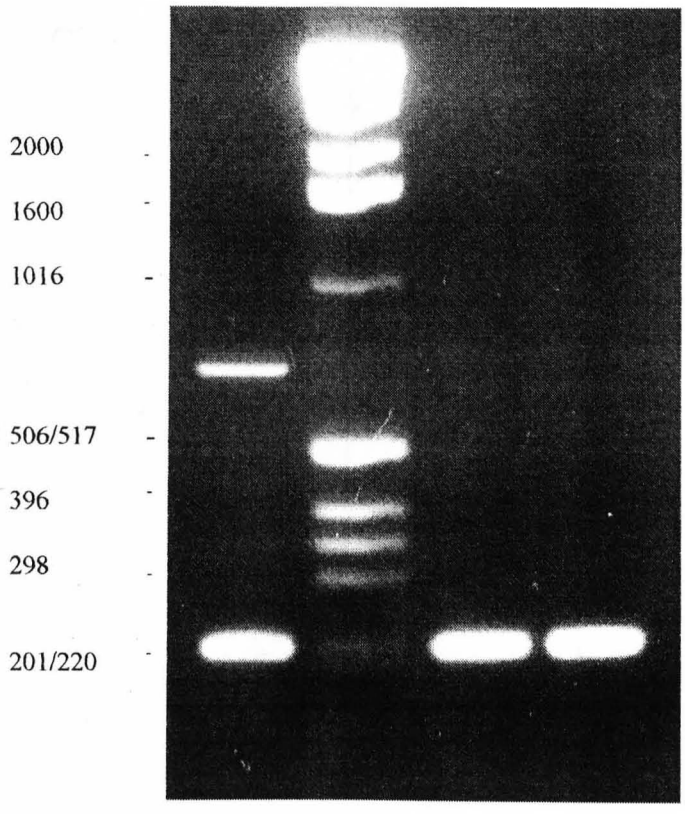


Figure 4-6 Check for multiple insertions of integrated plasmids at the attB^{Mx8} *site.*

1% agarose gel showing the products of PCR reactions performed using MULTI 1, MULTI 2 and MULTI 3 on various DNA templates. Lane 1: pAEB606 (positive control). Lane 2: Markers. Lane 3: DK101::pAEB606. Lane 4: DK101::pAEB607.

4.3. *In vivo* activity of mutant promoters.

Plasmid pAEB600 carries the wild-type *carQRS* promoter. The plasmid was introduced into the *attB*^{Mx8} site of DK101 and the resulting strain's production of β-galactosidase was assayed in both light and dark (Figure 4-7). As before, β-galactosidase assays were performed in at least duplicate and only consistent results are shown.

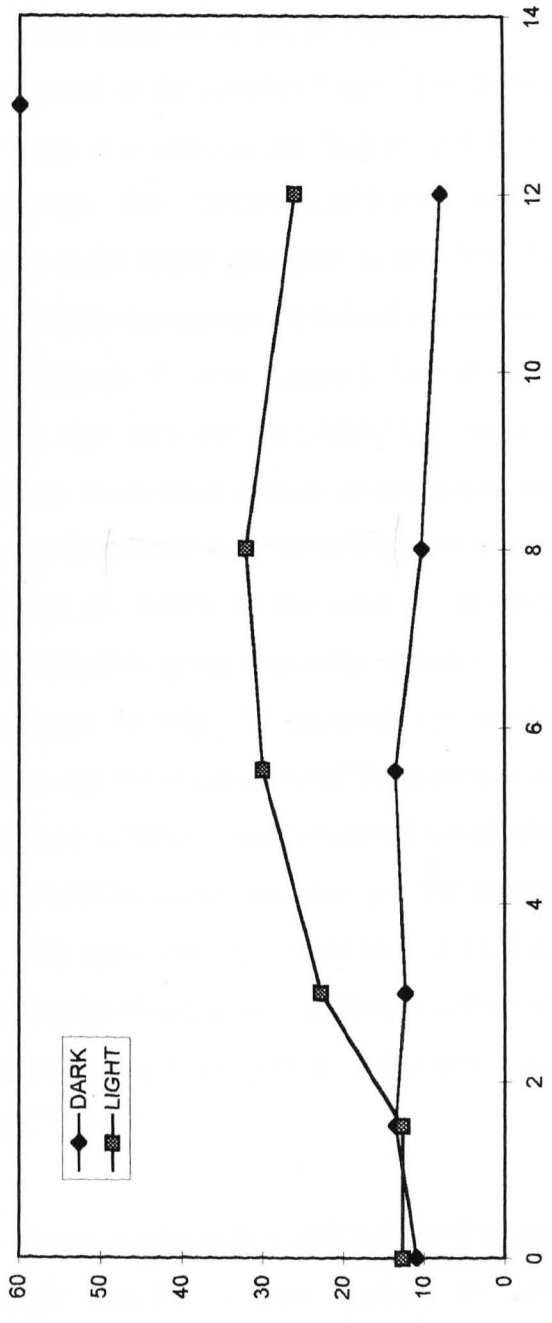


Figure 4-7 Activity of the carQRS *promoter at the AttB site (DK101::pAEB600).*

Y-axis is specific activity of β-galactosidase in units/min/mg protein. X-axis is time in hours.

If the promoter activity is compared to the situation in DK101::pDAH217 where the identical *carQRS* promoter is integrated at the *carQRS* locus, then there are two immediate observations. Firstly there is a very similar background level of expression (~10 units/min/mg protein for DK101::pAEB600 and for DK101::pDAH217). However, in most cases where promoter probes have been introduced into the $attB^{Mx8}$ site, very variable background levels of expression are observed (Fisseha *et al.,* 1996 and Youderian, P., pers. comm.). As will be seen later, this variable background level is also seen for the pAEB6XX constructs. Another observation that has been made from other studies of promoters at the $attB^{Mx8}$ site is that induction can be much reduced relative to the same promoter when at its wild-type locus (Fisseha *et al.,* 1996). In the case of the *carQRS* promoter, when at the *carQRS* locus, induction gives a maximal activity of ~400 units/min/mg protein whereas a maximum of only ~35 units/min/mg protein is seen when integrated at $attB^{Mx8}$. Despite the lower induction of the promoter when in the $attB^{Mx8}$ site, a significant induction (250%) is still observed which should prove sufficient to assess any loss-of-activity mutations amongst the promoters within the pAEB6XX constructs. The approximately three-fold induction on illumination corresponds well to the results obtained by McGowan (1992) which suggested that the *carQRS* promoter extending from -145 would exhibit roughly three-to-four fold induction at the $attB^{Mx8}$ site.

For ease of reference the mutations created in plasmids pAEB6XX will be refered to as AltX. For instance the mutant *carQRS* promoter carried on plasmid pAEB601 will be refered to as promoter Alt1.

4.3.1. Alt1 exhibits no light-induction.

Alt1 carries a mutation overlying positions +3 to +6 bases relative to the transcriptional start site of *carQRS*. When at the $attB^{Mx8}$ site, the promoter is not induced by light (Figure 4-8).

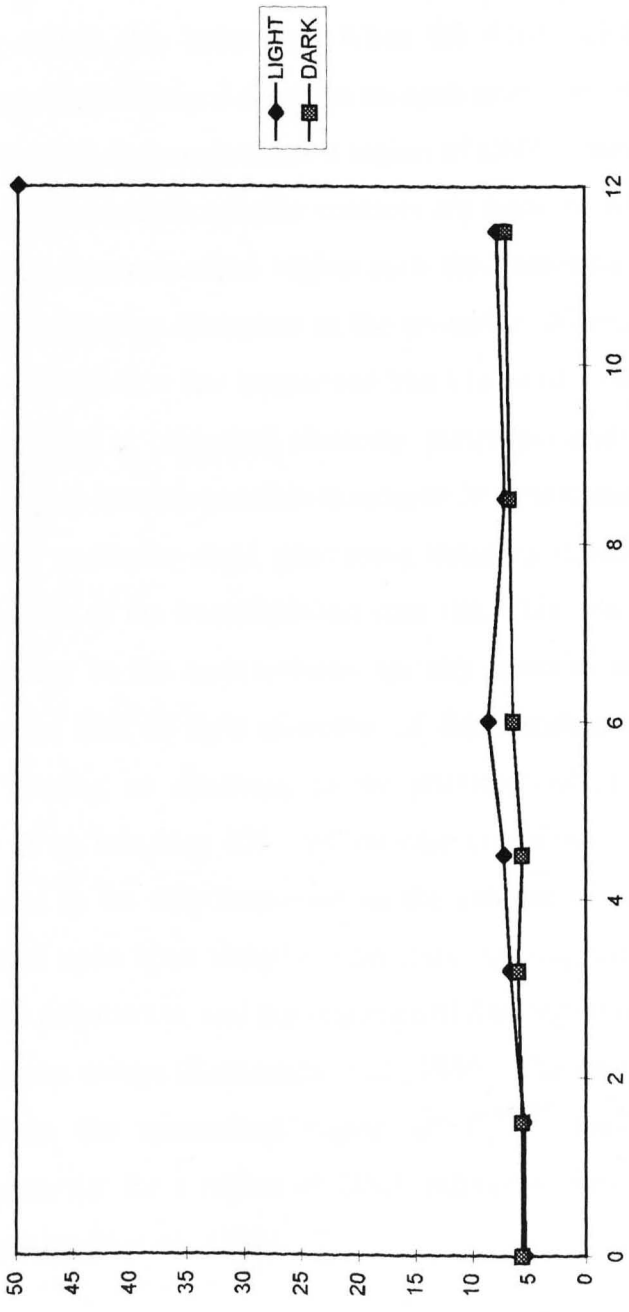


Figure 4-8 Activity of the Alt1 promoter (DK101::pAEB601).

Y-axis is specific activity of β-galactosidase in units/min/mg protein. X-axis is time in hours.

It is difficult to reconcile loss of light-induction with a mutation which is situated outside the promoter, within the transcribed region itself. However, there are possibilities which might explain this behaviour. When the RNA polymerase holoenzyme undergoes isomerisation from a closed to an open promoter complex, footprinting studies show that the enzyme protects a region of DNA extending to position +20. It is possible that sequence specific contacts are made between the enzyme and the DNA within the transcribed region such that mutation of this region prevents correct open complex formation at the promoter. Alternatively, there may be a requirement for the first few transcribed bases to be of a particular nature-either of specific sequence or conserved chemisty, purine/pyrimidine etc.. Buttner *et al.* (1987) have shown that it is possible to achieve *in vitro* transcription from the promoters of the *S. coelicolor dagA* gene using initiating dinucleotides that do not match the sequence of the transcriptional start site. This implies that there is considerable leniency in the requirements for any specific initiating nucleotides and therefore the lack of light-induction of Alt1 implies that the mutation is most likely causing an alteration in the promoter which affects formation and/or activity of an initiating RNA polymerase complex. The early transcribed region is known to be very important in the process of promoter clearance. It must be melted upon open complex formation, causing the loss of stabilising contacts with the polymerase, and the sequence of this region therefore modulates the rate of promoter escape (Kammerer *et al.*, 1986). The involvement of specific residues within the transcribed region of P^{carQRS} mirrors the requirement of the *crtI* promoter for a region of DNA extending well into the transcribed area (Martinez-Argudo *et al.*, 1998).

4.3.2. Alt2 shows wild-type induction.

The Alt2 promoter exhibits no fundamental differences from the wild-type *carQRS* promoter when integrated into $attB^{Mx8}$ (Figure 4-9).

The mutation carried by Alt2 lies centred at position -3. As such it lies outside of the promoter elements believed to have a role in promoter activity, except that it lies within the region of DNA that is melted during open complex formation. The

site-directed mutation causes AGT to be replaced with TAC and thus the GC content of the region is maintained. Thus any effect of the mutation on promoter melting is likely to be minimal and the lack of phenotype associated with the mutation is therefore not unexpected.

As with all the mutant Alt promoters which exhibit activity, it is possible that the activity (whether background or light-induced) is due to the mutant promoter being recognised by a sigma factor other than CarQ. For all the Alt promoters CarQ-dependence should be assessed by introduction into UWM303 (a *carQ* knock-out strain) and assaying for activity.

4.3.3. Alt3 is constitutively active in *M. xanthus*.

The mutation within Alt3 lies at position -14 to -17. This region is similar between the *carQRS* promoter and the 12bp-displaced *crtI* promoter (Figure 4-2) and may thus define the region recognised by region 2.4 of CarQ. Even though this region does not conform to the position expected to hold the -10 hexamer, it is possible that this is the region required for CarQ recognition. CarQ is an ECF sigma factor and as such carries a significant deletion relative to Sig70 between regions 2.4 and 4.2. It is possible that this deletion causes the recognition elements within SigECF-dependent promoters to be more closely spaced than those within Sig70-dependent promoters, and therefore the '-10' hexamer may not lie at -10 within the *carQRS* promoter. Whether this is the case or not is far from certain, as the poor sequence identity between the -10 elements of SigECF-dependent promoters has precluded their unambiguous assignment. Alternatively, the positions mutated in Alt3 may lie outside of the -10 recognition hexamer and would be postulated to have a minimal effect on promoter activity.

The Alt3 promoter was transduced into DK101 and the activity of the promoter assayed. Alt3 was found to be constitutively and highly active in DK101 (Figure 4-10).

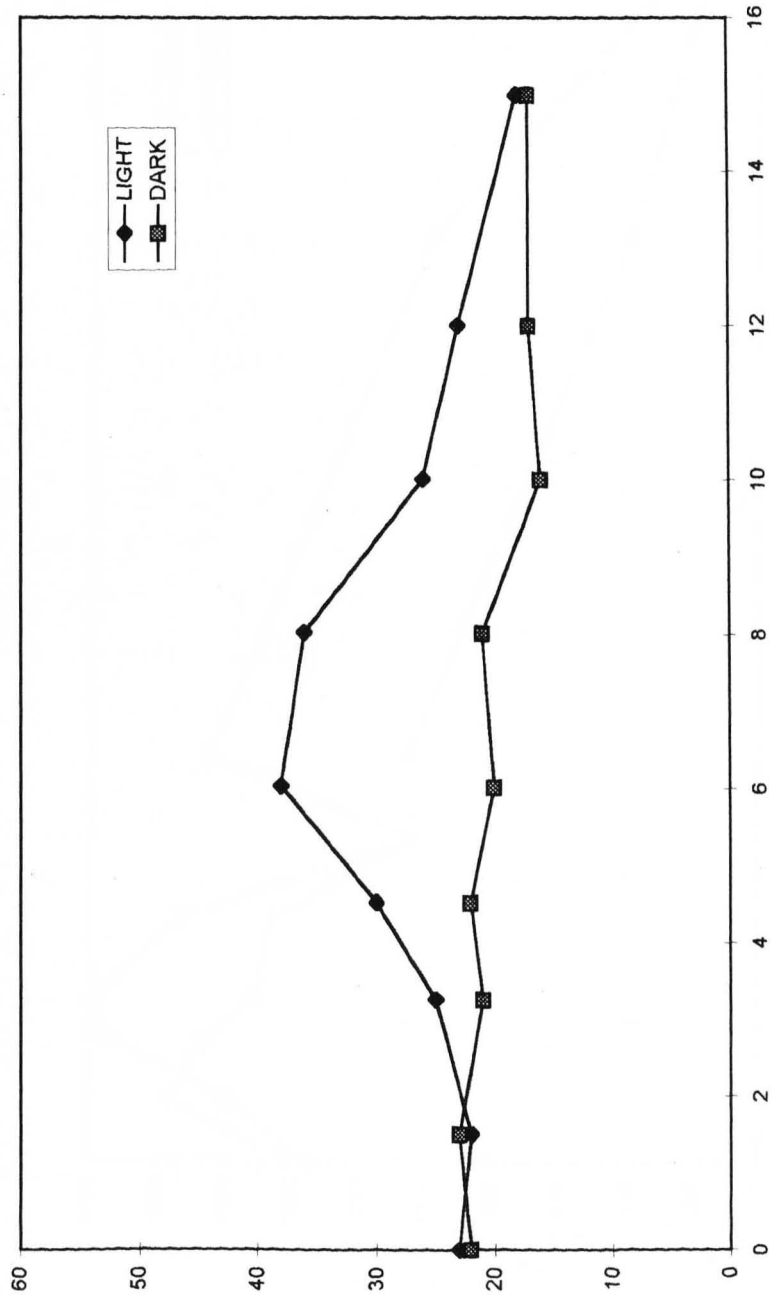


Figure 4-9 Activity of the Alt2 promoter (DK101::pAEB602).

Y-axis is specific activity of β-galactosidase in units/min/mg protein. X-axis is time in hours

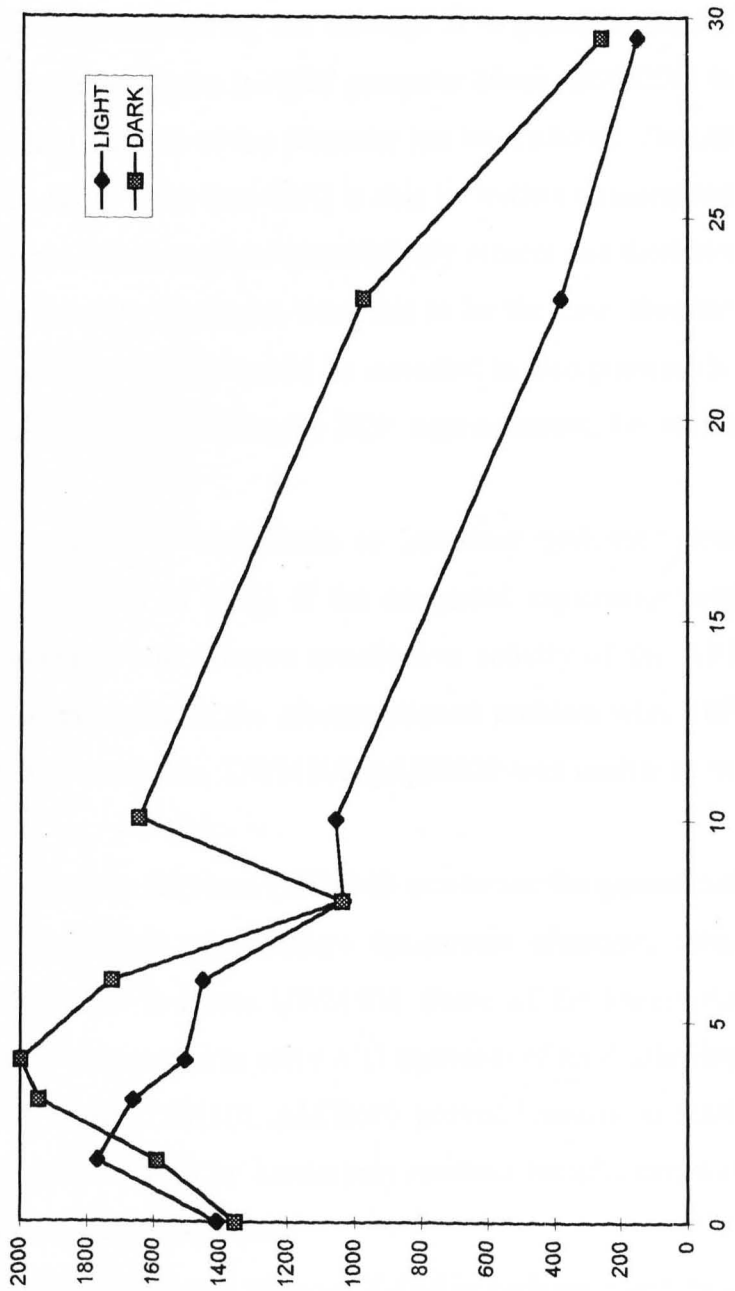


Figure 4-10 Activity of the Alt3 promoter (DK101::pAEB603).

Y-axis is specific activity of β-galactosidase in units/min/mg protein. X-axis is time in hours

There are several possible explanations for his activity. It is possible that by mutating the recognition sequence of the *carQRS* promoter (from GGCCCG to GTACCG), the sigma factor specificity of the promoter has been altered. Thus in the *M. xanthus* cell, a sigma factor other than CarQ is able to initiate transcription from the promoter. This sigma factor could be constitutively present and therefore activate the promoter constitutively. However, were this to be the case, then the sigma factor causing transcription of Alt3 would be expected to also possess the altered positional requirement postulated for the ECF sigma factors, i.e. be an ECF sigma factor itself.

To check this hypothesis, attempts were made to introduce pAEB603 into UWM303 which carries a deletion of *carQ*. If the suggested explanation was correct, it would be expected to still observe constitutive activity of the Alt3 promoter in this strain. However, due to the aforementioned problem with AttP integration of the pAEB6XX constructs, UWM303::pAEB603 was unable to be constructed.

An alternative strategy to introduce Alt3 into UWM303 was to use the generalised transducing phages Mx4 and Mx8 to transduce kanamycin resistance from DK101::pAEB603 into UWM303 to create UWM504. Some of the kanamycin resistant transductants would be expected to carry Alt3 upstream of *lacZ*, allowing assessment of Alt3 activity *in vivo*. DK101::pAEB603 proved resistant to Mx4, but was sensitive to Mx8 and a single Car$^-$ kanamycin resistant transductant was obtained through transduction mediated by Mx8.

UWM504 was grown on plates containing 20μg/ml X-Gal in both light and dark, and in both conditions colonies developed intense blue colouration. PCR reactions were performed using primers 1 and 2, confirming that UWM504 contained the Alt3 promoter upstream of *lacZ*. Time constraints prevented the determination of a time-course for β-galactosidase expression, however the hydrolysis of X-Gal in both light and dark-grown colonies of UWM504 implies that the Alt3 promoter is constitutively active in this *carQ* strain. It seems therefore that the initial hypothesis that Alt3 is expressed constitutively by a non-CarQ sigma factor of *M. xanthus* is correct.

An alternative possible explanation for the constitutive activity of Alt3 lies in the notion of an 'extended -10 promoter'. In some Sig70 promoters of *E. coli* there lies a conserved TG dinucleotide at -15 to -14, forming an extended -10 region sequence of TGnTATAAT. The presence of the extended -10 motif removes the requirement for a cognate -35 region and reduces the specificity required by the sigma factor for the promoter at the -10 region (Bown *et al.*, 1997 and Kumar *et al.*, 1993). It is possible that the mutation found in Alt3 has introduced a 'psuedo-extended -10' motif of TAcCGTAGA. Barne *et al.* (1997) changed the TG doublet by site-directed mutagenesis to TC and TT and in each case found that loss of promoter activity was compensated for by a mutation of Sig70 at position 458, within region 2.5.

CarQ may, or may not, have a region 2.5. Members of the ECF sigma factors have a significant deletion between regions 2.4 and 4.2 relative to all other Sig70-family sigma factors (Lonetto *et al.*, 1994). However, the vestigial region between 2.4 and 4.2 of CarQ had very little homology to any corresponding region of Sig70 and could thus actually form an entire region 2.5, or entire region 3.1, or may be interspersed randomly between the regions. It is possible that CarQ does possess an entire region 2.5 and may thus bind to an 'extended -10'-like promoter such as Alt3, however, CarQ is not constitutively available to cause transcription of *carQRS* promoters in DK101. If Alt3 is being transcribed by a non-CarQ sigma factor, it may be acting as an 'extended -10' promoter and thus the sigma factor which is causing its transcription is able to initiate transcription without requiring -35 and -10 regions similar to those of its cognate promoters.

4.3.4. Alt4 exhibits no light-induction.

The mutation within promoter Alt4 results in no promoter light-induction (Figure 4-11) although it seems to cause a general light-independent increase in the level of expression. It is possible that the lack of induction upon illumination is merely due to a masking of induction due to the higher basal promoter activity.

The mutated bases present in promoter Alt4 slightly overlap the 3' end of the -35 hexamer. That is, the wild-type sequence CGCCGGAAAC is altered to CGCCGGTACC. It is possible that by changing the -35 hexamer the promoter may no longer be recognised by CarQ. It is possible that the increased background level of transcription is due to Alt4 being recognised by a non-CarQ sigma factor, either by changing the sigma factor specificity of the extant promoter or by fortuitously creating a novel promoter. Determination of the transcriptional start site of Alt4 would allow elucidation of which possibility is the correct explanation. This situation is also true for the background activities of the other Alt promoters.

Berry (1998) has suggested using predictive programs that the *carQRS* promoter carries intrinsic curvature of around 16^0 at the CarD binding site due to the specific sequence of its DNA. The *Kpn*I sites introduced into the *carQRS* promoter by site-directed mutagenesis possess an intrinsic curvature of around 6^0 and this slight kink is able to influence the curvature of its surrounding region depending on the site of insertion. Thus in Alt1, Alt2 and Alt3, the *Kpn*I site causes no significant change in the curvature of the promoter. In Alt4 however, the *Kpn*I kink is positioned such that it is predicted to cause the introduction of a 13^0 bend in the DNA between positions -16 to -40. DNA bending is of considerable importance in transcription initiation (Perez-Martin and Espinosa, 1994). The process of promoter melting during conversion of open to closed complexes is enhanced by negative supercoiling and an over rotation of the DNA between -10 and -35 (deHaseth and Helmann, 1995). Bending and supercoiling of DNA are inter-related as bends cause a change in the degree of coiling of the DNA at the bend (Yang *et al.*, 1995) and supercoiling can ease or hinder the

melting of DNA, which can be viewed simplistically as a change in the bend of the DNA strands within the double helix. Thus as bending and supercoiling affect the ease of promoter melting it is conceivable that Alt4 is not light-inducible as a consequence of the introduced bend of its DNA betwen -35 and -10 hindering the melting of the downstream DNA during open complex formation. Alternatively, the bend between -10 and -35 may cause a reduced affinity of RNA polymerase holenzyme for the promoter, with CarQ being less able to make specific contacts simultaneously to both -10 and -35.

The observation of increased transcription from Alt4 in the dark may be a consequence of integration of pAEB604 into a different *attB*^{Mx8} site, or in an opposite orientation within the *attB*^{Mx8} site, from that of DK101::pAEB600. Not enough is known about the difference in transcriptional activity of promoters at the different *attB*^{Mx8} sites and in the alternative possible orientations, to be able to make any conclusions about this aspect of the activity of Alt4.

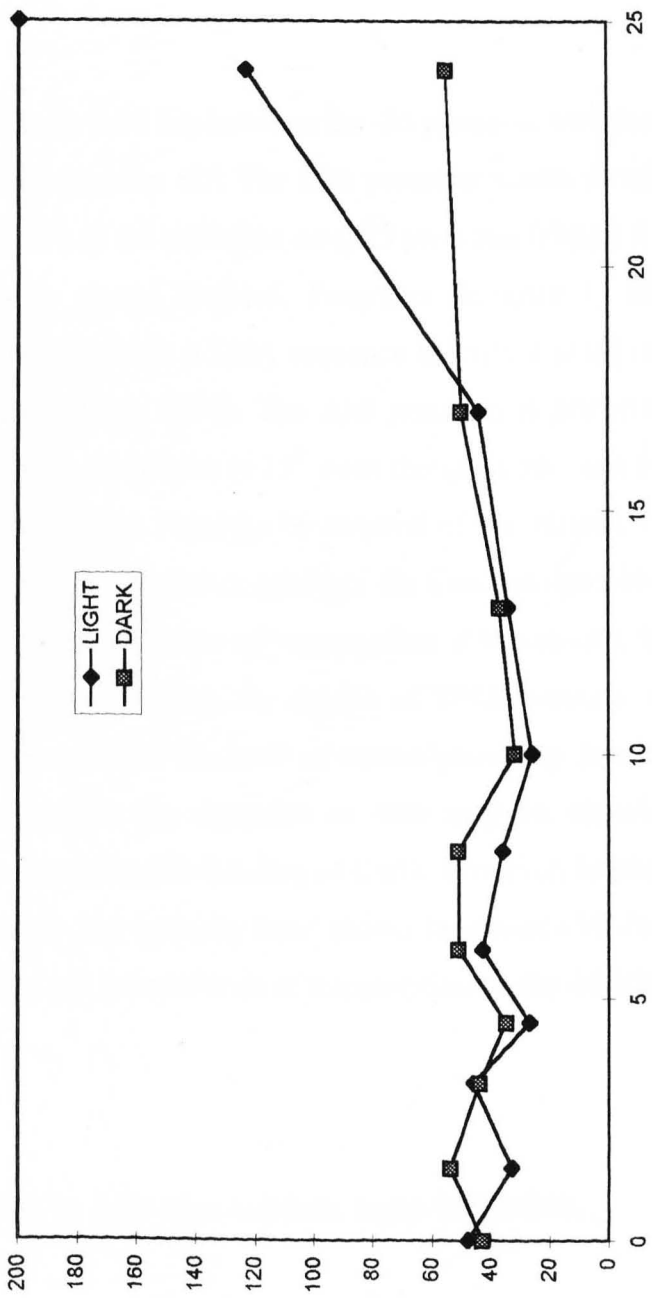


Figure 4-11 Activity of the Alt4 promoter (DK101::pAEB604).

Y-axis is specific activity of β-galactosidase in units/min/mg protein. X-axis is time in hours

4.3.5. The Alt6 promoter shows light-induction.

The mutation introduced to create Alt6 lies between the -35 promoter element and the CarD binding sites, around position -57. The Alt6 promoter seems to exhibit enhanced light-induction relative to the wild-type *carQRS* promoter (Figure 4-12), with the degree of induction almost doubled. Programs designed to model probable sequence specific bends within a DNA sequence identify a putative 16^0 kink at the CarD binding site (Berry, 1998). The Alt6 mutation is predicted to reduce the severity of the bend in this region to 13^0, even though it lies outside the area containing the CarD binding site. Possibly, by removal of the intrinsic DNA bend, the mutated promoter serves as a better substrate for CarD recognition and therefore causes an increase in the initiation of transcription. Alternatively, CarD may function by increasing or decreasing the degree of DNA bending at its binding site, which serves to modulate the level of transcription. By decreasing bending at the CarD binding site, the mutation in Alt6 may be stimulating transcription by intrinsially simulating the binding of CarD. However, bending of DNA upstream of the promoter has typically been shown to enhance binding of RNA polymerase and increase not reduce levels of transcription (Perez-Martin and Espinosa, 1994).

4.3.6. The mutation in Alt7 also exhibits light-induction.

The mutation in Alt7 overlies the CarD binding site at -65 and enhances light induction almost two-fold (Figure 4-13). It seems likely that the enhancement of induction is a consequence of alteration of the CarD binding site. As with Alt6, the mutation is predicted to cause a reduction in the predicted bending of the CarD binding site, and this may enhance promoter activity due to greater binding by CarD or by simulation of the action of CarD. In the case of Alt7, the predicted bend at the CarD binding region is predicted to be almost completely abolished.

The CarD binding site of P^{carQRS} is believed to be a tandem repeat of the sequence TTTCC (Berry, 1998). AT-rich regions have been shown to introduce intrinsic curvature into DNA and have also been shown to increase transcription when upstream of the RNA polymerase binding region (Rao *et al.*, 1994). The α sub-units of RNA polymerase are believed to bind to these AT-rich regions and pomote transcription (Ross *et al.*, 1993). They function by recognising either an AT-rich sequence, or the curvature induced by the AT-rich region.

By decreasing the predicted bend of the CarD binding site, it is possible that the Alt7 mutation provides a CarD binding site which has a greater affinity for CarD than does the wild-type promoter. This possibility could explain the increased activity of Alt7, if CarD binding to a wild-type promoter causes straightening of the promoter and allows CarD-RNA polymerase contacts to form, stabilising closed complex formation. However, it is normally the case that DNA binding proteins induce curvature within their target DNA. While true that many will bind to pieces of DNA that already exhibit curvature, they do not usually then cause removal of that curvature.

As for Alt6, it should be pointed out that whereas for Alt7, where a loss of DNA bending is associated with increased promoter activity, in general, bending of DNA upstream of the promoter typically enhances binding of RNA polymerase and increases not reduces levels of transcription (Perez-Martin and Espinosa, 1994).

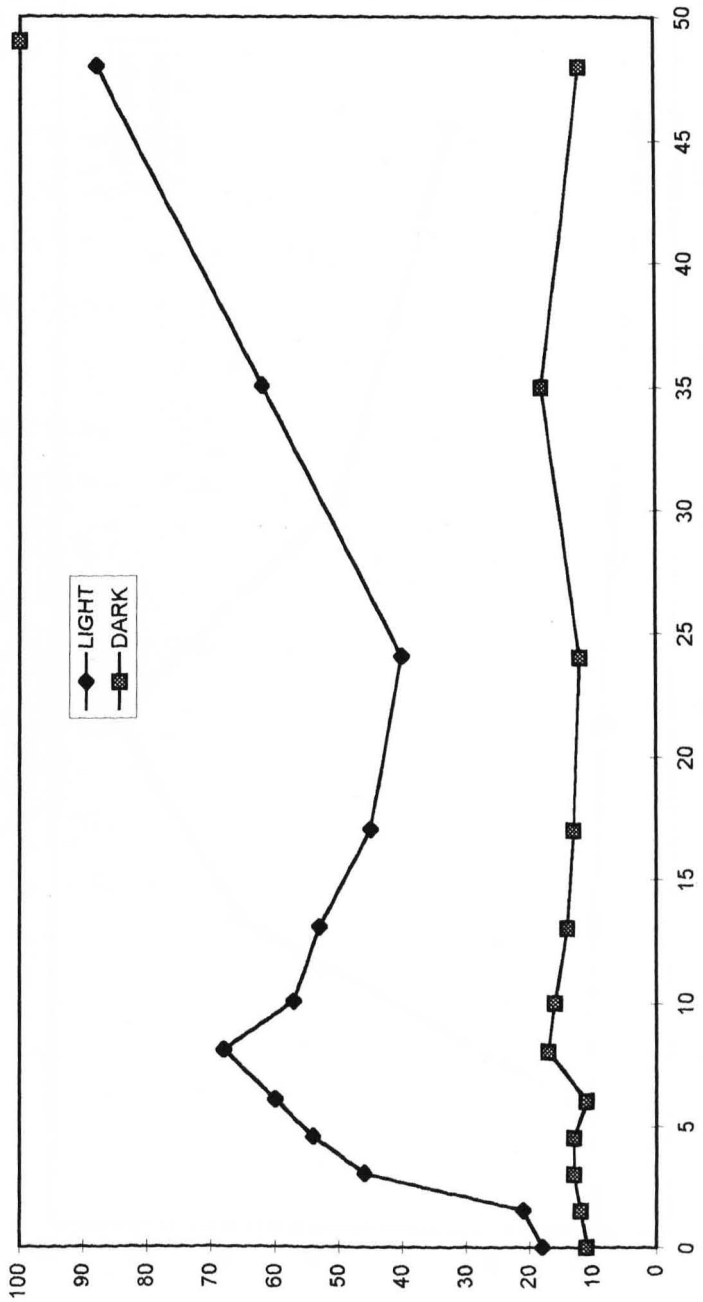


Figure 4-12 Activity of the Alt6 promoter (DK101::pAEB606).

Y-axis is specific activity of β-galactosidase in units/min/mg protein. X-axis is time in hours

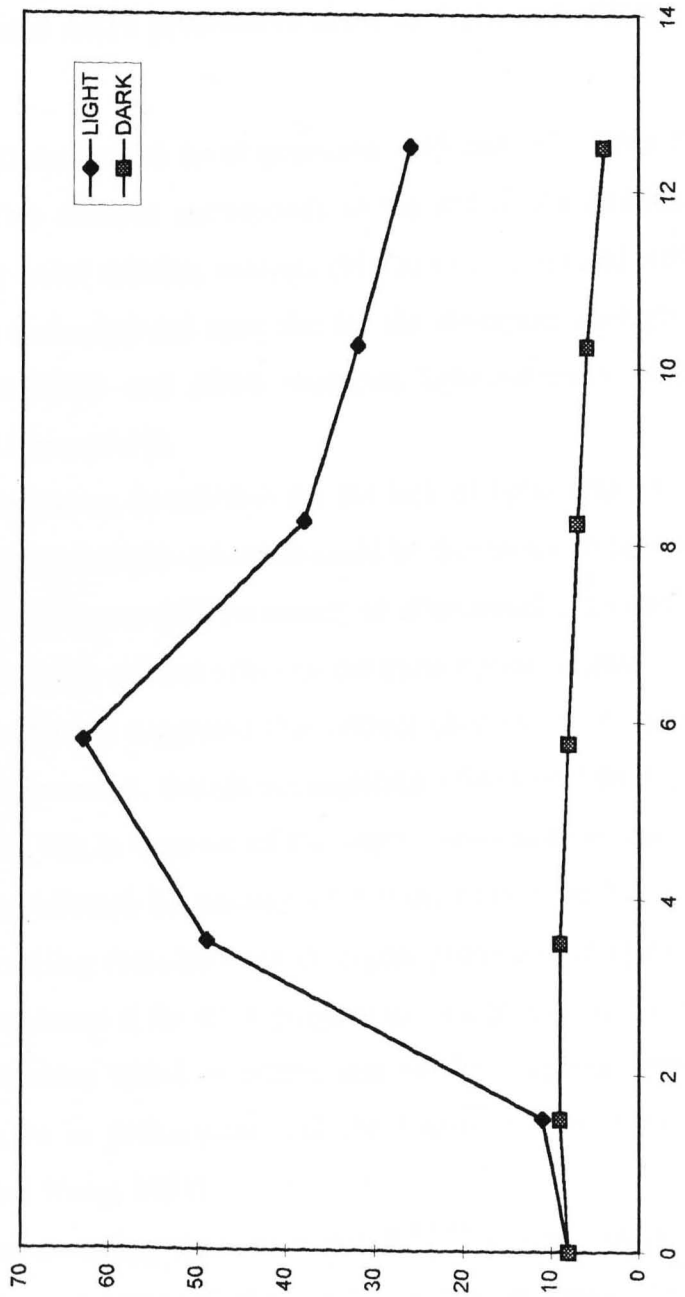


Figure 4-13 Activity of the Alt7 promoter (DK101::pAEB607).

Y-axis is specific activity of β-galactosidase in units/min/mg protein. X-axis is time in hours

4.3.7. The Alt15 and Alt16 promoters show no light-induction.

The mutations within Alt15 and Alt16 lie at positions -145 and -150 from the transcriptional start site. This distance corresponds to the end of the minimum promoter as determined by serial deletion analysis (McGowan, 1992) and Alt15 overlies the region of the transcriptional start site for the divergent *gufA* gene. Each of the mutations in Alt15 and Alt16 abolishes light-induction of the promoter (Figure 4-14 and Figure 4-15).

There are two main, non-exclusive, possiblities for the lack of light-induction of Alt15 and Alt16. The reduction in light-induction could be due to an alteration in an essential cis-acting part of the *carQRS* promoter, or alternatively, it could be due to an indirect effect caused by a direct effect on the transcription of *gufA*.

Initial work by McGowan (1992) suggested that correct expression of *carQRS* required a functioning *gufA* promoter, though not requiring a functional *gufA* gene product. It is probable that this is because of the supercoiling state of the two promoters, which is being affected by passage of a transcription bubble. High degrees of negative supercoiling form between divergent promoters (Mojica and Higgins (1996) which is enhanced if the RNA polymerase complex is anchored to the cell membrane. A situation which is likely, due to the coupling between transcription and translation in prokaryotes and the transmembrane nature of GufA and CarR (Lynch and Wang, 1993).

The evidence that there are cis-acting elements within P^{carQRS} around -160 is non-existent, however there is no evidence to the contrary either. Whether either or both of the explanations discussed above are correct will require a method for seperating the *gufA* and *carQRS* promoters - a far from trivial, if not impossible, exercise.

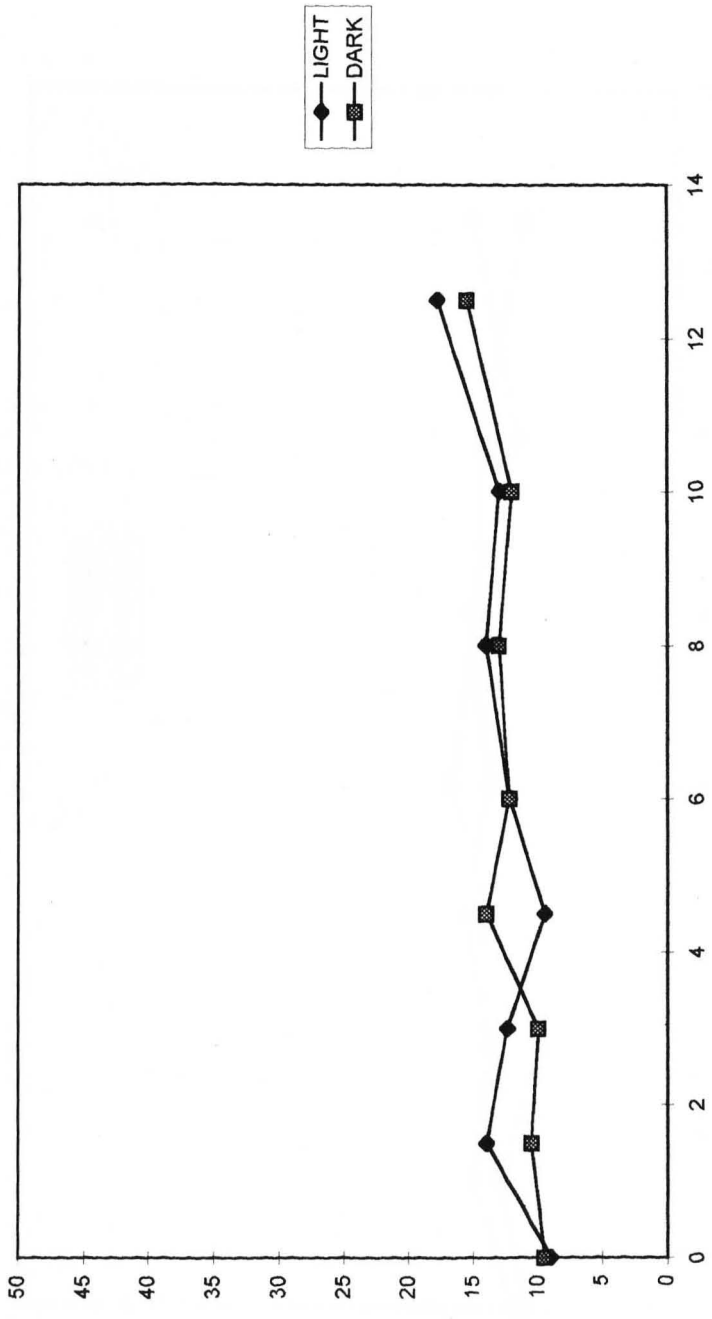


Figure 4-14 Activity of the Alt15 promoter (DK101::pAEB615).

Y-axis is specific activity of β-galactosidase in units/min/mg protein. X-axis is time in hours.

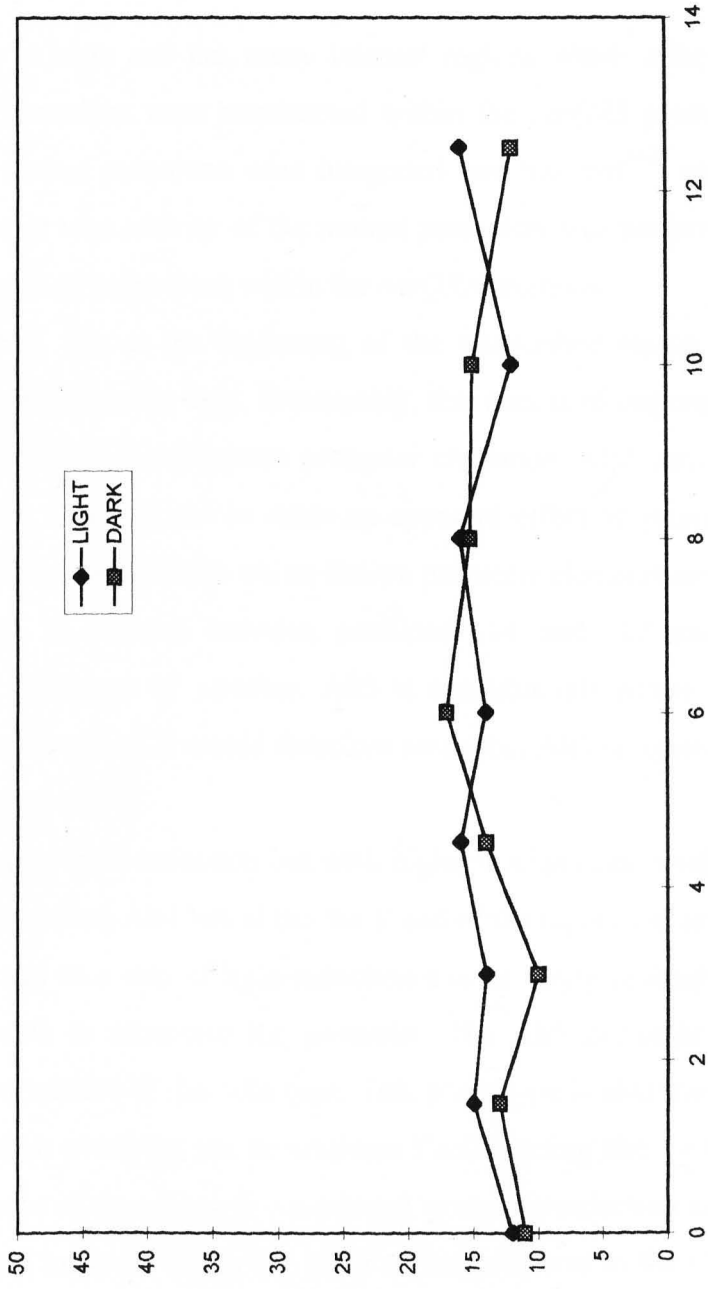


Figure 4-15 Activity of the Alt16 promoter (DK101::pAEB616).

Y-axis is specific activity of β-galactosidase in units/min/mg protein. X-axis is time in hours

4.4. Conclusions.

The *carQRS* promoter is large and has many internal regions which affect its function. Site-directed mutations were constructed within the *carQRS* promoter (Berry, 1998) and the mutant promoters were integrated into the $attB^{Mx8}$ site of DK101. Analysis of the *in vivo* activity of the mutant promoters was performed, allowing detection of areas of importance within the *carQRS* promoter.

The mutation within Alt1 lies at the beginning of the transcribed region and renders the promoter uninducible by light. Presumably, this area is of importance in initiation of transcription and subsequent promoter clearance. Alt2 carries a mutation around position -3 and seems to cause no apparent effect on promoter activity. This was to be expected as there are no known promoter elements around -3. The Alt3 promoter is mutated between positions -14 and -17 and is constitutively active in wild-type *M. xanthus*. Alt3 is constitutively active in a mutant strain deleted for *carQ* and it would therefore seem that Alt3 is expressed by a sigma factor other than CarQ.

Promoter Alt4 exhibited no light-induction but with higher background levels of expression. The mutation within Alt4 lies at the the 3' end of the highly conserved -35 promoter hexamer and so a loss of light-induction can be easily rationalised with CarQ no longer able to recognise the promoter. The Alt6 promoter has enhanced light induction relative to the wild-type. This phenotype is also true for Alt7, which has a mutation overlying the downstream CarD binding site. In both Alt6 and Alt7, the presence of the mutation is predicted to cause a reduction in the angle of a predicted bend intrinsic within the DNA of the promoter at the CarD binding site. It may be the case that the mutations reduce the angle of curvature of an inhibitory bend within the promoter, thus allowing enhanced induction. This would contrast with the conventional situation where curvature upstream of promoters typically enhances promoter activity.

Promoters Alt15 and Alt16 carry mutations around positions -145 and -151 respectively and they exhibit no light-induction in DK101. The mutation within promoter Alt15 overlies the transcriptional start site of *gufA* and the mutation carried by Alt16 lies around +6 relative to the *gufA* start site. The loss of light-

induction of these promoters could be due either to disruption of a cis-acting element of the *carQRS* promoter, or because of an indirect effect mediated by a change in the activity of the *gufA* promoter.

The ability to assess the effects of site-directed mutagenesis on the *carQRS* promoter *in vivo* provides a great deal of physiologically relevant infomation. However the analysis of the regions of importance within the *carQRS* promoter have been necessarily crude to date. Mutating clusters of three or four bases within the promoter allows a general dissection of regions of importance, with the assumption that the introduced mutations will be disruptive if they do lie at important regions. However, the conclusions that can be drawn from experiments of this nature are vague, with a typical conclusion being that the mutated area is of importance in promoter function. For a more refined analysis of promoter elements, single and double point mutations might perhaps provide more infomation about the precise nature of the elements within the promoter. This would be particularly interesting for the *carQRS* promoter, as its recognition by a member of the SigECF sigma factor family, makes the nature of its promoter-sigma factor interactions somewhat of an unknown quantity.

Further experiments need to precisely define the chromosomal context of the integrated Alt promoters and to determine whether their activities are truely CarQ-dependent. Only then can a meaningful, quantitative assessment of the effects of the mutations within the mutant promoters be made. Ideally the transcriptional start sites of each of the Alt promoters would also be checked to ensure that the transcriptional activity of the constructs is due to transcription from the *carQRS* promoter and not a novel, fortuitously-created promoter.

The first step along the road to a full understanding of the *carQRS* promoter has been taken with a gross mutational study having crudely identified regions of function within the promoter. Further refined mutagenesis may illustrate which elements of the promoter dictate the various features of *carQRS* expression, but only in conjunction with molecular techniques for the analysis of transcriptional intermediates and protein-protein / protein-DNA interactions.

5. Molecular analysis of promoter-CarQ interactions.

5.1. Introduction

One of the more widespread mechanisms whereby bacteria are able to alter their pattern of gene expression in response to a given stimulus, is through the production of alternative sigma factors. These generally lead to the expression of a set of genes of related function. For instance, σ^{32}, the heat shock sigma factor mediates expression of genes involved in the survival of stress (Gross, 1996). Similarly σ^{S}, the stationary phase sigma factor, is reponsible for expression of genes required for the entry into and survival of the stationary phase of the growth cycle (Loewen and Hengge-Aronis, 1994).

5.2. Sigma factors.

5.2.1.1. Diversity of sigma factors

Sigma or specificity factors are the polypeptides within the RNA polymerase which confer specificity for particular promoters sequences and as such they have various essential and conserved features. Sigma factors fall into two broad classes, the RpoD class and the RpoN class. These are very different proteins which recognise very different promoters. For instance the RpoD family sigma factors have recognition motifs consisting of hexamers around -35 and -10bp relative to the transcriptional start site. The RpoN family sigma factors however have recognition motifs centered at -12 and -24bp and unlike the RpoD proteins have been extensively characterised biochemically because of their ability to bind to their cognate promoters even in the absence of core RNA polymerase (Buck and Cannon, 1992). Sigma factors of the RpoN family are involved with functions such as flagella and pili formation, nitrogen fixation and utilisation of various metabolites (Ronson *et al.*, 1987). Transcription of RpoN dependent promoters always requires additional response regulator-type transcription factors to

overcome the energy barrier associated with open complex formation, which appears to be the rate-limiting step of initiation from all RpoN-dependent promoters (Geiselmann, 1997; Sasse-Dwight and Gralla, 1990).

Typically, bacteria contain a housekeeping or vegetative sigma factor (RpoD) which mediates transcription of housekeeping and essential genes. As might be expected for such an essential gene product, RpoD is highly conserved throughout the bacteria, recognises very similar promoters, and is functional in heterologous hosts (Inouye, 1990; Rudd and Zusman, 1982). There are also a wide variety of RpoD-family alternative sigma factors which cause expression of particular sub-sets of genes, e.g. the heat shock sigma factor σ^{32}. (For an example of the role of alternative sigma factors in a regulatory cascade see Haldenwang, 1995.)

Expression of an alternative sigma factor will cause an increase in the proportion of RNA polymerase holoenzymes containing that sigma factor and this leads to increased transcription of promoters recognised by that sigma. The dynamics of this process are uncertain. Alternative sigma factors may have higher affinity for core RNA polymerase than the housekeeping sigma. Alternatively, the proportion of holoenzymes containing a particular sigma factor could be purely determined by a dynamic equilibrium dependent on the relative concentrations of each sigma. This is not a trivial uncertainty, due to the low and limiting concentration of free core RNA polymerase in the cell. There are 2,000-10,000 copies of RNA polymerase per genome in *E. coli* (Ishihama, 1993; Bremer and Dennis, 1996) with most of it not 'free', since the majority of polymerase will be bound non-specifically to DNA or involved in the process of transcription.

There has been considerable confusion regarding the nomenclature of sigma factors.

Originally, sigma factors were described by a superscript denoting the molecular weight of the sigma factor. Thus *E. coli* σ^{70} refers to a sigma factor with a mass of 70kDa. The mass of sigma factors were originally assessed by SDS-PAGE. In actual fact these are usually incorrect since the sigma factor carries a preponderance of positive charge and this leads to an anomalously high mobility during SDS-PAGE. Thus for instance, CarQ has an apparent mass of around 27kDa but calculations based on translation of the gene sequence show its actual

mass should be 19.6kDa. Similarly, Sig70 has a mass of 70kDa although on SDS-PAGE it migrates at a rate consistent with a protein of 96kDa.

An alternative naming convention is to label the sigma factor with a superscripted capital letter e.g. σ^A. The letter used being dependent on the role of the factor within the cell and on the bacterial species the sigma factor is found in.

However, this has led to confusion with some similar sigma factors in different organisms being given different designations and the same letter designation being used for very different sigma factors in different species. Thus the vegetative sigma factor of *E. coli* is denoted SigD (σ^{70}, σ^D, encoded by *rpoD*), but is designated as SigA (σ^A, encoded by *sigA*) in *Bacillus subtilis,* while SigD of *Bacillus subtilis* is involved with flagella synthesis and is most similar to SigF of *E. coli.* (For a review of the current nomenclature of bacterial sigma factors see Lonetto and Gross, 1996.)

In 1988 there were only 13 putative sigma factors for which the primary sequence was known. Eleven years later, there are hundreds of sigma factor primary sequences known and several have also been proven to act as sigma factors, including two from *M. xanthus*, CarQ and SigA.

5.2.1.2. Structure of sigma factors

Sigma factors have four main functions. They bind to core RNA polymerase, recognise promoter DNA, mediate DNA melting during open complex formation and inhibit non-specific transcription. The first two of these properties are absolutely required for all sigma factors (Helmann and Chamberlin, 1988 and Gross *et al.*, 1996 for reviews).

Sigma factors of the RpoD family have a four domain structure as seen from homology matching. These are denoted regions 1, 2, 3 and 4 and have sub-domains denoted by decimals, for instance 2.4 for the fourth sub-domain of region 2. Region 2.4 has been shown to be involved in recognition of the -10 region of the promoter while region 4.2 binds to the -35 promoter region. The 4.2 region is likely to include a helix-turn-helix (HTH) motif in all sigma factors and at its C-

terminus has a highly basic region which may bind the phosphate backbone of DNA. Region 2.1 is suggested to have a role in binding to core polymerase while regions 2.1 and/or 2.3 are involved in melting the promoter DNA during open complex formation. Most isolated mutants of sigma factors have a mutation lying within the 2.4 or 4.2 regions. Siegele *et al.* (1989) were able to map interactions between sigma factors and their promoters to specific nucleotides and amino acid residues by a comparison of 13 missense mutants of *rpoD* (σ^{70} gene) and 37 mutant promoters. Their results confirmed a role for regions 2.4 and 4.2 in promoter hexamer binding.

Region 1 has an unknown function but is only of significant size in the housekeeping RpoD sigma factors. Dombroski *et al.* (1992) showed by filter binding assays of truncated forms of σ^{70} that region 1 seems to be required for the inhibition of specific and non-specific binding of the sigma factor to DNA. Interaction of sigma factor with core RNA polymerase would then be expected to relieve the inhibitory effect of region 1 allowing sigma factor-mediated binding to DNA. Severinova *et al.* (1996) suggest that the inhibition by region 1 is through interaction with the C-terminus of region 4. Region 1 has also been implicated in efficient conversion from open to elongating complex and in region 1.2 in particular, with isomerisation from the closed to open complex (Wilson and Dombroski, 1997). It should be noted that region 1.1 is essentially missing in all SigECF sigma factors, but region 1.2 is still present (Lonetto *et al.*, 1992).

Severinova *et al.* (1996) showed that a tryptic fragment of σ^{70} containing only region two was capable of binding to core RNA polymerase competitively with intact σ^{70}, and when bound to core was capable of binding specifically to oligonucleotides containing the -10 promoter consensus sequence (although not unless present with core) but wasn't able to initiate transcription. Comparison between the solely region 2 truncate and a truncated sigma lacking just region 4 showed that region 3 is essential for transcription whereas region 4 is not. The truncated protein lacking region 4 was able to initiate transcription from promoters containing only the -10 sequence.

Binding to core RNA polymerase is thought to be mediated by region 2.1 (Lesley and Burgess, 1989). Binding to core polymerase exposes the DNA binding domain of the sigma factor and also reduces affinity of the polymerase

holoenzyme to non-specific DNA. Region 2.3 has a preponderance of aromatic residues (44% aromatic as opposed to 7% over the entire protein). This suggests a possible role in binding the single-stranded DNA after open complex formation. A role for sigmas in the conversion of a closed complex into an active complex is suggested by the enhancement of the ability to photochemically cross-link sigmas to DNA when in the open complex compared to when in closed complex.

Apart from the analogous roles shared by sigma factors and the TATA-binding protein of eukaryotic RNA Pol II, disparate specificity factors are also found in some eukaryotes. For instance in *Saccharomyces cerevisiae* a 43kDa protein allows recognition of mitochondrial promoters (Schinkel *et al.*, 1987) while the core mitochondrial RNA polymerase is a single protein (Masters *et al.*, 1987). Specificity factors are found in other organisms including, for example, bacteriophage T4 has a sigma factor (Gp55) which mediates initiation of transcription from promoters which lack the -35 hexamer and have only the -10 region (Elliott and Geiduschek, 1984).

5.2.2. Sigma factors of *M. xanthus*.

The gene for the vegetative sigma factor (σ^A encoded by *sigA*) of *M. xanthus* was initially cloned by cross-hybridisation using a probe from the *E. coli rpoD*. SigA showed great similarity to the vegetative sigma factors of E. coli (σ^{70}) and *B. subtilis* (σ^{43}) (Inouye, 1990). Subsequent probing with *sigA* identified *sigB* and *sigC* which encode proposed development-specific sigma factors and the two putative sigma factors *sigD* and *sigE* (Apelian and Inouye, 1990; Apelian and Inouye, 1993; Ueki and Inouye, 1996). Probing with *rpoN* from *Caulobacter crescentus* identified the gene for σ^{54} from *M. xanthus*, which is essential in *M. xanthus* but in no other bacterium studied (Keseler and Kaiser, 1997). Recently, biochemical confirmation of the sigma factor activity of σ^A was provided by Biran and Kroos (1997) by transcriptional run-off assays from the *aphII* and *vegA* promoters, which have previously been shown to be expressed vegetatively.

The genes for two copies of σ^E have also been identified in *M. xanthus,* encoded by *rpoE1* and *rpoE2* (Ward *et al.,* 1998; Ward, M. pers.comm.). They are SigECF sigma factors (5.2.3) and appear to have a role in the regulation of motility.
Together with CarQ, this puts the number of sigma factors of *M. xanthus* identified to-date at nine, of which two (CarQ and SigA) have been demonstrated to act biochemically as sigma factors.

5.2.3. ECF sigma factors.

As mentioned earlier, CarQ is a member of a sub-family of the RpoD family known as the ECF sigma factors (Lonetto *et al.,* 1994). This group was originally identified as a group of 8 proteins responsible for the transcription of genes associated with an *e*xtra*c*ytoplasmic *f*unction. The original eight consisted of CarQ from *M. xanthus*, SigE of *Streptomyces coelicolor*, AlgT(AlgU) of *Pseudomonas aeruginosa*, HrpL of *Pseudomonas syringae*, SigE and FecI of *E. coli*, CnrH of *Alcaligenes eutrophus*, SigX of *Bacillus subtilis* and PbrA of *Pseudomonas* spp.. Since then a plethora of ECF sigma factors have been unearthed and they appear to be widely distributed (Missiakis and Raina, 1998). Sequencing of the entire *Bacillus subtilis* genome revealed 19 sigma factors, of which 7 were postulated to be members of the ECF sub-family (Kunst *et al.*, 1997). Likewise in *Mycobacterium tuberculosis*, 14 sigma factors were found including nine SigECFs (Cole *et al.,* 1998). However of these putative ECF sigma factors, very little is known except their homology to other sigma factors.
The greatest differences between the SigECF and other RpoD family factors lie in regions 1, 2.4 and 3. There is greater conservation of region 4.2 than of region 2.4 within the SigECF sigma factors, but not in comparison to non-ECF RpoD-family proteins. Consequently the promoter sequence profile for the SigECF sigmas is very clear at the -35 region and different from that of non-ECF factors. The promoter sequence profile derived from the known promoters of the first eight SigECF factors is CCGGAACTT at -35 and TCTNRt at -10. As can be seen in Figure 5-1, the sequence profile is virtually identical to the promoters of *carQRS* and *crtI* (+12) in the -35 region, but the -10 region shows very little similarity. Even less similarity is seen to the *crtI* (+0) promoter.

ECF profile	CCGGAACTT TCTnRt
carQRS	CGAGCGCCGGAAACACTTTCGCAGGTGGCCCGTAGAGGAGTC**G**
crtI (+12)	CCAGTGCCGGACGGTGCTCTTGTAACGTCCTGGCGGGTTCGC**G**
crtI (+0)	GGTGCTCTTGTAACGTCCTGGCGGGTTCGCGCGTTCGCCAGG**T**

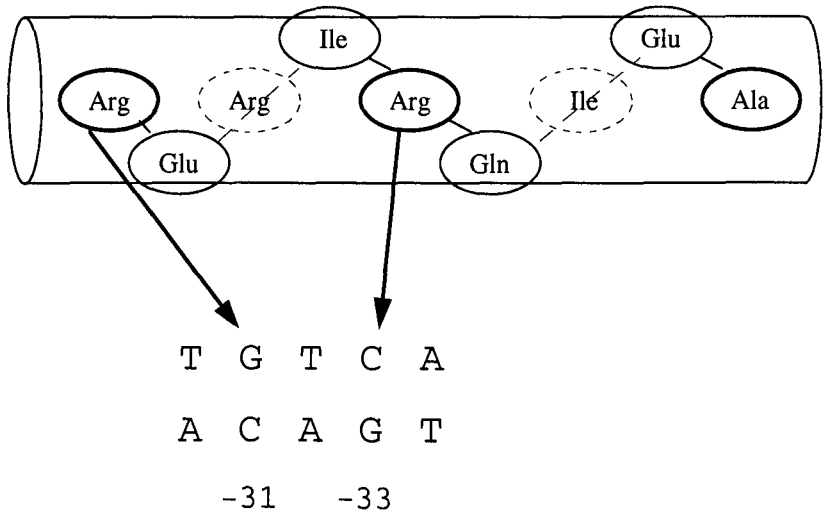
Figure 5-1 Comparison between SigECF-dependent, carQRS, crtI*(+0) and* crtI*(+12) promoters.*

Conserved bases in the carQRS, crtI *(+0) and* crtI *(+12) promoters are underlined. Transcriptional start sites are in bold face.*

Siegele *et al.* (1989) and Gardella *et al.* (1989) assessed the ability of a range of mutant Sig70 proteins, to mediate recognition of a set of mutated Sig70-dependent promoters, enabling the assignment of specific base-amino acid residue interactions (Figure 5-2). They showed that Arg584 of Sig70 contacted the base complementary to C(-31) and that Arg588 contacted the base complementary to the G at promoter position -33. In the helix-turn-helix motif of CarQ as shown by Lonetto *et al.* (1994), Arg584 is replaced with a proline residue while Arg588 is conserved. In the *carQRS* promoter, position -33 is a G and position -31 is an A. Therefore the Arg-C interaction at position -33 found for Sig70 is maintained for

CarQ, however the Arg-G interaction at -31 is not, instead being replaced with a putative Pro-T interaction. Whether this is a possible protein-DNA interaction or not is unknown, however, care must be taken when comparing DNA binding of Sig70 with that of a SigECF sigma factor as their helix-turn-helix motifs are very different. The HTH of Sig70 has a downstream highly basic region which is lacking in the SigECF sigmas (Lonetto *et al.*, 1994). The chemical character of the two recognition helices is also different, with polar, acidic, hydrophobic and basic residues lying at different positions. Since the 2.4 and 4.2 regions of the two types of sigma factor are separated by very different distances, a direct comparison of their DNA binding would be premature, as the exact regions of contact between the SigECFs and their promoters has not yet been subjected to experimental scrutiny.

A)



B)

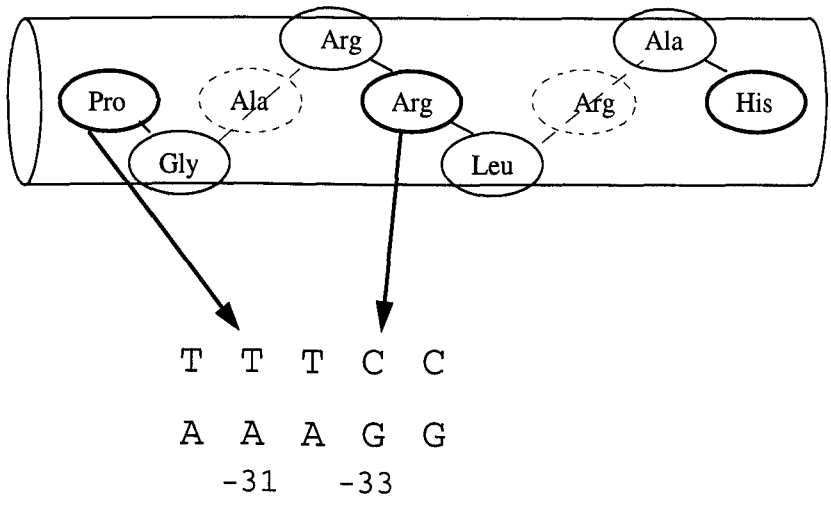


Figure 5-2 Recognition of promoter bases at -35 by the HTH of Sig70 and CarQ.

A) Interaction between the recognition helix of Sig70 region 4.2 and the consensus Sig70-dependent promoter. Adapted from Siegele et al. *(1989). B) Mirrored interactions between CarQ and the* carQRS *promoter.*

Some SigECFs are able to mediate transcription of SigECF-dependent promoters in heterologous species. Hershberger *et al.* (1995) showed that the ECF sigma factor AlgT of *Pseudomonas aeruginosa* was able to recognise the *E. coli* SigE-dependent promoter *rpoH* P3 *in vitro* (*P. aeruginosa* also contains the ECF sigma

factor PvdS). However, SigE of *Streptomyces coelicolor* was unable to mediate transcription of P^{carQRS} *in vitro* (Browning, 1997).

Another interesting feature noted by Lonetto *et al.* (1994) was that of the eight SigECFs, four (CarQ, AlgT(AlgU), *E. coli* SigE and CnrH), were thought to mediate transcription of negative regulators of themselves, i.e. anti-sigma factors. This has recently been shown to also be the case for *S. coelicolor* SigR (Buttner *pers. comm.*). In fact, of the 19 ECF sigma factors described by Missiakis and Raina (1998), 8 mediate transcription of their own anti-sigma factors and seem to share a common operon organisation, with the first gene in the operon coding for the sigma factor and the second gene in the operon encoding the anti-sigma factor. For instance, *carH/orf1* of *Alcaligenes eutrophus*, *algT(AlgU)/mucA* of *Pseudomonas aeruginosa* and of *Azotobacter vinelandii*, *rpoE/rseA* of *E. coli*, *Haemophilus influenza*, *carQ/carR* of *M. xanthus*, *sigR/rsrA* of *Streptomyces coelicolor*, *rpoE/orf2* of *Photobacterium* SS9 and *pupI/pupR* of *Pseudomonas putida*.

A good example of control by anti-sigma factor activity is the case of SigR of *Streptomyces coelicolor* A3(2) which mediates a response to oxidative stress through its activation of the *trxBA* operon (encoding thioredoxin reductase and thioredoxin) which serves to increase the reducing nature of the cytoplasm (Paget *et al.*, 1998). Downstream of the positively autoregulated *sigR* gene lies *rsrA* which is believed to encode an anti-sigma factor. The current model proposes that under a normal reducing environment, SigR is bound by the anti-sigma factor, RsrA, which contains seven thiols. In conditions where thiols are oxidised, the thiols in RsrA form a disulphide bond (Paget and Buttner, 1998) and oxidised RsrA releases SigR. The sigma factor then causes expression of *trxBA* which returns the redox poise within the cell to normal, whereupon RsrA is reduced and once again binds to SigR.

As most regulons involve homeostatic feedback loops, the fact that negative regulators of the SigECFs are transcribed by the SigECFs themselves may not be particularly surprising. Indeed, as part of co-ordinated regulons, it might be more surprising if the anti-sigma factors weren't transcribed by the actual sigma factors they inhibited. In a similar vein, the observation that some of the SigECF sigma factors seem to govern small regulons (Lonetto *et al.*, 1994) is probably a

consequence of the SigECFs being alternative sigma factors more than because it is a conserved feature of the SigECFs. It could be disputed whether the regulon governed by CarQ is indeed small. It is true that as far as we know CarQ only causes expression of the *carQRS* and *crtI* loci, but CarS, which is CarQ dependent mediates expression of *crtEBDC*, an operon of 6 genes. Therefore CarQ is responsible for expression of at least 10 genes which is not a small regulon.

The complement of SigECF sigma factors identified in *M. xanthus* has recently been increased to three with the discovery of *rpoE1* and *rpoE2* (Ward *et al.*, 1998 and Ward, M. pers.comm.). Both RpoE1 and RpoE2 share 40% identity with CarQ, particularly between regions 2.2 and 2.4. Of the three sigma factors, CarQ is smaller than RpoE2 by 23 residues, which is smaller than RpoE1 by 16 residues. All three sigma factors share homology across their length, including regions 2.4 and 4.2, but with more extensive identity at region 4.2 (Figure 5-3).

Alignments have also been suggested between the *carQRS* promoter and putative promoters upstream of the *rpoE* genes themselves (on the assumption that RpoE is responsible for its own expression) (Zusman, D. per. comm.). These possible alignments are shown below (Figure 5-4). As is true for the SigECFs generally, there is greater conservation at -35 than at -10, a reflection of the greater conservation of region 4.2 of the sigma factor over region 2.4. However, the benefits of assessing a consensus between the SigECFs of *M. xanthus* seems dubious, as the three sigma factors must recognise different promoters almost by definition. In addition it seems that the *rpoE1* and *rpoE2* promoters have been identified solely by their homology to the SigECF dependent promoter consensus (Ward *et al.*, 1998), and thus similarity to the *carQRS* promoter is a circular arguement.

Region 2.4

CarQ	68	**IAANAARD**ALR**H**QR**H**VDAYA	88
RpoE1	88	**IA**L**N**L**AKNH**AR**Q**V**Q**R**W**R**P**VL**	108
RpoE2	70	**I**LT**N**TF**I** **NR**Y**RR**K**V**K**ER**T**VV**	90

Region 4.2

CarQ	131	**VEGWSF**E**EIGA**LR**GIS**PG**AARLRAHRG**YE**KLR**E	163
RpoE1	156	DG**G**LA**FKDIAE**T**LGIT**ENN**AKV**QF**HHAM**KRLK**A	188
RpoE2	146	**LQEFSYKEIAEIL**ECPV**G**TVMS**R**LF**RGRKLL**QK	178

Figure 5-3 Alignment of CarQ, RpoE1 and RpoE2.

Regions 2.4 and 4.2 are shown as identified by Lonetto et al., *1994. Conserved and similar residues are in bold. For this purpose, the following amino acids are considered to be similar: (I,L,M,V), (H,K,R),(D,E,N,Q),(A,G),(F,Y,W) and (S,T). The helix-turn-helix motif of region 4.2 is underlined.*

carQRS	AG**C**GCC**GGAA** ACAC**TTTC**G**C**A**GG**TGGCCCG**T**AGA**GG**A**G**T**C**GGGT
rpoE1	TG**C**TAG**GGAA** TATG**TTTC**CCT**GG**CCGGC G**T**CGT**GTATT**CCCGA
rpoE2	AA**C**TCG**GGAA**GGCATG**TTTC CTGG**CTCAATA**T**CTT**G**C**AA**TTCACG
	-35 -10

carQRS	A**GC**GCC**GGAA**ACAC**TTTC**G**C**A**GG**TG**GC**CCG**T**AGA**GG**A**G**T**C**GG**GT**
rpoE1	T**GC**TAG**GGAA**TATG**TTTC**CCT**GG**CC**GGC GT**C**GT**G**T**A**T**TC**CC**G**A
	-35 -10

Figure 5-4 Alignment of the promoters for carQRS, rpoE1, *and* rpoE2.

Two alignments are shown between the carQRS *and* rpoE1 *promoters. Conserved bases are in bold. Underlined bases are at -35 and -10 relative to the presumed transcriptional start site.*

A common feature of the SigECF sigma factors is that they lack almost all of region 1. As region 1 is responsible for the inhibition of binding between σ^{70} and DNA (Dombroski *et al.*, 1992), it was thought possible that SigECF sigma factors lacking region 1 might be able to bind directly to their cognate promoters. This was shown by Dombroski *et al.* (1993) for FliA (SigF) of *S. typhimurium* which is a non-ECF sigma factor, but which has a very small region 1. Binding to DNA was noted by filter binding assays but was too weak to be visualised by a gel retardation approach. CarQ has a region 1 which is two residues smaller than that of FliA and as a member of the SigECF factors, CarQ may be different enough from FliA that it may bind to its cognate promoters with an affinity sufficient to be observed by band-shift assays.

5.3. Binding between CarQ and the *carQRS* promoter.

Since CarQ has a particularly small region 1, it was conceivable that any interaction with its cognate promoter may be strong enough to be visualised by a gel retardation approach. Browning (1997) provided a protocol for the production of pure, biologically active CarQ allowing an *in vitro* assessment of binding to P^{carQRS} using solely purified components.

5.3.1. Gel retardation assays.

Interactions between proteins and DNA can be detected in a number of ways. Filter binding assays are one of the oldest methods of demonstrating DNA-binding activity in proteins. In this assay, a radio-labelled piece of DNA is mixed with the protein of interest and passage of labelled probe through a nitrocellulose filter is followed. Binding of protein to the probe manifests as a reduction in the amount of radiation that passes through the filter. The technique is very sensitive but requires a large number of stringent controls (A good example of the use of filter binding assays is given in Dombroski *et al.*, 1992).

A less sensitive but more widespread method for assessing DNA binding by proteins is the gel retardation, or band-shift assay. In this method a radiolabelled probe piece of DNA is mixed with protein and the mixture resolved by non-denaturing gel electrophoresis. Protein bound to the DNA probe effectively increases the apparent molecular weight of the DNA and will reduce the rate of migration of the probe through a gel.

In both of these methods it is important to be able to discriminate between specific and non-specific binding. When assessing specific binding between a protein and a particular DNA sequence non-specific interactions must be eliminated. This is usually achieved by adding unlabelled 'non-specific' protein and DNA in high concentrations. Therefore if the protein is a non-specific DNA binding protein, due to the excess of unlabelled DNA, most of the protein will bind to unlabelled DNA and relatively little of the protein will bind to the labelled DNA and cause its retardation during electrophoresis. In such a fashion a relatively few experiments will allow discrimination between specific and non-specific binding.

As a simple and yet conclusive method for investigating interactions between DNA and protein, gel retardation assays were performed to determine whether specific interactions could be observed between CarQ and the *carQRS* promoter.

5.3.2. Production of CarQ

Plasmid pDFBT13 contains *carQ* cloned into the *Nde*I / *Bam*HI sites of pET-3A. The *Nde*I site overlies the Met^1 codon so no extra amino acids are introduced into CarQ. Two additional stop codons are located in-frame, downstream of the coding sequence and expression of *carQ* is under the control of the T7 promoter.

This plasmid was transformed into E. *coli* BL21::DE3, using the TSS method of Chung *et al.* (1989), using Amp^{100} selection. BL21::DE3 is protease deficient and contains the integrated λ phage DE3 which has a copy of the T7 polymerase gene under the control of the *lac* promoter.

A freshly transformed strain was used to produce CarQ as described in section 7.4.2. 200ml of selective medium was subcultured with an overnight culture of the transformed strain and incubated until an OD_{600} of 0.8 was attained, then

induction was achieved by addition of IPTG. After 4 hours, the cells were pelleted and frozen. Purification and resolubilisation of the inclusion bodies was then performed as described in section 7.4.3. Progress of the purification and refolding procedures was followed by SDS-PAGE and activity of the CarQ preparation confirmed by *in vitro* transcriptional run-off assay (section 5.4).

5.3.3. Gel Retardation of the *carQRS* promoter by CarQ.

Initial experiments used a range of concentrations of CarQ added to a probe containing the *carQRS* promoter. Binding reactions were performed as described in section 5.3.1.

Probe DNA was amplified from *M. xanthus* chromosomal DNA containing a wild-type copy of the *carQRS* promoter (primers 1 and 2, section 7.4.4.1) and end-labelled with ^{32}P by reaction with T4 polynucleotide kinase. Figure 5-5 demonstrates retardation of migration of the labeled DNA probe by CarQ.

Figure 5-5 Gel retardation of the carQRS *promoter by CarQ.*

Lane 1: No CarQ. Lanes 2-5: Increasing concentrations of CarQ (constant volume of 1000x, 100x, 10x and 1x dilutions of CarQ) used in binding reactions.

This was a promising result with the presence of CarQ retarding labelled P^{carQRS} and with more retarded band observed with greater quantities of CarQ. A reduction in the amount of unretarded probe could also be seen at high CarQ concentrations.

However, the bands due to retardation of the probe coincided with the bottom of the wells of the acrylamide gel. This is unusual as even large DNA/protein complexes can enter acrylamide upon electrophoresis (Dent and Latchman, 1993). To determine whether the effect was due to CarQ or a trace contaminant from the purification protocol, the purification procedure was repeated using BL21::DE3 untransformed with pDFBT13. The product of this purification was used in the place of CarQ during another attempt at the retardation assay. This time, no retarded bands were observed regardless of probe used, and the presence/absence of competitor DNA, so the initial retardation seen in Figure 5-5 must be CarQ-dependent.

These results suggested that CarQ was preventing entry of probe DNA into the acrylamide gel. This could be a specific or non-specific effect so retardation assays were repeated using i) a non-promoter-containing fragment of DNA as the probe. ii) addition of varying amounts of unlabeled competitor DNA. iii) addition of unlabelled non-specific DNA.

Retardation of probe by CarQ was seen irrespective of the probe used, even when the probe used was a non-promoter-containing fragment of completely synthetic DNA (provided by Dr. G. Owenson, pers.comm.) Thus the retardation effect caused by CarQ is non-specific. Addition of a range of amounts of unlabelled pIC19H or of unlabelled probe, showed a reduction in the intensity of the retarded band, again indicating that retardation was non-specific (data not shown).

In conclusion, it appears that CarQ causes the sequence-independent prevention of DNA entrance into a polyacrylamide gel. It is perhaps most probable that the preparation of CarQ, while producing active CarQ, also produced inactive denatured aggregates of protein which 'trapped' DNA in a state unable to penetrate the non-denaturing polyacrylamide gel. Browning (1997) showed that 6His-CarQ expressed in *E. coli* always precipitated, irrespective of concentration and even

following dialysis to remove sarkosyl after successful resolubilisation. It is possible that CarQ is also prone to precipitation, although not to the extent of 6His-CarQ, with the precipitation induced as the protein is electrophoretically induced to enter an acrylamide gel.

Potentially, if the inactive protein could be removed from the CarQ preparation it may prove possible to observe a sequence-dependent retardation of P^{carQRS} by CarQ. The experiments could alternatively be repeated using cell extracts, or using CarQ directly purified from a culture of *M. xanthus*. However, results obtained by Browning (1997) suggest that CarQ is a relatively unstable protein and present at very low levels within the cell. CarQ could not be detected in *M. xanthus* using anti-CarQ anti-sera even in strains carrying a *carR* lesion. Additionally, CarQ instability would explain the rapid down-regulation of *carQRS* expression on the production of carotenoids and the consequent transition from light to dark and also explain the much lower light-induction at loci far removed from the site of CarQ production, i.e. at *crtI* and at *M. xanthus* AttB (sections 2.2 and 4.3).

5.4. Transcription run-off assays.

Transcription run-off assays are a means of assessing the transcriptional efficacy of promoters and sigma factors *in-vitro*, using purified transcriptional components.

5.4.1. Nature of the assay.

For the *in vitro* pre-binding transcription run-off assay, a template fragment of DNA is combined with core RNA polymerase, a sigma factor and rNTPs. If the sigma factor is capable of directing initiation of transcription from a promoter contained on the DNA fragment, mRNA will be produced. Production of mRNA by the system is assessed by the incorporation of ^{32}P-α-CTP into the nascent macromolecule, resolution by gel electrophoresis and visualisation by autoradiography. The assay is a sensitive technique for the determination of sigma

factor activity in a protein sample and can also be used to assess the ability of a promoter sequence to be recognised by a specific sigma factor.

However, there can be problems associated with the interpretation of data derived from *in vitro* transcription run-off assays. If the assay is being used to determine whether a protein has sigma factor activity, the production of a specific mRNA product is incontravertible proof that the protein is a sigma factor (for example Browning, 1997). More care must be taken when looking at the ability of various promoters to be recognised by a specific sigma factor. This is a consequence of the non-physiological conditions employed in the assay, i.e. the artificial situation where there is a single sigma factor present with a single promoter. Additionally, the concentrations of each of these species and of the core RNA polymerase are typically much higher than those found within a cell. This can lead to anomalous results. For instance, work on the *dagA* gene of *Streptomyces coelicolor* used the transcriptional run-off assay to prove that SigE could mediate transcription from the P2 promoter of *dagA* (Buttner *et al.*, 1988). The implication was clearly that SigE mediated expression of *dagA* from P2 *in vivo* (Lonetto *et al.*, 1994). However, when a *sigE* knock-out mutant was created it was discovered that transcription from *dagA* P2 *in vivo* was unaffected (Jones *et al.*, 1997). Thus it would seem likely that it is possible to get a positive result from the transcriptional run-off assay by virtue of the assay conditions. This is not true in all cases however. When the ability of SigE to initiate transcription from the *carQRS* promoter was assayed, no transcription product was observed (Browning, 1997). This is probably since the promoters thought to be recognised by SigE and CarQ are very different around -10, even though very similar at -35 (Figure 5-6).

The implication is therefore that it is possible to obtain a transcription product from a promoter and sigma factor combination that may not produce mRNA *in vivo*. This would presumably be more likely if the promoter being assessed was similar to a promoter that *is* recognised by that sigma factor *in vivo*.

dagA P2	GCGTTCCGGAACTTTTTGCACGCACGCGAGCTCTCGAATTTT
carQRS	GAGCGCCGGAAACACTTTCGCAGGTGGCCCGTAGAGGAGTCG
crtI (+12)	CAGTGCCGGACGGTGCTCTTGTAACGTCCTGGCGGGTTCGCG

Figure 5-6 Comparison of the promoters whose recognition is mediated by CarQ and by SigE.

CarQ-dependent promoters shown are for crtI *and for* carQRS. *The SigE-dependent promoter shown is for* dagA *P2. Underlined bases correspond to bases conserved in* P^{carQRS}.

5.4.2. Transcription run-off assays of mutant *carQRS* promoters.

Bearing this in mind, attempts were made to mirror the *in vivo* assays of the mutant promoters constructed by Berry (1998), by assessing the *in vitro* activity of the same promoters. Browning, (1997) had previously used the transcription run-off assay to prove that CarQ had sigma factor activity and was able to mediate transcription from P^{carQRS}, although he was unable to show that CarQ could cause transcription of *crtI*.

PCR was used to amplify 500bp fragments of DNA from the pAEB6XX series of plasmids to use as templates for *in vitro* transcriptional run-off assays. The PCR products were each 480bp long, with the promoter located at such a position that any mRNA produced from promoter within the template would be 317bp long. Soluble CarQ was made as described in section 7.4.2 and assays were performed as described in section 7.4.5.

Using a PCR product derived from pAEB600 which carries the wild-type promoter, the run-off assay gave two bands of interest on the autoradiograph (Figure 5-7). The first at ~500nt corresponds to an RNA product transcribed from one end of the template to the other. This product is formed independently of sigma factor by core RNA polymerase and is always obtained in run-off experiments. It is produced as core RNA polymerase is capable of binding to and initiating transcription from the end of a fragment of DNA. The second product,

of ~300nt is only found if CarQ is present and is the result of transcription from the internal promoter within the template. The ability to obtain a transcript from P^{carQRS} using pure CarQ and core RNA polymerase is a poor reflection of what occurs *in vivo* already as transcription of *carQRS in vivo* absolutely requires the DNA binding protein CarD.

Assays were also performed for mutant promoter Alt4. In this case both full-length and internal transcripts were obtained from the assay (Figure 5-8). The Alt4 promoter contains a mutation around position -32 within the highly conserved -35 promoter hexamer. When the promoter activity of Alt4 was assessed *in vivo* it showed no light-induction and the inference was therefore that the mutation in Alt4 abolished light-induction of promoter activity. This conclusion is contradicted by the result of the *in vitro* run-off assay which suggests that CarQ can mediate expression from Alt4. It seems likely that the ability to produce a transcript from Alt4 is a result of the conditions employed in the run-off assay and not a true reflection of what occurs *in vivo*. The similarity of the Alt promoters to the wild-type *carQRS* promoter makes it likely that all run-off assays of Alt promoters will yield positive results, making the worthiness of the technique for this application doubtful. The presence of a large number of non-specific RNA products within the background noise seen in Figure 5-7 and Figure 5-8 implies that transcription is being initiated at multiple sites along the template DNA fragment, perhaps at 'initiation hot-spots', not at the promoter. This again argues for inappropriate conditions used in the run-off reactions. A further consideration is that the core RNA polymerase used is obtained from *E. coli* not *M. xanthus*. It is possible that the differences between the core RNA polymerase from the two species may influence the results of the transcription assays using CarQ.

The benefits of the *in vivo* assays of promoter activity over the *in vitro* assays are manifold. They provide temporal and quantitative data, obtained *in situ,* in the organism of interest, under whichever conditions are required. Conversely, the *in vitro* assays are purely qualitative, employing reaction conditions far removed from those in the organism and the results obtained can therefore be ambiguous to interpret.

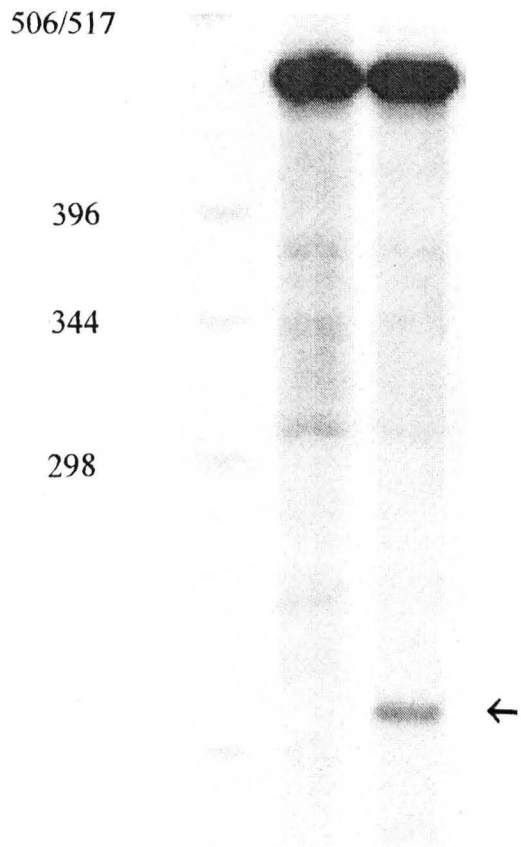


Figure 5-7 Transcriptional run-off of P^{carQRS} directed by CarQ.

Lane 1: Markers. Lane 2: Core polymerase only. Lane 3: CarQ and core polymerase.

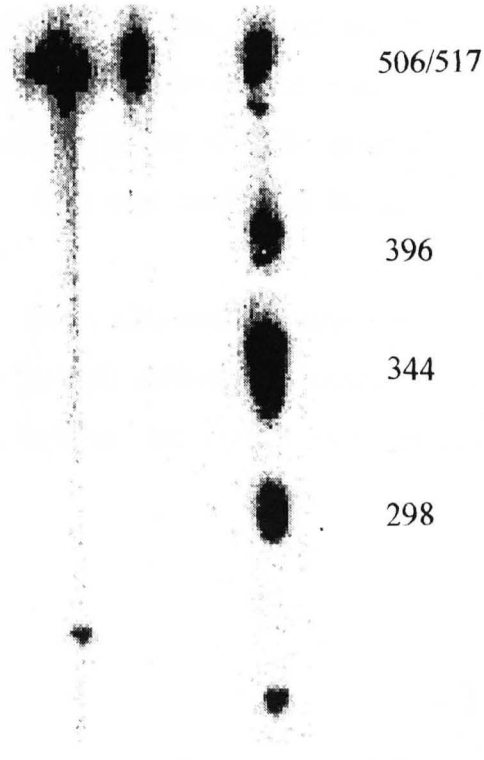


Figure 5-8 Transcriptional run-off of Alt4 directed by CarQ.

Lane:1 CarQ and core polymerase. Lane 2: Core polymerase only. Lane 3: Markers..

Due to time constraints run-off assays were not performed for any of the other Alt promoters. Another factor which hindered the performance of the run-of assays was the instability of the purified preparations of CarQ, which lost sigma factor activity at a significant rate and made any CarQ-mediated transcripts indistinguishable from the background products. Although it is difficult to determine an exact half-life of stability, the preparation had lost nearly all activity within two weeks of its production. This may be a consequence of the proteins natural instability or be due to precipitation at -20^0C during storage. Perhaps activity would not be lost at such a rate if the samples are stored at -80^0C. It should be noted that the CarQ preparations were no more than two days old when used in the binding assays described earlier and are assumed to have been biologically active at the time. This was confirmed by successful *in vitro* transcription run-off assays.

Future work will extend the use of the *in vitro* run-off transcription assay to assess all other Alt promoters. Whether this will yield data of comparable quality to that obtained from *in vivo* assays is unlikely, however it *may* generate qualitative results of interest.

5.5. Conclusions.

While *in vivo* assessment of the activity of mutated promoters provides a great deal of qualitative and quantitative data, *in vitro* studies have shed relatively little light on the nature of the sigma factor-promoter interactions between CarQ and P^{carQRS}. Experimental work proved unable to demonstrate direct binding between CarQ and the *carQRS* promoter. Whether this was a consequence of high levels of protein aggregates in the CarQ preparation or due to an inability of CarQ to bind directly to its cognate promoters is unknown. A more sensitive technique would have to be employed to determine that CarQ can not bind to P^{carQRS} without core RNA polymerase.

The use of *in vitro* transcription run-off assays to analyse the effects of mutations within P^{carQRS} also met with limited success. It is possible to obtain a specific transcription product due to the recogition of P^{carQRS} by CarQ. However it is also

which *in vivo* assays show to be defective in light-induction. This is likely to be due to the non-physiological conditions in the actual run-off transcription reactions and suggests that CarQ-mediated transcription products may be obtained from any promoter which is similar to wild-type P^{carQRS}.

A perhaps more appropriate method of dissecting functional regions within P^{carQRS} can be envisaged, made possible by the genetic tools available for manipulation of *M. xanthus*. A library of all possible point mutations within the -35 and/or -10 promoter hexamers could be generated easily using degenerate oligonucleotide-mediated mutagenesis (essentially a random-directed mutagenesis). If engineered appropriately, these mutant promoters could be introduced into *M. xanthus* such that selection favoured homologous recombination across P^{carQRS} resulting in replacement of wild-type P^{carQRS} with a mutant form. Isolation of all mutations causing a Car^C or Car^- phenotype would then provide a directory of base changes within the promoter which affect recognition by CarQ. This may provide a better strategy for the isolation of non-silent promoter mutations than the engineering of crude site-directed mutations and the detemination of their effects, if any.

6. Conclusions.

6.1. The role of *carA*

Car^C mutants caused by a *carA* mutation show greatly decreased levels of transcription from the *carQRS* and *crtI* promoters. By preventing the constitutive production of carotenoids in a *carA* strain and assessing promoter activities, it was apparent that the reduction in promoter activities due to the *carA* lesion was an indirect effect due to constitutive carotenogenesis, rather than due to loss of direct promoter activation by CarA. It seems that the only role of CarA in carotenogenesis is to repress *crtEBDC* expression in the dark.

The production of carotenoids was shown to be responsible for the feedback regulation of the carotenogenic regulon. Thus in a Car^- mutant, the negative feedback loop cannot be established and the regulon exhibits increased activity. Conversely, in a Car^C mutant, the constitutive production of carotenoids causes precocious establishment of the feedback loop and induction by light is diminished.

6.2. The role of *crtI*.

Previous work has shown that expresion of *crtI* is much more complicated than that of *carQRS*. The *crtI* locus has a biphasic pattern of induction, with the latter phase being enhanced by carbon limitation. Further differences between the *crtI* and *carQRS* promoters are suggested by sequence comparison and *in vitro* molecular studies. Browning (1997) showed it was possible to obtain transcriptional run-offs from the *carQRS* but not the *crtI* promoter.

Analysis of *crtI* promoter activity upon carbon limitation showed that in a *carR* mutant, the *crtI* promoter is still light-induced, implying a mechanism of light-induction separate from the CarQ-dependent mechanism.

The possible occurance of uncharacterised genes involved in carotenogenesis downstream of *crtI* was considered, and sequencing confirmed the presence of two genes downstream, and the start of a divergent gene upstream, of *crtI*.

Immediately downstream of *crtI* is the gene, *olpA*, for a thirty residue peptide whose initiation codon overlaps the termination codon of *crtI*. Downstream of *olpA* is a gene, *gufB*, whose mycobacterial homologues have no clear function. The termination codon of *olpA* lies over a putative ribosome binding site for *gufB*, suggesting a possible co-transcription and translational coupling of *crtI*, *olpA* and *gufB* and conceivably a shared involvement in carotenogenesis.

6.3. CarQ and the *carQRS* promoter.

The *carQRS* promoter is a large complicated promoter with many regions of importance. Site-directed mutation has enabled an *in vivo* confirmation of the areas of interest within the promoter. Mutation of the -35 hexamer, the early transcribed region, and the region within P^{carQRS} overlapping the *gufA* promoter start site all cause a loss of promoter function. Enhanced induction of the *carQRS* promoter was seen on mutation of the downstream CarD binding site, or on mutation of the region between the CarD binding site and the -35 hexamer. This is probably due to a change in the curvature of the DNA at the promoter which is predicted to have a sharp bend at the CarD binding site (Berry, 1998). The mutation introduced into the -10 hexamer caused constitutive promoter activity, presumably due to promiscuous activation by a non-CarQ sigma factor. Mutation of the region between the *carQRS* transcriptional start site and the promoter -10 region had no effect on promoter activity.

Attempts to detect direct binding between CarQ and the *carQRS* promoter by gel retardation assay were unsuccessful. Whether this is a result of the binding reaction conditions used, or because CarQ will not bind promoter DNA directly, is not known.

Upon performance of *in vitro* run-off transcription assays to assess the effect of mutations within the *carQRS* promoter, it was found that a mutant promoter (Alt4) which is uninduced *in vivo,* gave a transcription product *in vitro*. It appears that *in vitro* transcription is an inappropraiate technique to assess the activity of mutant promoters, most probably as a consequence of the non-physiological conditions used in the assay.

6.4. Future Perspectives.

The studies described in this work have important consequences and suggest further directions, for research into carotenogenesis in *M. xanthus*. The general conclusion from chapters two and three is that when assessing promoter activities of the *car* genes, the consequences of the Car phenotype of the strain must be borne in mind, as it has a drastic indirect effect on promoter activities.

Chapter three also identified at least two novel genes downstream of the *crtI* locus. Future work may enable elucidation of the functions of these genes and possibly implicate them with a role in carotenogenesis. Further sequencing around the *crtI* region, may identify other carotenogenic genes at this locus.

Chapters four and five attempted to elucidate molecular details of the interaction between CarQ and the *carQRS* promoter. The *in vivo* assays of mutant promoters described in chapter four provided infomation about which areas within the *carQRS* promoter are of imporance, and will enable future work to engage in a more fine-scale mutagenesis of the promoter.

Attempts to perfrom a variety of *in vitro* assays of interactions between P^{carQRS} and CarQ demonstrated a range of difficulties with the techniques used, when applied to the system at hand. Further analysis of CarQ / P^{carQRS} interactions will probably be best served by *in vivo* assays. Perhaps a more pressing area for study is the molecular nature of the action of CarS, which derepresses *crtEBC* in the light, and has no known homologues.

7. Materials and Methods.

7.1. Bacterial strains and plasmids.

7.1.1. Cultivation of *Escherichia. coli.*

7.1.1.1. Growth conditions and storage of *E. coli.*

E. coli strains were typically incubated in LB medium at 37°C, in liquid or on agar. For storage, plates were kept at 4°C and re-streaked every 2 months. Longer storage was achieved by centrifugation of an overnight culture followed by resuspension of pelleted cells in 1ml of 50% glycerol with storage at -80°C.

7.1.1.2. Materials required for the cultivation of *E. coli.*

LB medium ($1dm^3$).

5g Yeast Extract	15g agar (LB agar)
10g Tryptone	7.5g agar (LB soft agar)
5g NaCl	

7.1.2. Cultivation of *Myxococcus xanthus.*

7.1.2.1. Growth conditions and storage of *M. xanthus.*

Strains of *M. xanthus* were incubated in DCY or YT media at 33°C, in either liquid or on agar plates. Strains were stored on agar plates at 18°C and were re-streaked every two to three weeks. Storage of strains for longer periods was achieved by mixing 0.9ml of a two day old culture (late log phase) with 0.1ml of DMSO and freezing at -80°C.

7.1.2.2. Materials required for the cultivation of *M. xanthus*.

DCY medium ($1dm^3$).

20g Bacto casitone
2g yeast extract
10mM Tris HCl (pH. 8.0)
8mM $MgSO_4$

15g agar (DCY agar)
7.5g agar (DCY soft agar)

YT medium ($1dm^3$).

10g yeast extract
5g tryptone
lmM NaCl, 1mM $MgCl_2$ and 1mM $CaCl_2$.

7.1.3. Antibiotics

Stock solutions were made and stored as described in Sambrook *et al.* (1989). Antibiotics were used at a final concentration described in the appropriate protocol.

7.1.4. Bacteriophage used.

P1 *clr*-100 Tn*9* (Rosner, 1972)

Mx4-LA27 (Avery and Kaiser, 1983)

Mx8 (Martin *et al.*,1978)

7.1.5. Plasmids used.

Plasmids used during this study are detailed below. The name, description and key features of each plasmid are also given.

pAEB600	-P^{carQRS} (232bp *Bam*HI/*Eco*RI fragment from pAEB120), *lacZ*, P1*incΔ*, Mx8 *attP*, p15A *ori*, Km^R, *fd* term (Berry, 1998)
pAEB601-pAEB616.	Constructs identical to pAEB600 except that the copy of P^{carQRS} carried has been mutated. Thus pAEB601 carries mutant promoter Alt1 and pAEB607 carries promoter Alt7. (Berry, 1998).
pDAH217	-P^{carQRS}, *lacZ*, P1*inc*, Ap^R, p15A *ori*, Km^R + ¼ IS*50*L (Hodgson, 1993).

pDAH274	- *lacZ*, P1 *incΔ*, Ap^R, Km^R, p15A *ori*, (Hodgson, 1993). Identical to pDAH283.
pDFBT13	-T7 promoter (Φ10) driving expression of *carQ*, Ap^R, pBR322 *ori*. (Browning, 1998).
pDEW100	-P^{crtI} and *crtI*, Ap^R,Km^R p15A *ori* (This Study)
pDEW300	-*crtI* (internal 1.2Kb fragment), *lacZ*, P1 *incΔ*, Ap^R, Km^R, p15A *ori* (This Study).
pMAR202	-*crtI* (20kb *Kpn*I, *Eco*RI fragment), Km^R, (Fontes *et al.*, 1993).
pMAR206	-P^{crtI} and *crtI* (2.6kb *Pst*I fragment), *lacZ*, P1 *incΔ*, Ap^R, Km^R, p15A *ori* (Fontes *et al.*, 1993)
pPR108	-Mx8 *attP* (5.6kb *Pst*I fragment), Km^R, P^{carQRS}, *lacZ*. (Robson, 1992)

Ap^R, ampicillin resistance determinant; fd *term, major transcriptional terminator of coliphage fd; Km^R, kanamycin resistance determinant; Km^R +¼IS50L, kanamycin resistance determinant derived from Tn5 including the portion of IS50L containing the promoter; Mx8 AttP, myxophage Mx8 AttP phage attachment site;* P^{carQRS}*, light-inducible promoter of* carQRS; P^{crtI}*, promoter of* crtI*; p15A* ori*, origin of replication from plasmid p15A;P1* inc*Δ, P1* inc *region with a deletion between* Hpa*I and* Kpn*I; pBR322* ori*, origin of replication from plasmid pBR322.*

7.1.6. *E. coli* strains used.

MC1061	-*hsdR*, *mcrB*, *araD139*, Δ(*araABC-leu*)7679, *galU*, *galK*, *rpsL*, *thi*, Δ*lacX74(lacIPOZY)*.
BL21(DE3)	-F *ompT* [*lon*] $hsdS_B$ with DE3, a λ prophage carrying the T7 RNA polymerase gene.
DH5α	-Φ80d*lacZΔM15*, Δ*lacU169*, *recA1*, *endA1*, *hsdR17*, *supE44*, *thi-1*, *gyrA*, *relA1*.

7.1.7. *M. xanthus* strains used.

The strains used during the course of this study are detailed in Table 2.1. For each strain the genotype, cartenogenic phenotype, derivation and source are given.

Strain	Car Genotype	Car phenotype	Derivation	Source
DK101	wild-type	Car^+	spontaneous mutation of strain FB (DK100)	D.Kaiser.
DK1050	wild-type	Car^+	DK100 (sp.)	Ruiz-Vasquez and Murillo, 1984
DK406	*carR4*	Car^c	DK1050 (NTG)	Martinez -Laborda *et al.*, 1986.
DK717	*carA1*	Car^C	DK101 (UV)	Hodgson, 1993.
DK718	*carR5*	Car^c	DK101 (UV)	Martinez-Laborda *et al.*, 1986.
MR136	wild-type	Car^+	Mx8-mediated transduction of Km^R from DK2725 into DK1050	Martinez-Laborda *et al.*, 1986.
MR461	*carC1*	Car^-	Tc^R replacement of ΩMR403::Tn5-lac	Balsalobre *et al.*, 1987.
UWM303	*carQ2*	Car^-	Gene replacement of DK101 with pSJM122	McGowan *et al.*, 1993.
UWM501	*carC1, carA2.*	Car^-	Mx8-mediated transduction of Tc^R from DK461 into DK717	This study
UWM502	*carC1.*	Car^-	Mx8-mediated transduction of Tc^R from DK461 into DK101	This study
UWM503	*carC1*	Car^-	Mx8-mediated transduction of Tc^R from DK461 into DK1050	This study
UWM504	*carQ2*	Car^-	Mx8-mediated transduction of Km^R from DK101::pAEB603 into UWM303	This study

Table 7-1 M. xanthus *strains used.*

7.2. Basic Techniques.

7.2.1. Restriction endonuclease digestion of DNA.

DNA was digested with restriction enzymes using the buffers and conditions specified by the manufacturers. Digestion was carried out in a volume between 10µl and 30µl using 10 to 30 units of enzyme typically for 2 hours at 37^0C.

7.2.2. Reaction of Alkaline phosphatase with DNA.

DNA was 5'-dephosphorylated by incubation with calf intestine alkaline phosphatase using buffers and conditions as specified by the manufacturers.

7.2.3. Reaction of T4 DNA Kinase with DNA.

DNA was 5'-phosphorylated by addition of T4 DNA kinase and ATP, with conditions and buffers as suggested by the manufacturer.

7.2.4. Gel electrophoresis of DNA.

Gel electrophoresis was used to visualise or isolate DNA fragments. Agarose gels were typically between 0.5 and 1.4% and contained 1xTBE with 0.5µg/ml ethidium bromide. A long-wavelength transilluminator was used to visualise gels and isolate bands, whilst a short-wavelength transilluminator was employed to photograph gels with a Polaroid camera.

7.2.4.1. Materials needed for gel electrophoresis.

10 xTBE.	Loading buffer.
108g Tris base	50% glycerol
55g boric acid	0.5% xylene cyanol FF
9.3g EDTA per litre.	0.5% bromophenol blue.

7.2.5. Transformation of *E. coli* and *M. xanthus*.

7.2.5.1. Transformation of *E. coli* using TSS transformation.

TSS transformation was carried out as specified by Chung *et al.* (1989).

7.2.5.2. Transformation of *E. coli* using CaCl transformation.

The method used for CaCl transformation is described in Sambrook *et al.* (1989).

7.2.5.3. Transformation of *M. xanthus* by electroporation.

Electroporation of plasmid DNA into *M. xanthus* was performed according to the method of Kashefi and Hartzell (1995) using a Bio-Rad Gene Pulser electroporator.

7.2.5.3.1. Preparation of electrocompetent cells.

Exponentially growing cells were washed three times in one volume water and resuspended into 0.01 volumes of water. Manipulations were performed at 25^0C.

7.2.5.3.2. Electroporation conditions.

40µl of cells was added to 2µl of DNA in a 0.2cm electroporation cuvette. Conditions for electroporation were set at a capacitance of 25µF, a voltage of 0.65kV, a resistance of 400Ω and time constants in the range 0.5-0.9ms. Cells were immediately introduced into rich DCY media and after a period of 4 hours incubation at 33^0C, were overlayed in DCY soft agar onto selective DCY plates. Electroporants gave visible colonies after four days' incubation at 33^0C.

7.2.6. Preparation of plasmid DNA from *E. coli.*

The small scale preparation of plasmid DNA was carried out using the alkaline lysis QIAprep Spin Plasmid Kit (QIAGEN), as specified by the manufacturers.

7.2.7. Preparation of *M. xanthus* chromosomal DNA.

The method is based on that described by Hodgson (1993). An overnight culture was used to inoculate 100ml DCY containing appropriate antibiotic. This culture was incubated until stationary phase was reached. Cells were pelleted by centrifugation at 4°C, 16,000g for 15 minutes and resuspended in 10ml STE. 1mg Proteinase K and 2ml 0.5M EDTA were added, followed by 1.5ml 10% N-Lauryl sarcosine to induce lysis. Tubes were incubated at 60°C overnight.

28.5g of CsCl was then dissolved in the mixture and the volume made up to 38ml with water. 2ml 5mg/ml ethidium bromide was added and the sample was then incubated on ice for 2 hours. Cellular debris was pelleted by centrifugation at 4°C, 20,000g for 15 minutes and the supernatant decanted through muslin into a Beckman 'Quick-Seal' centrifuge tube. The tubes were topped up with 71.25% CsCl, balanced with paraffin oil and heat sealed. Centrifugation was performed at room temperature for 20 hours at 220,000g using a Beckman L8 ultracentrifuge (VTi50) rotor. The band of chromosomal DNA was removed using a 19 gauge needle and syringe. Successive washes with 5ml salt-saturated isopropanol removed ethidium bromide from the samples which were then made up to 12ml with water and the DNA precipitated by the addition of 24ml ethanol. After incubating at -20°C overnight, the DNA was pelleted by centrifuging at 4°C, 20,000g for 15 minutes and washing with 70% ethanol. The DNA was finally resuspended in 1ml of TE.

7.2.7.1. Solutions required for preparation of Chromosomal DNA.

STE	NTE
25% Sucrose	5M NaCl
50mM Tris-HCl pH 8.0	10mM Tris-HCl pH 8.0
5mM EDTA	1mM EDTA
TE	**Salt-saturated isopropanol**
10mM Tris-HCl pH 8.0	400ml isopropanol
5mM EDTA	200ml NTE

7.2.8. Preparation of P1 stock solution.

Stock solutions of P1 bacteriophage were prepared from single plaques of P1 grown on a lawn of MC1061 on LC agar. Plaques were picked into TM buffer and plated with MC1061 onto LGC agar to give a lawn of *E. coli*, with confluently lysed P1 plaques. Plates exhibiting confluent lysis were soaked with 6ml TM buffer and the resulting 'soak-out' stock solution of P1 stored above chloroform.

7.2.8.1. Materials needed for the preparation of P1 stock solutions.

LC medium.

As LB media (7.1.1.2)
+5mM CaCl

LGC medium.

As LC media
+2ml 50% Glucose

TM Buffer

10mM Tris-HCl pH8.0
8mM $MgSO_4$

7.2.9. P1 Packaging of plasmids for transduction of *M. xanthus*.

The P1 packaging of plasmids carried by *E. coli* strain MC1061 and the transduction of *M. xanthus* using P1 lysates was carried out as detailed in Hodgson (1993).

7.2.10. Mx8-mediated transductions.

Generalised transductions between strains of *M. xanthus* using bacteriophage Mx8 were performed as described in Hodgson (1993).

7.3. Physiological Studies.

7.3.1. Assays for the activity of β-galactosidase in strains of *M. xanthus*.

The protocol for assaying β-galactosidase activity in *M. xanthus* strains was performed according to the method detailed in Hodgson (1993).

7.3.2. Media replacement during assays of β-galactosidase activity.

Assays which required introduction of the cells into a medium lacking a carbon-source were performed as above (7.3.1). At time = 3 hours the cultures incubating in the light and in the dark were split into two 100ml daughter cultures. One of each pair of daughter cultures were subjected to centrifugation for 15 minutes at 16,000g at 20^0C and the cell pellets resuspended in MC7 buffer after washing with

100ml MC7 buffer. Sampling and enzyme activity determinations were continued as normal for a typical β-galactosidase assay.

7.3.2.1. Materials required for Media replacement β-galactosidase assay.

MC7 Buffer.

10mM MOPS
1mM $CaCl_2$
pH to 7.0

7.3.3. Developmental Assay.

Assays for starvation-induced fruiting body formation were performed as described in Yang and Kaplan (1997). A culture of *M. xanthus* in a late exponential phase of growth was centrifuged and the pelleted cells resuspended in 1/10 volume of TPM. Several 20µl drops of this solution were spotted onto TPM agar plates and incubated at 33^0C in the dark. Fruiting bodies typically formed after 72 hours.

7.3.3.1. Materials needed for Developmental Assays.

TPM ($1dm^3$)

10mM Tris-HCl pH 7.5
1mM KH_2PO_4
8mM $MgSO_4$
15g agar (TPM Agar)

7.4. Molecular Studies.

7.4.1. SDS PAGE gel electrophoresis.

7.4.1.1. SDS PAGE electrophoresis using the Tris-glycine system.

Tris-glycine SDS PAGE electrophoresis was carried out as detailed in Silhavy *et al.* (1984). Vertical SDS PAGE slab gels were run using a Bio-Rad Protean gel tank and two different percentages of separating gel were routinely used (7% and 12%). Their composition is indicated below:

	Acrylamide concentration	
	7%	12%
4x lower buffer	10ml	10ml
Acrylamide stock	9.3ml	16ml
Distilled water	20.5ml	13.8ml
10% ammonium persulfate	0.2ml	0.2ml

The stacking gel was poured on directly onto the top of the polymerised resolving gel. Its composition is defined below:

Stacking gel mix.

2.5ml 4x upper buffer
1ml acrylamide stock
6.4ml distilled water
0.2ml ammonium persulfate.

Protein samples were mixed with an equal volume of cracking buffer, boiled for 5 minutes and loaded onto the gel. Electrophoresis was carried out using 1x running buffer at 30-40mA constant current for 3-4 hours. Alternatively, gels were run overnight at 6mA.

7.4.1.2. Materials required for Tris-glycine SDS PAGE system.

Acrylamide stock (30%).

300g acrylamide
8g bis acrylamide
final volume of 1 litre.

4x upper buffer.

60.6g Tris base
40ml 10%(w/v) SDS
final volume of 1 litre.
Adjust to pH 6.8 with HCl
and add 1ml of TEMED.

4x lower buffer.

181.7g Tris base
40ml 10% (w/v) SDS
final volume of 1 litre.
Adjust to pH 8.8 with HCl
and add 1ml TEMED.

Running buffer.

24g Tris base
57.6g glycine.
0.66M sucrose
6% (w/v)SDS
final volume of 4 litres.

Cracking buffer.

133mM Tris-HCl pH8.8
3.3mM EDTA
40ml 10%(w/v) SDS
0.1M DTT
1.6% (v/v) 2-mercaptoethanol
0.01% bromophenol blue.

7.4.1.3. Coomassie blue R staining of SDS PAGE gels.

After electrophoresis, SDS PAGE gels were immersed in Coomassie blue R stain for 1-2 hours with shaking. Excess stain was removed by incubating gels in destain solution. This was replaced every 4 hours until there was little or no background staining within the gel.

7.4.1.4. Materials needed for Coomassie blue R staining.

Coomassie blue R stain.

90ml methanol
20ml acetic acid
90ml distilled water
0.45g Coomassie blue R.

Destain

400ml methanol
70ml acetic acid per litre.

7.4.2. Expression of recombinant CarQ using T7 polymerase.

Protein expression was also carried out using a T7 RNA polymerase expression system in the strain BL21(DE3) (Rosenberg *et al.*, 1987). Overnight cultures, containing the appropriate plasmid, were used to inoculate LB broth plus antibiotic at a dilution of 1:100. Cultures were incubated at 37°C with shaking until OD_{600} = 0.8, after which protein expression was induced by the addition of IPTG (1mM final concentration) for 4 to 5 hours.

7.4.3. Refolding of insoluble recombinant CarQ.

The expression of protein in 100ml *E. coli* cultures was carried out for 4 hours as detailed in section 2.7.4. The isolation of inclusion bound protein, its resolubilization using Sarkosyl and refolding was carried out as detailed by Nguyen *et al.* (1993). Cells were isolated by centrifugation at 4000 rpm in a MSE Hi-Spin 21 centrifuge at 4°C for 10 minutes and resuspended in 20ml of lysis buffer with 0.2% sodium deoxycholate and 200µg/ml lysozyme. This was kept on ice for 30 minutes and then sonicated using a Jencons sonicator with a 19mm diameter probe for 3x30s bursts (6 microns peak to peak) with 30s cooling on ice in between pulses. Inclusion bodies were isolated by centrifugation at 18000rpm in an MSE Hi-Spin 21 centrifuge at 4°C for 30 minutes. The pellet was resuspended in 25ml of TGED buffer containing 50mM NaCl and 2% sodium deoxycholate. This was incubated on ice for 1 hour with stirring and the inclusion bodies were collected by centrifugation as before. The pellet was again resuspended in 25mls of TGED with 50mM NaCl and 2% sodium deoxycholate and incubated on ice for 1 hour with stirring. Inclusion bodies were collected by centrifugation at 18000rpm and resuspended in 25mls of TGED containing 50mM NaCl and 0.25% Sarkosyl. After 1 hour incubation on ice with stirring, the sample was centrifuged at 18000 rpm in an MSE Hi-Spin 21 centrifuge at 4°C for 30 minutes. The supernatant was dialysed overnight at 4°C against 2 litres of HGED

buffer, which was changed twice and then dialysed overnight at 4°C against 1 litre of storage buffer. Samples were stored at -20°C until required.

7.4.3.1. Materials required for refolding of insoluble recombinant CarQ.

Lysis buffer.

50mM Tris-HCl (pH 7.9)
10mM EDTA
50mM NaCl
1mM DTT
1mM PMSF.
5% glycerol.

TGED buffer

50mM Tris-HCl (pH 7.9)
0.1mM EDTA
0.1mM DTT
5% glycerol

HGED buffer.

50mM sodium HEPES (pH 7.8.)
0.1mM EDTA
0.1mM DTT
5% glycerol.

Storage buffer.

50mM Na HEPES (pH 7.8)
0.1mM EDTA
0.1mM DTT
50% glycerol
50mM NaCl.

7.4.4. Gel retardation assays.

Assays were performed according to a method based on that of Marshak *et al.* (1996).

7.4.4.1. Preparation of DNA probe

Probe fragments of DNA were made by PCR amplification from strains of *M. xanthus*. The PCR products used were initially 5'-labelled with ^{32}P from ^{32}P-γATP by the action of T4 polynucleotide kinase according to supplier's specifications. Polynucleotide kinase was then removed by phenol extraction (Sambrook *et al.*, 1989).

7.4.4.2. Binding reactions.

Binding reactions were performed in a total volume of 10µl with incubation at 37°C for 10 minutes prior to loading on a 6% non-denaturing polyacrylamide gel. Contents of the binding reactions are described below. Electrophoresis was

performed at 110V for 20 hours before the gel was dried onto filter paper and visualised by autoradiography.

Binding Reaction

1µl 10x shift buffer
2µl glycerol
2µg pBR322
5µl Protein preparation
1µl labelled DNA probe

7.4.4.3. Materials required for gel retardation assays.

10x Shift Buffer (for 10 reactions).

0.1M HEPES
0.6M KCl
40mM $MgCl_2$
1mM EDTA
1mg/ml BSA
2.5mM DTT

7.4.5. Pre-binding *in vitro* transcription run-off assays.

7.4.5.1. The *in vitro* transcription run-off assay.

Template DNA containing P^{carQRS} or the Alt promoters were prepared by PCR from the pAEB6XX series of plasmids using primers 1 and 2 (Section 7.4.6.2). CarQ protein was prepared as in sections 7.4.2 and 7.4.3.

0.5pmol of DNA template was mixed with 25µl of reaction mix and the volume made up to 26µl. This was incubated at 30°C for 2 minutes, after which 0.5µl (2 units) of *E. coli* core RNA polymerase (Epicentre Technologies) and purified sigma factor were added. (When using purified CarQ protein, 7µl of sample was added.) The reaction mixture was incubated for 5 minutes at 30°C and 1µl (10µCi) of [α-^{32}P]rCTP (Amersham, UK) and 5µl of RTP mix was added. This was incubated for 2 minutes at 30°C. After the addition of 1µl of 5mg/ml heparin, the reaction was placed at 30°C for 5 minutes. 4µl of 5mg/ml CTP mix was added and incubated for 10 minutes at 30°C. Reactions were stopped by placing mixtures

on ice and RNA products were precipitated with the addition of 10µl of precipitation mix and 60µl of isopropanol. Samples were left on ice for at least 15 minutes and then centrifuged at maximum speed using a bench top microcentrifuge for 30 minutes. The pellet was resuspended in 10µl of sequencing 'stop' solution (Sequenase ^{TM} Version 2.0 DNA Sequencing Kit (United States Biochemical)). Samples were stored at -20°C until required.

2-3µl of RNA sample was incubated at 90°C for 2 minutes before electrophoresis was carried out using a 6% acrylamide, 8M urea, 1xTBE sequencing gel at 40W constant voltage for 2.5 hours. RNA products were detected by autoradiography using intensifier screens.

Molecular weight markers were made by end-labelling 1Kb ladder (Pharmacia, UK) by incubating with ^{32}P-γ-dATP and T4 DNA kinase in 1x exchange buffer (see manufacturers instructions).

7.4.5.2. Solutions required for transcription run-off experiments

2x master mix.

Tris base	1.21 g
EDTA	9.3 mg
$MgCl_2.6H_2O$	487.2 mg

This mixture was made up to 35ml with water and adjusted to pH 7.9 with HCl. To this 25ml of 4mM potassium phosphate (pH 7.5) was added and adjusted to pH 7.9. This was made up to a final volume of 70ml, filter sterilised and stored in aliquots at -20°C.

Reaction mix (for 10 reactions).

2x master mix	115µl
BSA (RNase-free, 2.9mg/ml)	6µl
100mM DTT	6µl
100mM EDTA	2µl
glycerol	80µl
water	41µl

RTP mix.

RTP mix is an equal mixture of:

5mg/ml rATP (pH7.0)
5mg/ml rGTP (pH7.0)
5mg/ml rUTP (pH7.0)

Precipitation mix.

tRNA (10mg/ml)	10µl
3M sodium acetate	53µl
water	17µl.

7.4.6. PCR of DNA from *M. xanthus* and *E. coli* strains

7.4.6.1. PCR amplification of DNA.

PCRs were performed in a 50 µl mixture containing 1x *Taq* polymerase buffer, 4mM $MgCl_2$, 25µM dNTPs, and 12.5 pmoles of each primer. Template DNA was either 1µl of a 100x dilution of a previous PCR product, 1µl of a 10x dilution of a plasmid preparation, or 5µl of a 100µl solution containing a single colony picked from plate and resuspended in water. Reactions were performed in a Perkin-Elmer thermal cycler according to two programs. Program one was used for amplification between primers 1 and 2, whereas program two was used as the basis for all other reactions.

Program one involved a hot-start drop-down approach with annealing temperatures reducing from 72°C to 62°C at a rate of 0.5°C per cycle, followed by 10 cycles with an annealing temperature of 62°C. Each cycle consisted of a 15 second melting step at 96°C, a 15 second annealing step and finally a 30 second extension step at 72°C. After the final cycle a five minute extension step at 72°C was performed.

Program two was also hot-start with 30 cycles and final extension step as described above but with a constant annealing temperature of 58°C.

7.4.6.2. Primers.

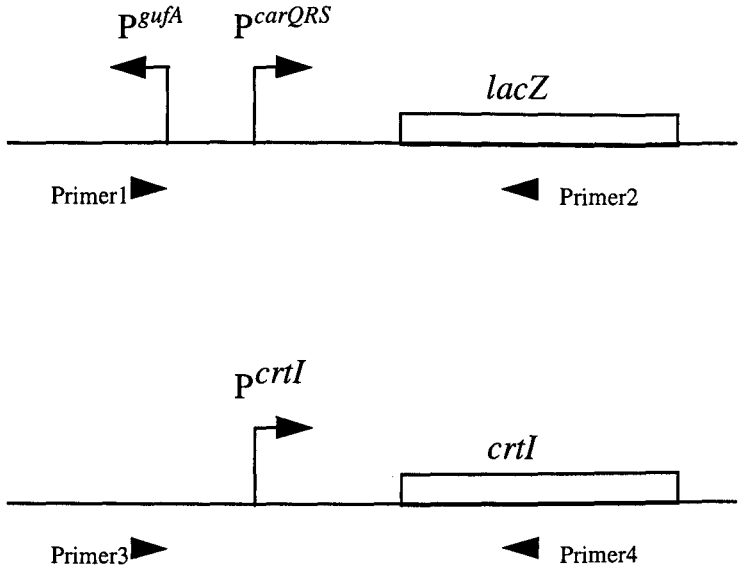
Primers were obtained from GIBCO (UK) and stored at -20°C as a 100 pmol/ml solution.

7.4.6.2.1. Primers for amplification of DNA containing promoters from *M. xanthus*.

Primer 1	GCGTCCGAGGTGCCTCCG
Primer 2	CGATCGGTGCGGGCCTCT
Primer 3	CCAGGACCACGCAGTAGG
Primer 4	CTCCCTCGCTTGCCATGC

Pairs of primers used to amplify pieces of DNA containing promoters of interest are shown below with a description of the promoters located on the PCR products and the size of the predicted products.

Primer 1 - Primer 2	480bp Product.	P^{carQRS}, P^{gufA}
Primer 3 - Primer 4	231bp Product.	P^{crtI}



The annealing site for primer 1 lies at the 5'-end of the minimum *carQRS* promoter around position -170, while the annealing site for primer 2 lies on the opposite strand within *lacZ* such that amplification of template DNA will only occur on a template that contains P^{carQRS} upstream of a copy of the *lacZ* gene (for instance any of the pAEB6XX constructs).

7.4.6.2.2. Primers used for the amplification of DNA to determine multiple insertions of pAEB6XX promoter constructs.

Primers were designed to be complementary to sites within the *attP* region as shown in Figure 7-1.

Primer MULTI1	CGGACTCACGATGCTCGC
Primer MULTI2	GACTCCTCTGGCTGGGTG
Primer MULTI3	CAGCTTGGCACTCGGCAG

PCR reactions were performed as described above but including 12.5 pmol of each of MULTI2 and MULTI3 and with 25 pmol of MULTI1, with a final concentration of 10% DMSO to eliminate problems due to secondary structure formation..

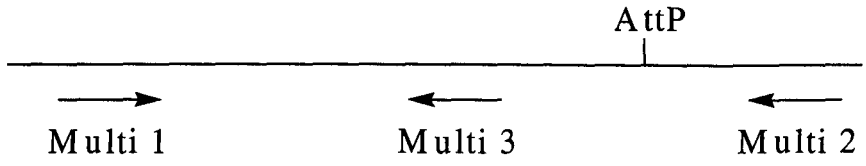


Figure 7-1 The Mx8 attP *region and annealing sites for primers.*

Amplification between primers MULTI1 and MULTI2 is expected to give a product of 694bp if the template strain contains an intact *attP* region, due to multiple insertions of plasmids carrying *attP* into the $attB^{Mx8}$ site. Amplification between MULTI1 and MULTI3 is expected to give a product of 213bp from all *M. xanthus* strains. Lack of a 694bp PCR product does not prove the non-existance of multiple insertions since it is possible that that PCR reaction failed for some reason. However, if the strain under test gives a product between primers MULTI1 and MULTI3 in the same reaction, it is almost certain that primers MULTI1 and MULTI2 aren't giving a product because there aren't multiply inserted plasmids. In this way, amplification between MULTI1 and MULTI3 serves as an internal positive control to each PCR product.

7.4.7. Sequencing of *M. xanthus* DNA

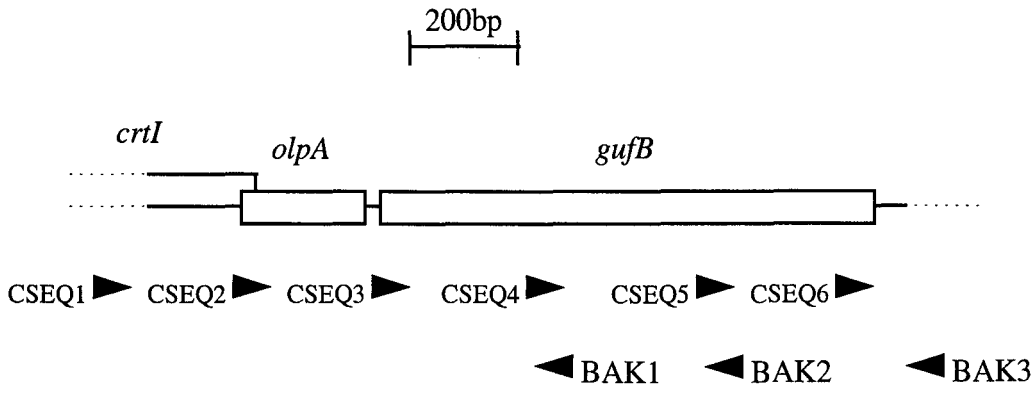
Plasmids containing *M. xanthus* DNA for sequencing were cloned into *E. coli* strain DH5α. Template DNA was then prepared as described in 7.2.6.

7.4.7.1. Sequencing primers for the *crtI* region.

Primers for sequencing were obtained from GIBCO (UK) and stored at -20°C as a 100 pmol/ml solution. Primer sequences are shown below along with a schematic representation of the positions of the primer annealling sites within the *crtI* region.

CSEQ1	GACTACCTGACGCGCGAG
CSEQ2	GCTGGAGCCGGAGTCGCC
CSEQ3	CGGTGGGTGGCCTGGTGC
CSEQ4	GGCACGTCCAGGAAGCCG
CSEQ5	GGGACGCCGCTGTACGCA
CSEQ6	GTCTGACGCCCATGCTCG

BAK1	CGGCTTCCTGGACGTGCC
BAK2	TGCGTACAGCGGCGTCCC
BAK3	CGAGCATGGGCGTCAGAC
BAK4	CACCACGCAGCAGGCAGC



The *crtI* region

7.4.7.2. Sequencing reactions

Sequencing reactions were performed using the method of Sanger *et al.* (1977). 1.2 µg of template DNA was added to 3.2 pmol of primer and made up to 6µl with water. *Taq* cycle sequencing was then performed using an Applied Biosystems 373A DNA sequencer (Alta Bioscience, Birmingham, UK).

8. Appendices

8.1. Appendix 1: Inhibition of Development by Light.

Light has been shown to have two effects on cells of *Myxococcus xanthus*. The first response is the initiation of carotenogenesis described in the bulk of this work. The second effect is the prevention of starvation-induced development. In *M. xanthus*, starvation induces aggregation of large numbers of cells which form a fruiting body containing cells, a subset of which then differentiate into myxospores - cells relatively resistant to environmental stresses (Sudo and Dworkin, 1968). Disrupting Tn*5-lac* insertions have allowed isolation and characterisation of a large number of developmentally expressed genes. Initial events seem dependent on a starvation signal involving (p)ppGpp and a cell-density dependent signal mediated through the CsgA protein (Harris *et al.,* 1998). Together, these signals induce expression of genes necessary for A-signal production, and activate the *frz* genes responsible for co-ordinated cell motions and aggregation. Subsequent steps generate and respond to the C, D and E signals and lead to the completion of fruiting body formation (for a recent review see Dworkin, 1996).

Previous work (Shimkets, L., pers. comm.) has shown that light inhibits the normal process of fruiting body formation of *M. xanthus* and that the developmental block occurs early in the developmental cascade. In all other myxobacteria, light is absolutely required for fruiting body formation. In *Stigmatella aurantica*, light photosensitises starving cells to a pheromone which stimulates fruiting body formation (Inouye *et al.,* 1980 and Stephens *et al.,* 1982). The possibility that this block in fruiting body formation is a consequense of gene expression within the Car regulon has never been systematically investigated, although the existance of a link between carotenogenesis and development is clear, since fruiting bodies contain high levels of carotenoids.

Another suggested line of evidence implicating a link between carotenogenesis and development is the existance of CarD which is required both for fruiting body

formation and for carotenogenesis. However, CarD is most probably only a requirement for both processes since it is a general DNA-binding protein and has a non-specific mode of action.

D. A. Hodgson (pers. comm.) has shown that the inhibition of development is not due directly to the production of carotenoids themselves, by showing that a Car^c *carR* mutant would fruit in the dark but not in the light. This also suggests that the wavelength of light causing the inhibitory effect is probably not blue light. To determine whether the inhibition of development by light is a phenomenon which depends on products of the carotenogenic regulon, the wild-type (DK101) and a CarQ mutant (UWM303) were assessed for their ability to form fruiting bodies in the light and dark. If light-mediated inhibition of development was due to a product of the Car regulon (or absence of CarR), it would be expected that UWM303 would be able to develop normally in the light, unlike DK101. This would be the case if the required product was a gene product such as CarS or CrtI, or whether it is a carotenoid produced by the regulon's structural genes.

This was shown not to be the case. Figure 8-1 displays the results of a standard developmental assay for both strains, in the dark and in the light. Both strains develop normally in the dark with multiple fruiting bodies forming from each colony, but neither strain forms any fruiting bodies under illumination.

PAGE NUMBERING AS ORIGINAL

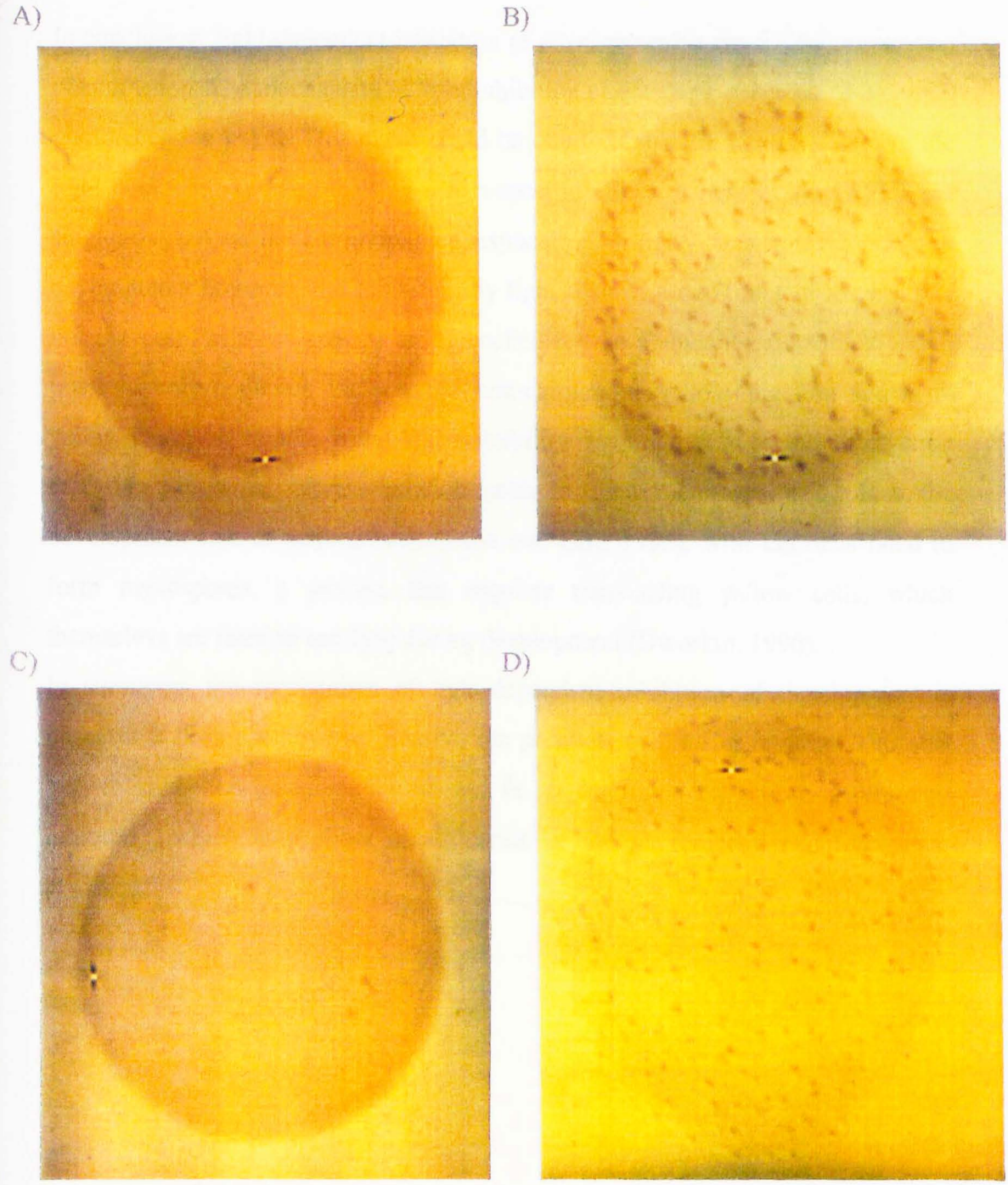


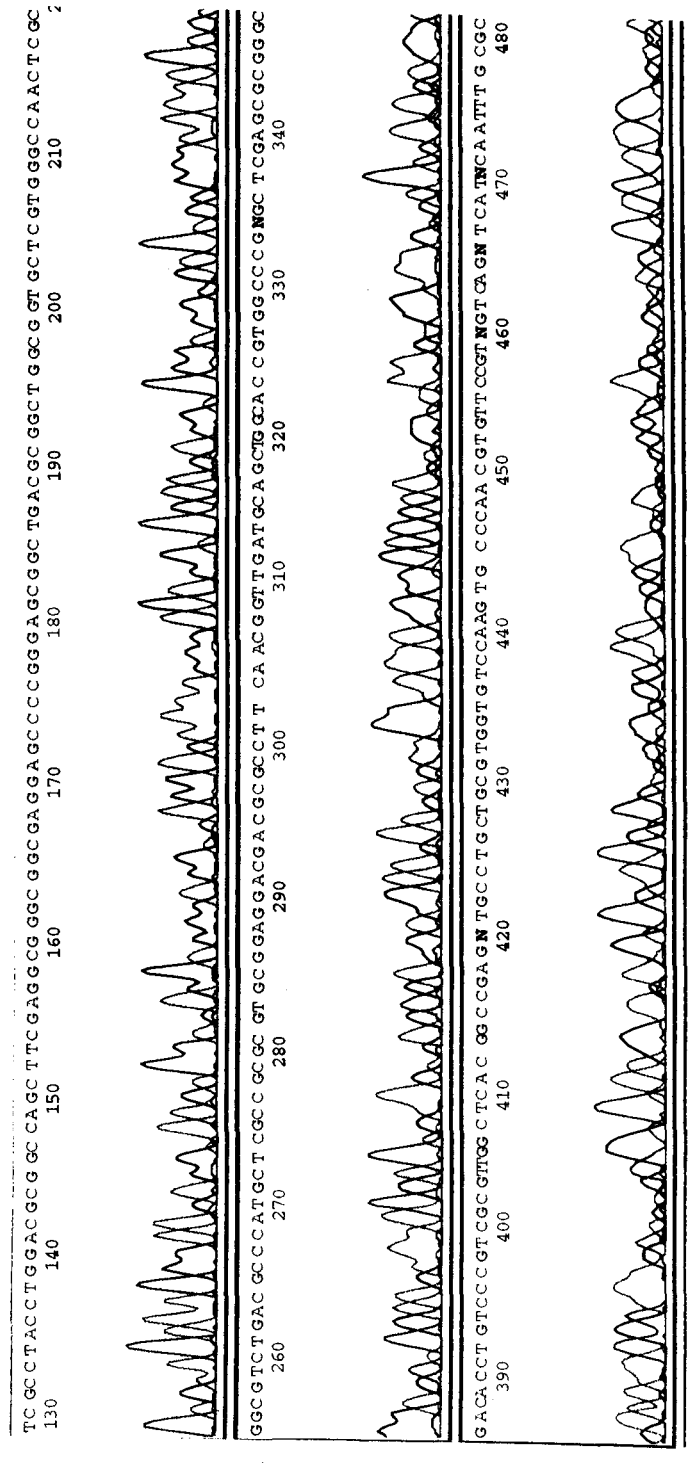
Figure 8-1 Development of wild-type and Car⁻ strains in light and dark.

A): DK101 (wild-type) under illumination. B): DK101 in the dark.
C): UWM303 (Car⁻) under illumination. D): UWM303 in the dark.

In conclusion, light-dependent inhibition of development is not a CarQ-dependent phenomenon. The mechanism of light inhibition of development requires an initial photoreception event. This event could be due to a specific chromophore, or the light could be absorbed causing a non-specific effect. It may be that PPIX and singlet oxygen are the chromophores responsible for light-dependent inhibition of development however, the inhibition by light seen in a *carR* mutant implies this not the case, as these species are quenched by the abundant carotenoids. Other possible chromophores include the non-carotenoid yellow pigment found in yellow phase *M. xanthus* cells. This possibility is strengthened by the observation that the yellow/tan phase variation seen in mixed cultures determines the development fate of starved cells (Laue and Gill, 1995), with tan cells fated to form myxospores, a process that requires trans-acting yellow cells, which themselevs are fated to autolyse during development (Dworkin, 1996).

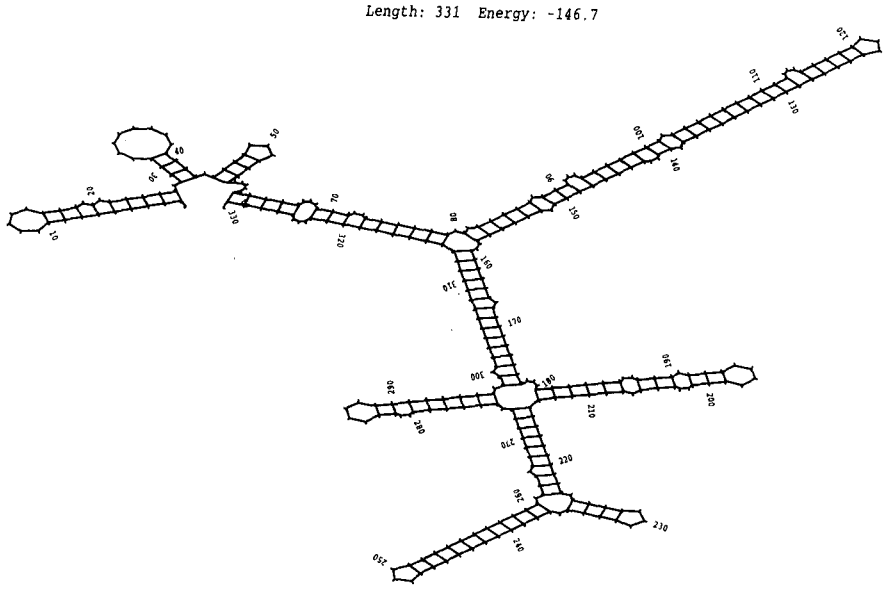
In summary, the mechanism of light-dependent inhibition of development is obscure. It does not however involve the products of the Car regulon. The next step to study this phenomenon should be to determine an action spectrum of inhibition, which may allow identification of the photoreceptive chromophore involved.

8.2. Appendix 2: Example of Sequence data.

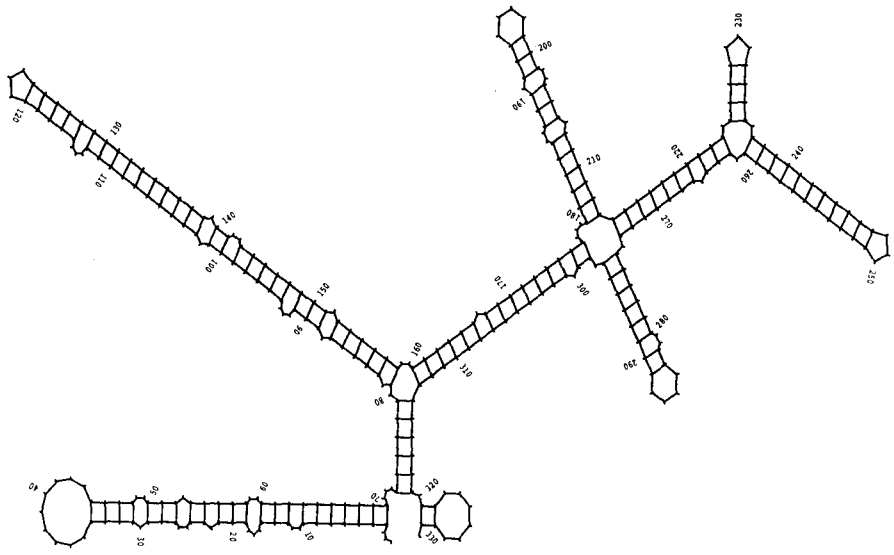


8.3. Appendix 3: Examples of predicted secondary structure around *olpA*.

Squiggle plot of: crtigufb.mfold February 10, 1999 11:53
(Linear) MFOLD of: crtigufb.gcg T: 37.0 Check: 1489 from: 1 to: 331 February 10, 1999 11:49
Length: 331 Energy: -146.7



Squiggle plot of: crtigufb.mfold February 10, 1999 11:53
(Linear) MFOLD of: crtigufb.gcg T: 37.0 Check: 1489 from: 1 to: 331 February 10, 1999 11:49
Length: 331 Energy: -146.2



8.4. Appendix 4: LP405 is Car^C and does not contain a *carR* lesion.

Strain LP405 was recently created by Linda Plamann's group by UV mutagenesis of a Car$^+$ A-signalling mutant, during an attempt to find supressors of an *asgA* mutant (Kessler and Plamann, 1998). It carries a 'suppressor of A-signalling' (*sas*) mutation and is Car^c. The double phenotype introduced by UV-mutagenesis in LP405 is probably a consequence of two independent mutations, not the result of a single pleiotropic mutation and could therefore provide another *carR* or *carA* mutant allele. Alternatively, the mutation could lie at an as yet unidentified locus encoing a negative regulator of carotenogenesis.

In order to determine whether the Car^c mutation within LP405 lay at *carR*, a wild-type *carR* allele which is linked to a Tn*5* insertion (Martinez-Laborda *et al.*, 1986) was transduced from MR136 into LP405 by Mx4-LA27-mediated generalised transduction. Around 100 transductants were isolated and each one maintained a Car^c phenotype. Had LP405 contained a lesion in *carR*, around 80% of the transductants would have been expected to acquire a Car$^+$ phenotype due to the close proximity of *carR* and Tn*5* in MR136. This enabled the conclusion that the mutation causing the Car^c phenotype in LP405 does not lie within *carR*.

A suitable plasmid containing wild-type *carA* has recently been provided by F. J. Murillo (pMAR100, Martinez-Laborda *et al.*, 1989) which has been shown to allow complementation of a *carA* mutant. This plasmid will be introduced into LP405 to determine whether the Car^c mutation within LP405 lies within *carA* or at a novel locus.

9. References

Ahern, H. and Cunningham, R. P. (1995) Genetic regulation of antioxidant defenses in *Escherichia coli* and *Salmonella typhimurium*. In: Oxidative Stress and Antioxidant Defenses in Biology. Ahmad, S. (Ed.) Chapman and Hall, (New York).

Ahmad, S. (1995) Antioxidant mechanisms of enzymes and proteins. In: Oxidative Stress and Antioxidant Defenses in Biology. Ahmad, S. (Ed.) Chapman and Hall, (New York).

Albertini, A. M., Caramori, T., Crabb, W. D., Scoffone, F. and Galizzi, A. (1991) The *flaA* locus of *Bacillus subtilis* is part of a large operon coding for flagellar structures, motility functions and an ATPase-like polypeptide. Journal of Bacteriology **173** 3573-3579.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990) Basic local alignment search tool. Journal of Molecular Biology **215** 403-410.

Altschul, S. F., Madden, T. L., Schaffer, J. Y., Zhang, Z., Miller, W. and Lipman, D. J. (1997) Gaed BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research **25** 3389-3402.

Apelian, D. and Inouye, S. (1990) Development -specific σ-factor essential for late-stage differentiation of *Myxococcus xanthus*. Genes and Development **4** 1396-1403.

Apelian, D. and Inouye, S. (1993) A new putative sigma factor of *Myxococcus xanthus*. Journal of Bacteriology **175** 3335-3342.

Armstrong, G. A. (1997) Genetics of eubacterial carotenoid biosynthesis. Annual Review of Microbiology **51** 629-659.

Armstrong, G. A., Alberti, M., Leach, F. and Hearst, J. E. (1989) Nucleotide sequence, organisation and nature of the protein products of the carotenoid biosynthesis gene cluster of *Rhodobacter capsulatus*. Molecular and General Genetics **216** 254-268.

Armstrong, G. A., Alberti, M. and Hearst, J. E. (1990) Conserved enzymes mediate the early reactions of carotenoid biosynthesis in non-photosynthetic prokaryotes. Proceeding of the National Academy of Science (USA) **87** 9975-9979.

Avery, L. and Kaiser, D. (1983) *In situ* replacement and isolation of a spontaneous tandem genetic duplication. Molecular and General Genetics **191** 99-109.

Bairoch, A., Bucher, P. and Hofmann, K. (1997) The PROSITE database, its status in 1997. Nucleic Acid Research **25** 217-221.

Balsalobre, J. M., Ruiz-Vasquez, R. and Murillo, F. J. (1987) Light induction of gene expression in *Myxococcus xanthus*. Proceeding of the National Academy of Science (USA) **84** 2359-2362.

Banghart, L. R., Chamberlin, C. W.,Verlade, J. Korobko, I. V., Ogg, S. L., Jacj, L. J. W., Vakharia, V. N. and Mather, I. H. (1998) Butyrophilin is expressed in mammary epithelial cells from a single-sized messenger RNA as a type I membrane glycoprotein. Journal of Biological Chemistry **273** 4171-4179.

Barne, K. A., Bown, J. A., Busby, S. J. W. and Minchin, S. D. (1997) Region 2.5 of the *Escherichia coli* RNA polymerase s^{70} subunit is resposible for the recognition of the 'extended -10' motif at promoters. The EMBO Journal **16** 4034-4040.

Berry, A. E. (1998) Towards a molecular mechanism for light induction of gene transcription in *Myxococcus xanthus*. PhD thesis. University of Warwick, UK.

Bhriain, N. N., Dorman, C. J. and Higins, C. F. (1989) An overlap between osmotic and anaerobic stress responses: a poential role for DNA supercoiling in the coordinate regulation of gene expression. Molecular Microbiology **3** 933-942.

Bibb, M. J., Findlay, P. R. and Johnson, M. W. (1984) The relationship between base composition and codon usage in bacterial genes and its use for the simple and reliable identification of protein-coding sequences. Gene **30** 157-166.

Biran, D. and Kroos, L. (1997) *In vitro* transcription of *Myxococcus xanthus* genes with RNA polymerase containing σ^A, the major sigma factor in growing cells. Molecular Microbiology **25** 463-472.

Botella, J. A., Murillo, F. J. and Ruiz-Vasquez, R. (1995) A cluster of structural and regulatory genes for light-induced carotenogenesis in *Myxococcus xanthus*. European Journal of Biochemistry **233** 238-248.

Bown, J., Barne, K., Minchin, S. and Busby, S. (1997) Extended -10 promoters. Nucleic Acids Molecular Biology **11** 41-52.

Brandner, J. P. and Kroos, L. (1998) Identification of the Ω4400 regulatory region, a develpomental promoter of *Myxococcus xanthus*. Journal of Bacteriology **180** 1995-2004.

Bremer, H. and Dennis, P. P. (1996) Modulation of chemical composition and other parameters of the cell by growth rate. In: *Escherichia coli* and *Salmonella typhimurium:* Cellular and molecular biology. Neihardt, F. C., Curits, R. I., Ingraham, J. L., Lin, E. C. C., Low, K., Reznikoff, W. S., *et al.,* (Eds.) American Society for Microbiology. Washington D.C.

Browning, D. F. (1997) A molecular investigation of light induced carotenoid Production in *Myxococcus xanthus*. PhD Thesis. University of Warwick, UK.

Buck, M. and Cannon, W. (1992) Specific binding of the transcription factor σ^{54} to promoter DNA. Nature **358** 422-424.

Burchard, R. P., Burchard, A. C. and Parish, J. H. (1977) Pigmentation phenotype instability in *Myxococcus xanthus*. Canadian Journal of Microbiology **23** 1657-1662.

Burchard, R. P., Gordon, S. A. and Dworkin, M. (1966) Action spectrum for the photolysis of *Myxococcus xanthus*. Journal of Bacteriology **91** 896-897.

Burchard, R. P. and Dworkin, M. (1966) Light-induced lysis and carotenogenesis in *Myxococcus xanthus*. Journal of Bacteriology **91** 535-545.

Burchard, R. P. and Hendricks, S. B. (1969) Action spectrum for carotenogenesis in *Myxococcus xanthus*. Journal of Bacteriology. **97** 1165-1168.

Buttner, M. J. (1989) RNA polymerase heterogeneity in *Streptomyces coelicolor* A3(2). Molecular Microbiology **3** 1653-1659.

Buttner, M. J., Fearnley, I. M. and Bibb, M. J. (1987) The agarase gene (*dagA*) of *Streptomyces coelicolor* A3(2): nucleotide sequence and transcriptionl analysis. Molecular and General Genetics **209** 101-109.

Buttner, M. J., Smith, A. M. and Bibb, M. J. (1988) At least three different RNA polymerase holoenzymes direct transcription of the agarase gene (*dagA*) of *Streptomyces coelicolor* A3(2). Cell **52** 599-607.

Carattoli, A., Romano, N., Ballario, P. Morelli, G. and Macino, G. (1991) Carotenoid biosynthetic gene (*albino-3*) reveals highly conserved regions among prenyltransferases. Journal of Biological Chemistry **266** 5854-5859.

Cayuela, M. L. and Murillo, F. J. (1998) New findings on the *M. xanthus* gene *carD* and its product (a high mobility group I(Y)-like protein) Abstracts of the 25th International Conference on the Biology of the Myxobacteria. June 1998. Arahova, Greece.

Cervantes, M. and Murillo, F. J. (1998) Structural and regulatory genes in the *carA* locus, a gene cluster for light-induced carotenogenesis in *M. xanthus*. Abstracts of the 25th International Conference on the Biology of the Myxobacteria. June 1998. Arahova, Greece.

Chen, H. W., Keseler, I. M and Shimkets, L. J. (1991) A physical map of the *Myxococcus xanthus* chromosome. Journal of Bacteriology **173** 4206-4213.

Chung, C. T., Niemela, S. L. and Miller, R. H. (1989) One-step preparation of competent *Escherichia coli*: Transformation and storage of bacterial cells in the same solution. Biochemistry **86** 2172-2175.

Codgell, R. J. and Frank, H. A. (1987) How carotenoids function in photosynthetic bacteria. Biochimica et Biophysica Acta **895** 63-79.

Cole, S. T., Brosch, R., Parkhill, J. and 39 others (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. **393** 537-544.

Craig, M. L., Suh, W. C. and Record, M. T., Jnr. (1995) HO· and DNase I probing of $E\sigma^{70}$ RNA polymerase-lambda PR promoter open complexes: Mg^{2+} binding and its structural consequences at the transcription start site. Biochemistry **34** 15624-15632.

Cunningham, F. X. and Grant, E. (1998) Genes and enzymes of carotenoid biosynthesis in plants. Annual Review of Plant Physiology and Plant Molecular Biology **49** 557-583.

Dailey, T. A. and Dailey, H. A. (1998) Identification of an FAD superfamily containing protoporphyrin oxidases, monoamine oxidases and phytoene dehydrogenase: Expression and characterisation of phytoene desaturase of *Myxococcus xanthus*. The Journal of Biological Chemistry **273** 13658-13662.

Davis, J. M., Mayor, J. and Plamann, L. (1995) A missense mutation in *rpoD* results in an A-signalling defect in *Myxococcus xanthus*. Molecular Microbiology **18** 943-952.

deHaseth, P. L. and Helmann, J. D. (1995) Open complex formation by *Escherichia coli* RNA polymerase: the mechanism of polymerase-induced strand separationof double-helical DNA. Molecular Microbiology **16** 817-824.

deHaseth, P. L., Zupnancic, M. L. and Record, M. T., -Jnr. (1998) RNA polymerase-promoter interactions: the coming and goings of RNA polymerase. Journal of Bacteriology **180** 3019-3025.

Dent, C. L. and Latchman D. S. (1993) The DNA mobility shift assay. In: Latchman, D. S. (Ed.) Transcription Factors: A Practical Aroach. Oxford University Press.

Di Mascio, P., Devasagayam, T. P. A., Kaiser, D. and Sies, H. (1990) Carotenoids, tocopherols and thiols as biological singlet molecular oxygen quenchers. Biochemical Society Transactions **18** 1054-1056.

Dombroski, A. J. (1997) Recognition of the -10 promoter sequence by a partial polypeptide of s^{70} *in vitro*. Journal of Biological Chemistry **272** 3487-3493.

Dombroski, A. J., Walter, W. A., Record Jnr., M. T., Siegele, D. A. and Gross, C. A. (1992) Polypeptides containing highly conserved regions of transcription initiation factor σ^{70} exhibit specificity of binding to promoter DNA. Cell **70** 501-512.

Dombroski, A. J., Walter, W. A. and Gross, C. A. (1993) N-terminal amino acids modulate sigma factor DNA binding activity. Genes and Development **7** 2446-2455.

Downard, J., Ramaswamy, S. V. and Kil, K. -S. (1993) Identification of *esg*, a genetic locus involved in cell-cell signalling during *Myxococcus* development. Journal of Bacteriology **175** 7762-7770.

Dworkin, M. (1959) Function of carotenoids in photosynthetic bacteria. Nature **184** 1891-1892.

Dworkin, M. (1962) Nutritional requirements for vegetative growth of *Myxococcus xanthus*. Journal of Bacteriology **84** 250-257.

Dworkin, M. (1996) Recent advances in the social and developmental biology of the Myxobacteria. Microbiology Reviews **60** 70-102.

Elliott, T. and Geiduschek, E. P. (1984) Defining a bacteriophage T4 late promoter: absence of a '-35' region. Cell **36** 211-219.

Fiel, R. J., Datta-Gupta, N., Mark, E. H. and Howard, J. C. (1981) Induction of DNA damage by porphyrin photosensitisers. Cancer Research **41** 3543-3545.

Fisseha, M., Gloudemans, M., Gill, R. E. and Kroos, L. (1996) Characterisation of the regulatory region of a cell-interaction-dependent gene in *M. xanthus*. Journal of Bacteriology **178** 2539-2550.

Fontes, M., Ruiz-Vasquez, R. and Murillo, F.J. (1993) Growth phase dependance of the activation of a bacterial gene for carotenoid synthesis by blue light. The EMBO Journal **12** (4) 1265-1275.

Foote, C. S. (1991) Definition of type I and type II photosensitised oxidation. Photochemistry and Photobiology **54** p659.

Foster, H. A., Yasouri, F. N. and Daoud, N. N. (1992) Antibiotic activity of soil myxobacteria and its ecological implications. FEMS Microbiology ecology **101** 27-32.

Fridovich, I. (1978) The biology of oxygen radicals. Science **201** 875-880.

Gardella, T., Hoyle, H. and Susskind, M. M. (1989) A mutant *Escherichia coli* s^{70} subunit of RNA polymerase with altered promoter specificity. Journal of Molecular Biology **26** 579-590.

Gatz, C. and Quail, P. H. (1988) Tn*10*-encoded *tet* repressor can regulate an operator-containing plant promoter. Proceedings of the National Academy of Science (USA) **85** 1394-1397.

Geiselmann, J. (1997) The role of DNA conformation in transcriptional initiation and activation in *Escherichia coli*. Biological Chemistry **378** 599-607.

Gill, R. E. and Shimkets, L. J. (1993) Genetic approaches for analysis of Myxobacterial behaviour. In Dworkin, M. and Kaiser, D. (ed.), *Myxobacteria II.* American Society for Microbiology, Washington D.C.

Gorham, H. C., McGowan, S. J., Robson, P. R. H. and Hodgson, D. A. (1996) Light-induced carotenogenesis in *Myxococcus xanthus*: light-dependent membrane sequestration of ECF sigma factor CarQ by anti-sigma factor CarD. Molecular Microbiology **19** 171-186.

Gross, C. A. (1996) Function and regulation of the heat shock proteins. In *Escherichia coli* and *Salmonella typhimurium:* Cellular and molecular biology. Neihardt, F. C., Curits, R. I., Ingraham, J. L., Lin, E. C. C., Low, K., Reznikoff, W. S., *et al.,* (Eds.) Washington D.C. American Society for Microbiology, 1382-1399.

Gross, C. A., Chan, C. L. and Lonetto, M. A. (1996) A structure/function analysis of *Escherichia coli* RNA polymerase. Philosophical Transactions of the Royal Society (London) B **351** 475-482.

Grossman, A. D. (1995) Genetic networks controlling the initiation of sporulation and the development of genetic competence in *Bacillus subtilis*. Annual Review of Genetics **29** 477-508.

Hader, D. -P. and Tevini, M. (1987) General photobiology. Pergamon Press (UK).

Haldenwang, W. G. (1995) The sigma factors of *Bacillus subtilis*. Microbiology Reviews **59** 1-30.

Harris, B. Z., Kaiser, D. and Singer, M. (1998) The guanosine nucleotide (p)ppGpp initiates development and A-factor production in *Myxococcus xanthus*. Genes and Development **12** 1022-1035.

Havaux, M. (1998) Carotenoids as membrane stabilisers in chloroplasts. Trends in Plant Science **3** 147-151.

Helman, J. D. and Chamberlin, M. J. (1988) Structure and function of bacterial sigma factors. Annual Review of Biochemistry **57** 839-872.

Hershberger, C. D., Ye, R. W., Parsek, M. R., Xie, Z. -D. and Chackrabarty, A. M. (1995) The *algT* (*algU*) gene of *Psuedomonas aeruginosa*, a key regulator

involved in alginate biosynthesis, encodes and alternative sigma factor (σ^E). Proceeding sof the National Acadamy of Science (USA) **92** 7941-7945.

Hidalgo, E., Leautaud, V. and Demple, B. (1998) The redox-regulated SoxR protein acts from a sinle DNA site as a repressor and an allosteric activator. EMBO Journal **17** 2629-2636.

Hodgkin, J. and Kaiser, D. (1979) Genetics of gliding mobility in Myxococcus xanthus (Myxobacteriales): genes controlling movement of single cells. Molecular and General Genetics **171** 167-176.

Hodgson, D. A. (1993) Light-induced carotenogenesis in *Myxococcus xanthus*: genetic analysis of the *carR* region. Molecular Microbiology **7** (3) 471-488.

Hodgson, D. A. and Berry, A. E. (1998) Light regulation of carotenoid synthesis in *Myxococcus xanthus*. In Caddick, M. X., Baumberg, S., Hodgson, D. A. and Phillips-Jones, M. K. (Eds) Microbial Responses to light and time. Society for General Microbiology, UK.

Hodgson, D. A. and Murillo, F. J. (1993) Genetics of regulation and pathway of synthesis of carotenoids. In Dworkin, M. and Kaiser, D. (ed.), *Myxobacteria II.* American Society for Microbiology, Washington D.C.

Hood, D. W., Heidstra, R., Swoboda, U. K. and Hodgson, D. A. (1992) Molecular genetic analysis of proline and tryptophan biosynthesis in *Streptomyces coelicolor* A3(2): interaction between primary and secondary metabolism - a review. Gene **115** 5-12.

Houssaini-Iraqui, M., Clavel-Seres, S., Rastogi, N. and David, H. L. (1992) The expression of the *Mycobacterium aurum* carotenogenesis operon is not repressed by the repressor of *Mycobacterium vaccae* photoinducible carotenogenesis. FEMS Microbiology Letters **99** 233-236.

Humayun, M. Z. (1998) SOS and mayday: multiple inducible mutagenic pathways in *Escherichia coli*. Molecular Microbiology **30** 905-910.

Inouye, S. (1990) Cloning and DNA sequence of the gene coding for the major sigma factor from *Myxococcus xanthus*. Journal of Bacteriology **172** 80-85.

Inouye, S., White, D. and Inouye, M. (1980) Development of *Stigmatella aurantiaca*: effects of light on gene expression. Journal of Bacteriology **141** 1360-1365.

Ishihama, A. (1993) Protein-protein communication within the transcription aaratus. Journal of Bacteriology **175** 2483-2489.

Jones, G. H., Paget, M. S. B., Chamberlin, L. and Buttner, M. J. (1997) Sigma-E is required for production of the antibiotic actinomycin in *Streptomyces antibioticus*. Molecular Microbioloy **23** 169-178.

Kammerer, W., Deuschle, U., Gentz, R. and Bujard, H. (1986) Functional dissection of *Escherichia coli* RNA polymerase promoters: information in the transcribed region is involved in late steps of the overall process. EMBO Journal **5** 2995-3000.

Kashefi, K. and Hartzell, P. L. (1995) Genetic suppression and phenotypic masking of a *Myxococcus xanthus* FrzF$^-$ defect. Molecular Microbiology **15** 483-494.

Kato, F., Hino, T., Nakaji, A., Tanaka, M. and Koyama, Y. (1995) Carotenoid synthesis in *Streptomyces setonii* ISP5395 is induced by the gene *crtS* whose product is similar to a sigma factor. Moleular and General Genetics **247** 387-390.

Keseler, I. M. and Kaiser, D. (1995) An early A-signal-dependent gene in *M. xanthus* has a sig-54-like promoter. Journal of Bacteriology 4638-4644.

Keseler, I. M. and Kaiser, D. (1997) Sig-54, a vital protein for *M. xanthus*. Proceedings of the National Academy of Science (USA) **94** 1979-1984.

Kessler, V. and Plamann, L. (1998) Genetic suppression analysis of *asgA* in *M. xanthus*. Abstract: 25th International Conference on the Biology of the Myxobacteria, Arahova, Greece.

Kleinig, H. (1974) Inhibition of carotenoid synthesis in *Myxococcus fulvus* (*Myxobacterales*). Archives of Microbiology **97** 217-226.

Kleinig, H. and Reichenbach, H. (1973) Biosynthesis of carotenoid glucoside esters in *Myxococcus fulvus* (Myxobacterales): Inhibition by nicotine and carotenoid turnover. Biochimica et Biophysica Acta (Amst.) **306** 249-256.

Knox, J. P. and Dodge, A. D. (1985) Singlet oxygen and plants. Photochemistry **24** 889-896.

Kohen, E., Santus, R. and Hirschberg, J. G. (1995) Photobiology. Academic Press. San Diego, California, USA.

Komano, T., Franceschini, T. and Inouye, S. (1987) Identification of a vegetative promoter in *Myxococcus xanthus*. Journal of Molecular Biology **193** 517-524

Konigsberg, W. and Godson, G. N. (1983) Evidence for use of rare codons in the *dnaG* gene and other regulatory genes of *Escherichia coli.* Proceedings of the National Academy of Science (USA) **80** 687-691.

Krummel, B. and Chamberlin, M. J. (1989) RNA chain initiation by *Escherichia coli* RNA polymerase. Structural transitions of the enyme in early ternary complexes. Biochemistry **28** 7829-7842.

Kuhlbrandt, W., Wang, D. N. and Fujiyoshi, Y. (1994) Atomic model of plant light-harvesting complex by electron crystallography. Nature **367** 614-621.

Kumar, A., Malloch, R. A., Fujita, N., Smillie, D. A., Ishihama, A. and Hayward, R. S. (1993) The -35 binding region of sig-70 is inessential for initiation of transcription from an 'extended -10 promoter'. Journal of Molecular Biology **232** 406-418.

Kunst, F., Ogasawara, N., Moszer, I. and 148 others (1997) The complete genome sequence of the gram-positive bacterium *Bacillus subtilis* Nature **390** 249-256.

Lampson, B. C. (1993) Retron elements of the Myxobacteria. In Dworkin, M. and Kaiser, D. (ed.), *Myxobacteria II.* American Society for Microbiology, Washington D.C.

Landschulz, W. H., Johnson, P. F. and McKnight, S. L. (1988) The leucine zipper: A hypothetical structure common to a new class of DNA binding proteins. Science **240** 1759-1764.

Laue, B. E. and Gill, R. E. (1995) Using a phase-locked mutant of *Myxococcus xanthus* to study the role of phase variation in development. Journal of Bacteriology **177** 4089-4096.

Lesley, S. A. and Burgess, R. R. (1989) Characterisation of the *Escherichia coli* transcription factor sigma 70: Localisation of a region involved in the interaction with core RNA polymerase. Biochemistry **28** 7728-7734.

Liaaen-Jensen, S. (1963) Carotenoids of photosynthetic bacteria-distribution, structure and biosynthesis. In: Bacterial Potosynthesis. Guest, H., San Pietro, A and Vernon (Eds.), 19-34. (Antioch Press, Yellow Springs, Ohio.)

Linden, H., Misawa, N., Chamovitz, D., Pecker, I., Hirschberg, J. and Sandmann, G. (1991) Functional complementation in *Escherichia coli* of different phytoene desaturase genes and analysis of accumulated carotenes. Z. Naturforsch **46c** 1045-1051.

Linden, H., Misawa, N., Saito, T. and Sandmann, G. (1994) A novel carotenoid biosynthesis gene coding for ζ-carotene desaturase: functional expression, sequence and phylogenetic origin. Plant Molecular Biology **24** 369-379.

Linden, H. and Macino, G. (1997) White collar 2, a partner in blue-light signal transduction, controlling expression of light-regulated genes in *Neurospora crassa*. The EMBO Journal **16** 98-109.

Loewen, P. C. and Hengge-Aronis, R. (1994) The role of the sigma factor σ^S (KatF) in bacterial global regulation. Annual Review of Microbiology **48** 53-80.

Lonetto, M. A. and Gross, C. A. (1996) Nomenclature of sigma factors from *Escherichia coli* and *Salmonella typhimurium* and relationships to sigma factors from other organisms. In *Escherichia coli* and *Salmonella typhimurium:* Cellular and molecular biology. Neihardt, F. C., Curits, R. I., Ingraham, J. L., Lin, E. C. C., Low, K., Reznikoff, W. S., *et al.,* (Eds.) Washington D.C. American Society for Microbiology, 1382-1399.

Lonetto, M., Gribskov, M. and Gross, C. A. (1992) The σ^{70} family: sequence conservation and evolutionary relationships. Journal of Bacteriology **174** 3843-3849.

Lonetto, M. A., Brown, K. L., Rudd, K. E. and Buttner, M. J. (1994) Analysis of *S. coelicolor sigE* gene reveals the existance of a sub-family of eubacterial RNA polymerase sigma factors involved in the regulation of extracytoplasmic functions. Proceedings of the National Academy of Science (USA) **91** 7573-7577.

Lynch, A. S. and Wang, J. C. (1993) Anchoring of DNA to the bacterial cytoplasmic membrane through cotranscriptional synthesis of polypeptides encoding membrane proteins or proteins for export: a mechanism of plasmid hypernegative supercoiling in mutants deficient in DNA topoisomerase I. Journal of Bacteriology **175** 1645-1655.

Mandel, M. and Leadbetter, E. R. (1965) Deoxyribonucleic acid base composition of myxobacteria. Journal of Bacteriology **90** 1795-1796.

Marshak *et al.,* (1996) Strategies for protein purification and characterisation. CSHL Press, CA, USA.

Martin, S., Sodergren, R., Masuda, T and Kaiser, D. (1978) Systematic isolation of transducing phages for *Myxococcus xanthus*. Virology **88** 44-53

Martin, D. W., Holloway, B. W. and Deretic, V. (1993) characterisation of a locus determining the mucoid status of *Psuedomonas aeroginosa*: AlgU sequence similarities with a *Bacillus* sigma factor. Journal of Bacteriology **175** 1153-1164.

Martinez-Argudo, I., Ruiz-Vazquez, R. M. and Murillo, F. J. (1998) The structure of an ECF-σ-dependent, light-inducible promoter from the bacterium *Myxococcus xanthus*. Molecular Microbiology **30** 883-893.

Martinez-Labordez, A., Elias, M., Ruiz-Vazquez, R. and Murillo, F. J. (1986) Insertions of Tn*5* linked to mutations affecting carotenoid synthesis in *Myxococcus xanthus*. Molecular and General Genetics **205** 107-114.

Martinez-Labordez, A. and Murillo, F. J. (1989) Genic and allelic interactions in the carotenogenic response of *Myxococcus xanthus* to blue light. Genetics **122** 481-490.

Martinez-Labordez, A., Balsalobre, J. M., Fontes, M. and Murillo, F. J. (1990) Accumulation of carotenoids in structural and regulatory mutants of the bacterium *Myxococcus xanthus*. Genetics **223** 205-210.

Masters, B. S., Stohl, L. L. and Clayton, D. A. (1987) Yeast mitochondrial RNA polymerase is homologous to those encoded by bacteriophages T3 and T7. Cell **51** 89-99.

McGowan, S. J., Gorham, H. C. and Hodgson, D. A. (1993) Light-induced carotenogenesis in *Myxococcus xanthus*: DNA sequence analysis of the CarR Region. Molecular Microbiology **10** (4) 713-735.

McGowan, S. J. (1992) Further genetic studies of the *carQRS* region of *Myxococcus xanthus*. PhD thesis. University of Warwick, UK.

Misawa, N., Nakagawa, M., Kobayashi, K., Yamano, S., Izawa, Y., Nakamura, K. and Harashima, K. (1990) Elucidation of the *Erwinia uredevora* carotenoid biosynthetic pathway by functional analysis of gene products expressed in *E. coli*. Journal of Bacteriology **172** 6704-6712.

Missiakas, D. and Raina, S. (1998) The extracytoplasmic function sigma factors: role and regulation. Molecular Microbiology **28** 1059-1066.

Mojica, F. J. M. and Higgins, C. F. (1996) Localised domains of DNA supercoiling: topological coupling between promoters. Molecular Microbiology **22** 919-928.

Munoz-Dorado, J., Inouye, S. and Inouye, M. (1991) A gene encoding a protein serine/threonine kinase is required for normal development of *M. xanthus*, a Gram negative bacterium. Cell **67** 995-1006.

Nguyen, L. H., Jensen, D. B. and Burgess, R. R. (1993) Overproduction and purification of σ^{32}, the *Escherichia coli* heat shock transcription factor. Protein Expression and Purification **4** 425-433.

Nicolas, F. J., Ruiz-Vasquez, R. and Murillo F. J. (1994) A genetic link between light response and multicellular development in the bacterium *Myxococcus xanthus.* Genes and Development **8** 2375-2387.

Nicolas, F. J., Cayuela, . L., Martinez-Argudo, I. M., Ruiz-Vazquez, R. M. and Murillo, F. J. (1996) High mobility group I(Y)-like DNA binding domains on a bacterial transcription factor. Proceedings of the National Academy of Science (USA) **93** 6881-6885.

Nunoshiba, T., Hidalgo, E., Amabile Cuevas, C. F. and Demple, B. (1992) wo-stage control of an oxidative stress regulon: the *Escherichia coli* SoxR protein triggers redox-inducible expression of the *soxS* regulatory gene. Journal of Bacteriology **174** 6054-6060.

Orndorff, P., Stellwag, E., Starich, T., Dworkin, M. and Zissler, J. (1983) Genetic and physical characterisation of lysogeny by bacteriophage Mx8 in *Myxococcus xanthus*. Journal of Bacteriology **154** 772-779.

O'Connor K. A. and Zusman, D. R. (1983) Coliphage P1-mediated transduction of cloned DNA from *Escherichia coli* to *Myxococcus xanthus*: Use for complementation and recombinatorial analyses. Journal of Bacteriology **155** 317-329.

O'Connor, K.A., McBride, M. J., West, M., Yu, H., Trinh, L.,Yuan, K., Lee, T. and Zusman, D. R. (1996) Photolyase of *M. xanthus*, a Gram-negative eubacterium is more similar to photolyases found in archaea and 'higher' eukaryotes than to photolyases of other eubacteria. Journal of Biological Chemistry **271** 6252-6259.

Oyaizu, H. and Woese, C. R. (1985) Phylogenetic relationships among the sulphate respiring bacteria, myxobacteria and purple bacteria. Systematic and Applied Microbiology **6** 257-263.

Paget, M. S. B. and Butner, M. J. (1998) A redox-sensitive σ factor: anti-σ factor couple that modulates expression of the thioredoxin system in response to oxidative stress in *Streptomyces coelicolor.* In: Abstracts of the 141st Ordinary Meeting of the Society for General Microbiology. Norwich, (UK).

Paget, M. S. B., Kang, J. -G., Roe, J. -H. and Buttner, M. J. (1998) σ^R, an RNA polymerase sigma factor that modulates expression of the thioredoxin system in response to oxidative stress in *Streptomyces coelicolor* A3(2). The EMBO Journal **17** 5776-5782.

Paietta, J. and Sargent, M. I. (1981) Photoprotection in *Neurospora crassa* - correlation of reduced light sensitivity with flavin deficiency. Proceedings of the National Academy of Science (USA) **78** 5573-5577.

Pan, J-M., Haring, M. A. and Beck, C. F. (1996) Dissection of the blue-light-dependent signal-transduction pathway involved in gametic differentiation of *Chlamydomonas reinhardtii.* Plant Physiology **112** 303-309.

Parkhill, J., Lawley, B., Hobman, J. L. and Brown, N. L. (1998) Selection and characterisation of mercury-independent activation mutants of the Tn*501* transcriptional regulator, MerR. Microbiology **144** 2855-2864.

Perego, M. and Hoch, J. A. (1996) Cell-cell communication regulates the effects of protein aspartate phosphatases on the phosphorelay controlling development in *Bacillus subtilis.* Proceedings of the National Academy of Science (USA) **93** 1549-1553.

Perez-Martin, J. and Espinosa, M. (1994) Correlation between DNA bending and transcriptional activation at a plasmid promoter. Journal of Molecular Biology **241** 7-17.

Porter, J. W. and Lincoln, R. E. (1950) I. *Lycopersicon* selections containing a high content of carotenes and colourless polyenes. II. The mechanism of carotene biosynthesis. Archives of Biochemistry **27** 390-403.

Poster J. B. and Dean, N. (1996) The yeast VRG4 gene is required for normal golgi functions and defines a new family of related genes. Journal of Biological Chemistry **271** 3837-3845.

Quinones, M. A. and Zeiger, E. (1994) A putative role of the xanthophyll, zeaxanthin, in blue light photoreception of corn coleoptiles. Science **264** 558-561.

Ramakrishnan, L., Tran, H. T., Federspiel, N. A. and Falkow, S. (1997) A *crtB* homolog essential for photochromicity in *Mycobacterium marinum*: characterisation and gene disruption via homologous recombination. Journal of Bacteriology **179** 5862-5868.

Rao, L., Ross, W., Aleman, J. A., Gaal, T., Leirmo, S., Schlax, P. J., Record, Jnr., M. T. and Gourse, R. L. (1994) An 'extended' promoter with an upstream element that dramatically increases promoter strength. Journal of Molecular Biology **235** 1421-1435.

Reichenbach, H. (1993) Biology of the myxobacteria: ecology and taxonomy. In Dworkin, M. and Kaiser, D. (ed.), *Myxobacteria II.* American Society for Microbiology, Washington D.C.

Reichenbach, H. and Hofle, G. (1993) production of bioactive secondary metabolites. In Dworkin, M. and Kaiser, D. (ed.), *Myxobacteria II.* American Society for Microbiology, Washington D.C.

Reichenbach, H. and Kleinig, H. (1984) pigments of myxobacteria, *In* Rosenberg, E. (ed.), *Myxobacteria: Development and Cell Interactions*. Springer-Verlag, New York.

Rex, G., Surin, B., Besse, G., Schnepe, B. and McCarthy, J. E. G. (1994) The mechanism of translational coupling in *Escherichia coli*. Journal of Biological Chemistry **269** 18118-18127.

Robertson, B. D. and Meyer, T. F. (1992) Genetic variation in pathogenic bacteria. Trends in Genetics **8** 422-427.

Robson, P. R. H. (1992) Towards a mechanism of carotenogenesis in *Myxococcus xanthus*. PhD thesis. University of Warwick, UK.

Romeo, J. M. and Zusman, D. R. (1991) Transcription of the myxobacterial hemagglutinin gene is mediated by a σ^{54}-like promoter and a *cis*-acting upstream regulatory region of DNA. Journal of Bacteriology **173** 2969-2976.

Ronson, C. W., Nixon, B. T., Albright, L. M. and Ausubel, F. M. (1987) *Rhizobium meliloti ntrA* (*rpoN*) gene is required for diverse metabolic functions. Journal of Bacteriology **169** 2424-2431.

Rosenberg, A. H., Lade, B. N., Chui, D., Lin, S., Dunn, J. J. and Studier, F. W. (1987) Vectors for selective expression of cloned DNAs by T7 RNA polymerase. Gene **56** 125-135.

Rosner, J. L. (1972) Formation, induction and curing of bacteriophage P1 lysogens. Virology **48** 679-689.

Ross, W., Gosink, K. K., Saloman, J., Igarashi, K., Zou, C., Ishihama, A., Severinov, K. and Gourse, R. L. (1993) A third recognition element in bacterial promoters: DNA binding by the α subunit of RNA polymerase. Science **262** 1407-1413.

Rost, B., Casadio, R., Fariselli, P. and Sander, C. (1995) Transmembrane helices predicted at 95% accuracy. Protein Science **4** 521-533.

Rudd, K. E. and Zusman, D. R. (1982) RNA polymerase of *M. xanthus:* purification and selective transcription *in vitro* with bacteriophage templates. Journal of Bacteriology **151** 89-105.

Ruiz-Vasquez, R., Fontes, M. and Murillo, F.J. (1993) Clustering and co-ordinated activation of carotenoid genes in *Myxococcus xanthus* by blue light. Molecular Microbiology **10** (1) 25-34.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular cloning: a laboratory manual. Second Edition. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, New York, (USA).

Sancar, A. (1994) Structure and function of DNA photolyase. Biochemistry **33** 2-9.

Sancar, A. and Sancar, G. B. (1988) DNA repair enzymes. Annual Review of Biochemistry **57** 29-67.

Sanger, F. S., Nicklen, S. and Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Science (USA). **74** 5463-5467.

Sasse-Dwight, S. and Gralla, J. D. (1990) Role of eukaryotic-type functional domains found in the prokaryotic enhancer receptor factor σ^{54}. Cell **62** 945-954.

Schickor, P., Metzger, W. , Werel, W., Lederer, H. and Heumann, H. (1990) Topography of intermediates in transcription initiation in *E. coli.* The EMBO Journal **9** 2215-2220.

Schieferstein, U. and Thoma, F. (1998) Site-specific repair of cyclobutane pyrimidine dimers in a positioned nucleosome by photolyase and T4 endonuclease V *in vitro*. EMBO Journal **17** 306-316.

Schinkel, A. H., Groot Koerkamp, M. J. A., Touw, E. P. W. and Tabak, H. F. (1987) Specificity factor of yeast mitochondrial RNA polymerase. Journal of Biological Chemistry **262** 12785-12791.

Schmidthauser, T. J., Lauter, F. R., Russo, V. E. A. and Yanofsky, C. (1990) Cloning, Sequence and photoregulation of *al-1*, a carotenoid biosynthetic gene of *Neurospora crassa*. Molecular and Cellular Biology **10** 5064-5070.

Schumann, G., Nurnberger, H., Sandmann, G. and Krugel, H. (1996) Activation and analysis of cryptic *crt* genes for carotenoid biosynthesis from *Streptomyces griseus*. Molecular and General Genetics **252** 658-666.

Schroeder W. A. and Johnson, E. A. (1993) Antioxidant role of carotenoids in *Phaffia rhodozyma*. Journal of General Microbiology **139** 907-912.

Schroeder W. A. and Johnson, E. A. (1995) Singlet oxygen and peroxyl radicals regulate carotenoid biosynthesis in *Phaffia rhodzyma*. Journal of Biological Chemistry **270** 18374-18379.

Severinova, E., Severinov, K., Fenyo, D., Marr, M., Brady, E. N., Roberts, J. W., Chait, B. T. and Darst, S. A. (1996) Domain organisation of the *Escherichia coli* RNA polymerase sig-70 subunit. Journal of Molecular Biology **263** 637-647.

Shimkets, L. J., Gill, R. E. and Kaiser, D. (1983) Developmental cell interactions in *Myxococcus xanthus* and the *spoC* locus. Proceedings of the National Acadamy of Science (USA) **80** 1406-1410.

Siegele, D. A., Hu, J. C., Walter, W. A. and Gross, C. A. (1989) Altered promoter recognition by mutant forms of the sig-70 of *E. coli* RNA polymerase. Journal of Molecular Biology **206** 591-603.

Silhavy, T. J., Berman, M. L. and Enquist, L. W. (1984) Experiments with gene fusions. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, New York, (USA).

Soloman, J. M. and Grossman, A. D. (1996) Who's competent and when: regulation of natural genetic competence in bacteria. Trends in Genetics **12** 150-155.

Stephens, K., Hegeman, G. D. and White, D. (1982) Pheromone produced by the myxobacterium *Stigmatella aurantiaca*. Journal of Bacteriology **149** 739-747.

Stephens, K., Hartzell, P. and Kaiser, D. (1989) Gliding motility in *Myxococcus xanthus*: *mgl* locus, RNA and predicted protein products. Journal of Bacteriology **171** 819-830.

Sudo, S. Z. and Dworkin, M. (1969) Resistance of vegetative cells and microcysts of *Myxococcus xanthus*. Journal of Bacteriology **98** 883-887.

Suwa, K., Kimura, T. and Schaap, A. P. (1977) Reactivity of singlet molecular oxygen with cholesterol in a phospholipid membrane matrix: a model for oxidative damage of membranes. Biochemical and Biophysical Research Communications **75** p785.

Tartaglia, L. A., Gimeno C. J., Storz, G. and Ames, B. D. (1992) Multidegenerate DNA recognition by the *oxyR* transcriptional regulator. Journal of Biological Chemistry **267** 2038-2045

Thaxter, R. (1892) Contributions from the cryptogamic laboratory of Harvard University XVIII. On the Myxobacteriaceae, a new order of Schizomycetes. Botanical Gazette **14** 389-406.

To, K., Lai, E., Lee, L., Lin, T., Hung, C., Che, C., Chang, Y. and Liu, S. (1994) Analysis of the gene cluster encoding carotenoid biosynthesis in *Erwinia herbicola* Eho13. Microbiology **140** 331-339.

Tojo, N., Sanmiya, K., Sugawara, H., Inouye, S. and Komano, T. (1996) integration of bacteriophage Mx8 into the *Myxococcus xanthus* chromosome causes a structural alteration at the C-terminal region of the IntP protein. Journal of Bacteriology **178** 4004-4011.

Trudeau, K. G., Ward, M. J. and Zusman, D. R. (1996) Identification and characterisation of FrzZ, a novel response regulator, necessary for swarming and fruiting body formation in *Myxococcus xanthus*. Molecular Microbiology **20** 645-655.

Ueki, T. and Inouye, S. (1996) New sigma-factors, SigD and SigE from *Myxococcus xanthus*. Abstracts of the 23rd Annual Meeting on the Biology of the Myxobacteria, June 1996 Angel Fire, New Mexico, p. 37.

Valenzeno, P. D. (1987) Photomodification of biological membranes with emphasis on singlet oxygen mechanisms. Photochemistry and Photobiology **46** 147-160.

Varon, M., Paitan, Y. and Rosenberg, E. (1997) Trans-acting regulation of antibiotic TA genes in *Myxococcus xanthus*. FEMS Microbiology Letters **155** 141-146.

von Heijne G., (1989) Control of topology and mode of assembly of a polytopic membrane protein by positively charged residues. Nature **341** 456-458.

von Hiel, P. H. and Berg, O. G. (1989) Facilitated target location in biological systems. Journal of Biological Chemistry **264** 675-678.

Wakarchuk, W. W., Muller, F. W. and Beck, C. F. (1992) Two GC-rich DNA elements of *Chlamydomonas reinhardtii* with complex arrangements of dirctly repeated sequence motifs. Plant Molecular Biology **18** 143-146.

Walker, G. C. (1984) Mutagenesis and inducible responses to DNA damage in *Escherichia coli*. Microbiology Reviews **48** 60-93.

Ward, M. J., Lew, H., Treuner-Lange, A. and Zusman, D. R. (1998) Regulation of motility behaviour in *Myxococcus xanthus* may require an extracyoplasmic-function sigma factor. Journal of Bacteriology **180** 5668-5675.

Weigle, J. J. (1953) Induction of mutagenesis in a bacterial virus. Proceedings of the National Academy of Science (USA) **39** 628-635.

Wilson, C. and Dombroski, A. J. (1997) Region 1 of σ^{70} is required for efficient isomerisation and initiation of transcription by *Escherichia coli* RNA polymerase. Journal of Molecular Biology **267** 60-74.

Wright, F. and Bibb, M. J. (1992) Codon usage in the G+C-rich *Streptomyces* genome. Gene **113** 55-65.

Yager, T. D. and von Hiel, P. H. (1991) A thermodynamic analysis of RNA transcript elongation and termination in *Escherichia coli*. Biochemistry **30** 1097-1118.

Yang, C. and Kaplan, H. B. (1997) *Myxococcus xanthus sasS* encodes a sensor histidine kinase required for early developmental gene expression. Journal of Bacteriology **179** 7759-7767.

Yang, Y., Westcott, T. P., Pedrson, S. C., Tobias, I. and Olson W. K. (1995) Effects of localised bending on DNA supercoiling. Trends in Biological Sciences **20** 313-319.

Yang, Z., Ma, X., Tong, L., Geng, Y., Sun, H., Xu, D., Kaplan, H. and Shi, W. (1998) Further characterisation of *Myxococcus xanthus dif* locus required for social motility. Abstracts of the 25th International Conference on the Biology of the Myxobacteria. June 1998. Arahova, Greece.

Zhang, W., Inouye, M. and Inouye, S. (1996) Reciprocal regulation of the differentiation of *M. xanthus* by Pkn5 and Pkn6, eukaryotic-like Ser/Thr protein kinases. Molecular Microbiology **20** 435-447.

Zheng, M., Aslund, F. and Storz, G. (1998) Activation of the OxyR transcription factor by reversible disulfide bond formation. Science **279** 1718-1721.

Zhu, Y. S. and Hearst, J. E. (1986) Regulation of expression of genes for light-harvesting antenna proteins LH-I and LH-II; reaction centre polypeptides RC-L, RC-M and RC-H; and enzymes of bacteriochlorophyll and carotenoid biosynthesis in *Rhodobacter capsulatus* by light and oxygen. Proceedings of the National Acadamy of Science (USA) **83** 7613-7617.

Zucker, M. (1989) On finding all suboptimal foldings of an RNA molecule. Science **244** 48-52.