The Molecular Biology of Carotenoid Biosynthesis in Bacteria

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor of Philosophy by Catherine Haycock.

September 1996

The Molecular Biology of Carotenoid Biosynthesis in Bacteria
Catherine Haycock

ABSTRACT

Carotenoids are naturally occurring pigments, biosynthesized by plants, fungi, algae and bacteria. Due to their antioxidant properties, many commercial applications of carotenoids are currently being promoted, including their use as sunscreen and anti-disease agents. Commercial exploitation of carotenoid production by bacteria using recombinant DNA technology is a rapidly growing field of research.

The bacterium *Erwinia herbicola* biosynthesizes the carotenoid zeaxanthin; the genes for which, have been cloned and expressed in *Escherichia coli*. Genetic manipulations were performed increasing gene expression, to facilitate increased carotenoid production. This involved cloning the genes such that they were under the control of the bacteriophage T7 promoter. Bacterial carotenoid biosynthesis enzymes are thought to have complex formation either integral to, or associated with the membrane. Cells of *E. coli* carrying the carotenoid (*crt*) genes of *E. herbicola* were fractionated to discover true cellular location of the enzymes. Although some membrane association was apparent the proteins were largely located in the cytoplasm. This could reflect a transient association with membranes or low level contamination of cell fractions with the cytoplasm. Until monoclonal antibodies and subsequent *in situ* hybridizations in a homologous host are available, the intriguing question of localization may go unanswered.

The *crt* genes of *E. herbicola* were used as heterologous probes to screen the bacteria *Brevibacterium linens*, *Micrococcus luteus*, *Rhodococcus ruber* and *Sphingobacterium multivorum* for equivalent *crt* genes. Initially this was performed with the use of Southern blotting, where hybridization was observed. Fully representative genomic libraries were constructed to isolate specific *crt* genes from the above four bacteria. The carotenoid composition of the bacteria chosen was also analysed. *R. ruber* was found to biosynthesize the carotenoid dihydroxyisorenieratene, previously only thought to be produced by *B. linens*. *M. luteus* was found to have isomers of sarcinaxanthin that were not geometrical, as previously thought. The carotenoid content of *B. linens* and *S. multivorum* was verified.

The characterization, to atomic resolution of the light-harvesting pigment-protein complex of the photosynthetic bacterium *Rhodopseudomonas acidophila* has recently been achieved. The carotenoid rhodopin is present in these complexes. A genomic library was screened to isolate the carotenoid biosynthesis genes responsible for rhodopin production. Recombinant phage were isolated that hybridized with the *crtI* gene of *Rhodobacter sphaeroides* and were passed on for further characterization.

Acknowledgements

Thanks to Julie and George for their constant help, enthusiasm and encouragement. Also I am grateful to my 'foster' supervisor Phil. To Jon and John for the early days and Elaine in the later, thanks for putting up with me, hopefully I won't moan quite as much now! Steve (who has always been a top mate) for his help with the antibodies and running the odd errand for me, cheers!

Thanks must go to, Bhupinder Hundle for supplying the *E. herbicola crt* cluster, Neil Hunter and Helen Lang for help with designing degenerate primers and Richard Cogdell for supplying the genomic library of *R. acidophila*.

Thanks must also go to various people in the Department who have helped along the way, Bodi and Damaso for their patience in helping me use the HPLC, and Mark Prescott for running all the mass-spec. analyses. Prof. Rees and his boys Lee and Geoff for letting me become an honorary member of their lab., and Mark Wilkinson for his help in the analysis of Westerns are all owed a huge thankyou. Lastly, Maureen Wilde for her patient photography, thankyou.

Lastly special thanks to Mum and Dad who've always supported, loved, encouraged and believed in me. I will always be grateful.

CONTENTS

		Page
Abstract		i
Acknowledgements		ii
List of contents		iii
List of abbreviations		x
CHAPTER 1		
INTRODUCTION		
1.1	Carotenoids	1
1.1.1	Structure and nomenclature	1
1.1.2	Distribution	2
1.1.3	Properties of carotenoids	3
	(i) General	3
	(ii) Light absorption	3
1.1.4	Functions of carotenoids	4
	(i) Photoprotection	4
	(ii) Light harvesting	4
	(iii) Other functions	5
1.1.5	Uses of carotenoids	5
	(i) Food colorants	5
	(ii) Anti-disease agents	5
	(a) Cancer	5
	(b) Heart disease	6
	(c) Cataract formation and macular degeneration	6
	(d) Erythropoietic protoporphyria	6
1.2	Biosynthesis of carotenoids	7
1.2.1	The formation of phytoene	7
1.2.2	Phytoene desaturation	8
1.2.3	Cyclization reactions	9
1.2.4	Ring modification	9
1.2.5	Regulation of carotenoid biosynthesis	10
	(i) Light	10
	(ii) Genetic	11
	(iii) Growth and development	12
	(iv) Chemical control	12

(v) Herbicides 12
(vi) Feedback inhibition 13
1.2.6 Carotenoids biosynthesised in phototrophic bacteria 13
1.2.7 Carotenoids biosynthesised in non-phototrophic bacteria 14
1.3 Structure and expression of genes involved in carotenoid biosynthesis 15
1.3.1 Phototrophic bacteria 15
1.3.2 Non-phototrophic bacteria 17
(i) *Erwinia herbicola* and *Erwinia uredovora* 17
(ii) *Myxococcus xanthus* and *Myxococcus fulvus* 19
(iii) Others 20
1.3.3 Cyanobacteria 21
1.3.4 Plants 22
(i) *Lycopersicon esculentum* (tomato) 22
(ii) *Zea mays* (corn) 23
(iii) *Capsicum annuum* (pepper) 24
(iv) *Arabidopsis thaliana* 24
(v) Others 25
1.3.5 Fungi 25
1.3.6 Algae 27
1.4 Aims of the thesis 27

CHAPTER 2
MATERIALS AND METHODS

2.1 Materials 28
2.1.1 Solutions, media and materials 28
2.1.2 Growth media 28
(i) Nutrient medium 28
(ii) Nutrient agar 28
(iii) Luria Bertani medium 28
(iv) M9 minimal medium 28
(v) Tryptone top agar 29
2.1.3 Antibiotics 29
2.1.4 Bacterial strains 29
2.1.5 Plasmids 29
2.2 Methods 30
2.2.1 Rapid $CaCl_2$ transformation procedure 30
2.2.2 Rapid preparation of plasmid DNA 30

2.2.3 Preparation of plasmid DNA by alkaline lysis 31
2.2.4 Purification of plasmid DNA by means of NACS Cartridges 32
2.2.5 Preparation of genomic DNA 33
2.2.6 Agarose gel electrophoresis 33
2.2.7 Recovery of DNA from agarose gels by use of the Geneclean II® kit 34
2.2.8 Restriction endonuclease digestion of DNA (plasmid and genomic) 34
2.2.9 Treatment of DNA with alkaline phosphatase 35
2.2.10 Ligation of DNA by T4 DNA ligase 35
2.2.11 PCR amplification 36
2.2.12 Treatment of PCR fragments for blunt ended cloning 37
2.2.13 Analysis of genomic DNA by Southern hybridization 38
(i) Transfer of DNA to membranes via capillary action (Southern blotting) 38
(ii) Radiolabelling of DNA probes 38
(iii) Hybridization of DNA to radiolabelled probes 38
2.2.14 Construction of a genomic library 39
(i) Optimization of conditions for partial digestion with *Sau*3AI 39
(ii) Partial digestion and size fractionation of genomic DNA 40
(iii) Ligation of genomic DNA to *Bam*HI half-site arms 40
(iv) Packaging of ligated DNA, and titration 41
(v) Plaque lifts and hybridization 41
(vi) Preparation of phage lysate 42
2.2.15 RNA Preparation, dot blotting and hybridization 42
2.2.16 Maxicell *in vivo* expression system 43
2.2.17 Induction of cultures by use of IPTG 43
2.2.18 Preparation of periplasmic fraction from ^{35}S-labelled cells, by osmotic shock 43
2.2.19 Cell fractionation of ^{35}S-labelled cells 44
2.2.20 *In vitro* transcription/translation 44
2.2.21 SDS polyacrylamide gel electrophoresis (SDS-PAGE) 45
2.2.22 Staining of SDS polyacrylamide gels 46
2.2.23 Fluorography of ^{35}S-methionine labelled proteins separated by SDS-PAGE 46
2.2.24 Production of antibodies 46
(i) Polypeptide synthesis 47
(ii) Conjugation of polypeptide to Keyhole limpet haemocyanin (KLH) 47
(iii) Immunization of hens 47
(iv) Purification of antibodies from hens egg yolk 47

2.2.25 Western blotting 48
2.2.26 Characterization of pigments 48
(i) Pigment isolation 48
(ii) Reversed-phase HPLC of carotenoid pigments 49
(iii) Thin layer chromatography of carotenoids 49
(iv) Mass spectrometry 49

CHAPTER 3
INCREASING THE EXPRESSION OF CAROTENOID BIOSYNTHESIS ENZYMES IN *E. COLI*

3.1 General introduction 50
3.2 The T7 expression system 51
3.3 Genetic manipulation of *crt* genes 52
(i) Placing the *crt* genes under the control of T7 RNA polymerase 52
(ii) Amplification of *crtZ* by PCR 53
(iii) Transcription all *crt* genes in the same direction under the control of T7 54
3.4 *In vivo* induction of the carotenoid biosynthesis genes 55
3.5 RNA synthesis of the *crt* transcript after *in vivo* induction 56
3.6 Transcription/translation reactions of the *crt* cluster *in vitro* 57
3.7 Quantitative analysis of carotenoid production 57
3.8 Carotenoid biosynthesis in stationary phase 58
3.9 Discussion 60

CHAPTER 4
CELLULAR LOCATION OF THE ENZYMES OF CAROTENOID BIOSYNTHESIS

4.1 General introduction 65
4.2 Predicting possible membrane spanning regions 68
4.3 Maxi-cell *in vivo* expression 68
4.4 Cell fractionation 69
4.5 The generation of polyclonal antibodies against CrtY and CrtI 70
(i) Designing polypeptides 70
(ii) Titration of antisera 71
(iii) Western blotting of cell-fractionated recombinant pT7-CH7 72

(iv) Localizing phytoene dehydrogenase in *E. herbicola* 72
(v) Screening other bacteria for phytoene dehydrogenase 73
4.6 Discussion 74

CHAPTER 5
CHARACTERIZATION OF UNUSUAL CAROTENOIDS AND THE IDENTIFICATION OF THEIR BIOSYNTHETIC GENES

5.1 General introduction 79
5.2 Bacterial strains used in this work 80
(i) *Rhodococcus ruber* 80
(ii) *Brevibacterium linens* 80
(iii) *Sphingobacterium multivorum* 81
(iv) *Micrococcus luteus* 81
5.3 Analysis of pigments extracted from the bacteria 84
(i) Extraction of carotenoids 84
(ii) Reversed-phase HPLC 84
(iii) TLC and Mass spectrometry 85
(iv) Discussion 86
5.4 Analysis of genomic DNA 88
(i) Extraction of DNA 88
(ii) Design of suitable probes 88
(iii) Hybridization experiments 89
(iv) Discussion 90

CHAPTER 6
THE CONSTRUCTION AND SCREENING OF BACTERIAL GENOMIC LIBRARIES

6.1 General introduction 93
(i) Bacteriophage λ 93
(ii) Infection of *E. coli* cells 94
(iii) Bacteriophage growth, lysis and lysogeny 94
(iv) Replication 95
(v) λ-GEM-11 replacement vector 95
(vi) *In vitro* packaging 96
(vii) Bacterial hosts 96

(viii) Size of library 97
6.2 The preparation of genomic DNA for use in constructing genomic ibraries 98
(i) Optimising conditions for partial digestion of bacterial genomic DNA 98
(ii) Large-scale preparation of genomic DNA 98
(iii) Size fractionation of partially digested genomic DNA 99
6.3 Ligation of genomic DNA to λ-GEM-11 *Bam*HI arms 99
6.4 Packaging and titration of recombinant λ-GEM-11 phage 100
6.5 Amplification of the libraries 100
6.6 Discussion 102

CHAPTER 7

SCREENING A GENOMIC LIBRARY FROM *RHODOPSEUDOMONAS ACIDOPHILA*

7.1 General introduction 106
7.2 Design of a suitable probe 108
(i) Degenerate PCR 108
(ii) Redesigned degenerate primers 110
(iii) Other possible probes 110
7.3 Screening the *R. acidophila* library 111
(i) Primary screening 111
(ii) Secondary screening 112
(iii) Tertiary screening 112
7.4 Characterization of putative positive isolates 112
7.5 Discussion 113

CHAPTER 8

SUMMARY 116

FINAL COMMENTS 117

REFERENCES 118

Appendix I 146
Appendix II 148
Appendix III 149

ABBREVIATIONS

A	absorbance
A_{550nm}	absorbance at 550nm
A_{600nm}	absorbance at 600nm
Amp^r	ampicillin resistant
ATP	adenosine 5'-triphosphate
BCIP	5-bromo-4-chloro-3-indolyl phosphate
bp	base pair
cAMP	cyclic adenosine 5'-monophosphate
Cm^r	chloramphenicol resistant
Ci	Curie
CIP	calf intestinal phosphatase
crt	carotenoid
DCM	dichloromethane
dd	double distilled
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dCTP	deoxycytidine 5'-triphosphate
dNTP	deoxyribonuleoside 5'-triphosphate
EB	extraction buffer
EDTA	ethylenediaminetetra-acetic acid
EtBr	ethidium bromide
g	gravitational force
g/l	grams per litre
HPLC	high-performance liquid chromatography
Ig	immunoglobulin
in	inch
IPTG	isopropyl β-D-thiogalactoside
Kan^r	kanamycin resistant
Kb	kilobase
KDa	kilodalton
KLH	keyhole limpet haemocyanin
λ	lambda
lbs	pounds
µg	microgram
mg	miligram
ml	mililitre

mM	milimolar
M	molar
MBS	m-maleimidobenzoic acid N-hydroxysuccinimide ester
NBT	nitro blue tetrazolium
ng	nanogram
nm	nanometer
PCR	polymerase chain reaction
pfu	plaque forming units
PPO	2,5-diphenyloxazol
RNA	ribonucleic acid
RNase	ribonuclease
rpm	revolutions per minute
SDS-PAGE	sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SSPE	standard saline-phosphate-EDTA
TAE	tris-acetate-EDTA buffer
TB	tryptone top agar
TBS	tri-buffered saline
TE	tris-EDTA buffer
TEG	tris-EDTA-glucose buffer
TEMED	*N,N,N',N'*-tetramethylenediamine
Tet^r	tetracycline resistant
tlc	thin layer chromatography
TCA	trichloroacetic acid
Tris	tris (hydroxy) methylamine
U	units
uv	ultraviolet
v	volume
vis	visible
w	weight
wt	wild-type
x	times

CHAPTER 1

INTRODUCTION

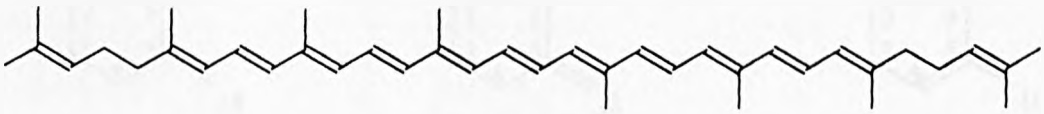
1.1 Carotenoids

Carotenoids form one of the most widely distributed classes of naturally occurring pigments. Over 600 structurally distinct compounds have been isolated. They are biosynthesized by plants, bacteria, fungi and algae; animals cannot biosynthesize carotenoids but some are able to modify dietary carotenoids. Because of their ability to absorb light in the visible region carotenoids appear red, orange or yellow.

1.1.1 Structure and nomenclature

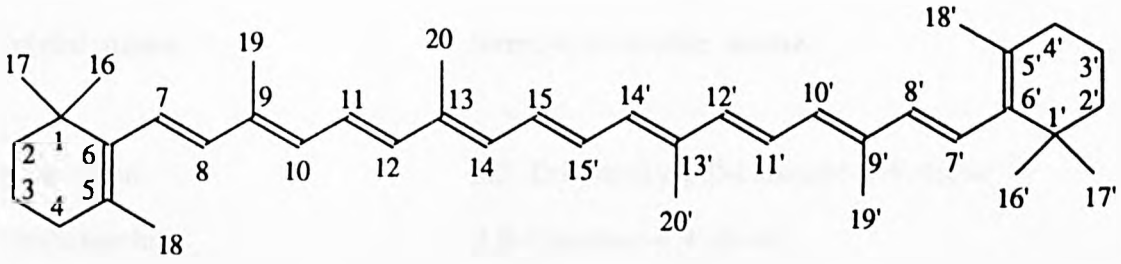
Carotenoids have a tetraterpenoid structure formed by the joining of eight isoprenoid units. The arrangement of the units is reversed about the centre giving a symmetrical molecule exemplified by lycopene in Figure 1.1. This typical C_{40} backbone structure possesses a highly delocalised π-electron system, with 3 to 15 conjugated double bonds and is extremely hydrophobic. Carotenoids containing no oxygen function are known as carotenes and can be modified by cyclization at one end or both ends of the molecule. Xanthophylls are oxygenated derivatives of carotenes modified by hydroxy-, methoxy-, epoxy-, oxo-, aldehyde or carboxylic acid groups, usually in the end ring. Suitable groups can then become esterified or glycosylated. Skeletal modifications of the C_{40} backbone, by either elongation or degradation, give rise to C_{30}, C_{45} and C_{50} carotenoids.

Carotenoids can be named semi-systematically although many are commonly known by their trivial names. For the purpose of this thesis trivial nomenclature will be used with the semi-systematic names given in Table 1.1. Carotenoid structures not depicted in the thesis chapters will be given in Appendix I. The semi-systematic nomenclature is based on the stem name carotene with one or two of 7 different end groups (Figure 1.2). Figure 1.2 also shows the numbering system. The end groups



Lycopene

Figure 1.1 The acyclic, symmetrical structure of lycopene.



β-carotene (β,β-carotene)

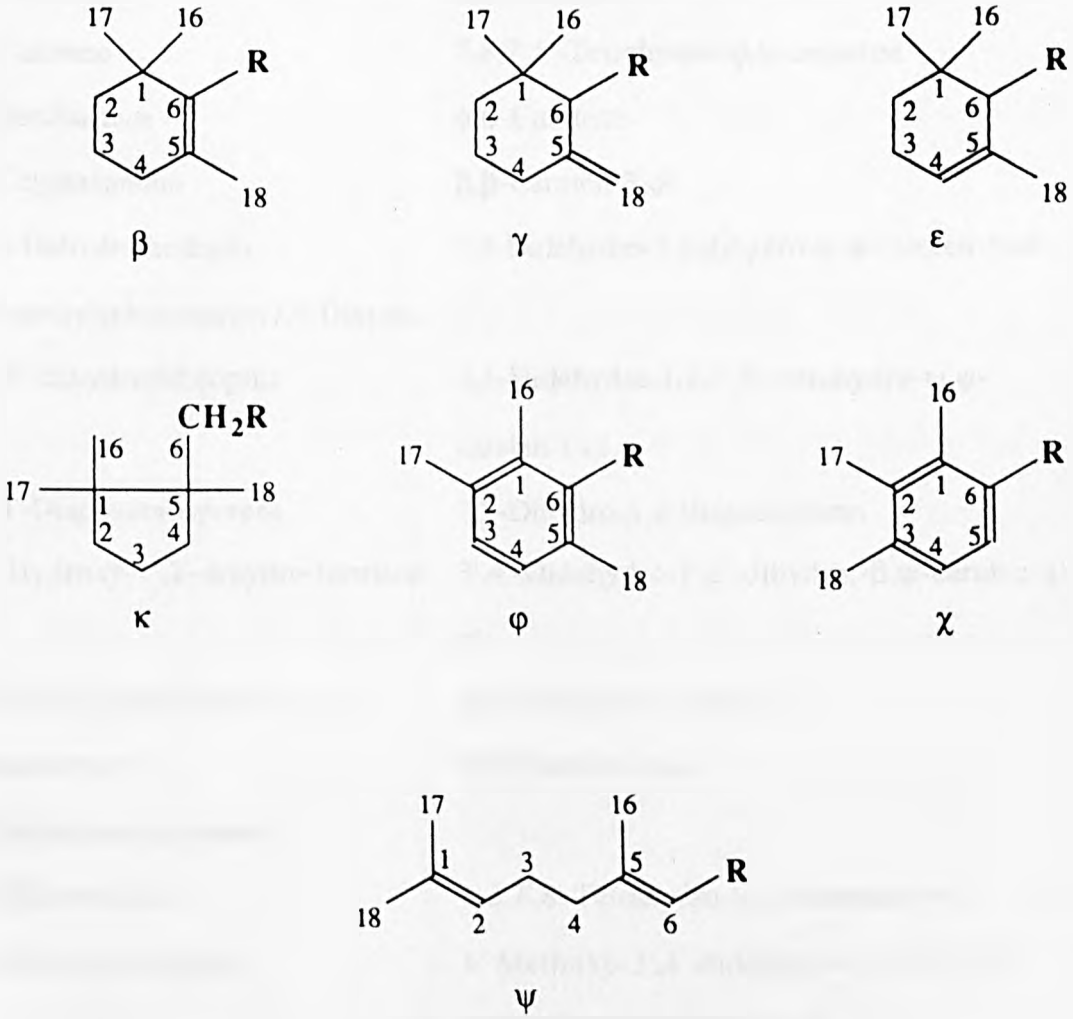


Figure 1.2 Structure and carbon skeleton numbering of β-carotene and end group designations. R - rest of molecule. (Britton, 1983)

Table 1.1 Semi-systematic names of carotenoids mentioned in this thesis

Trivial name	Semi-systematic name
Astaxanthin	3,3'-Dihydroxy-β,β-carotene-4,4'-dione
Canthaxanthin	β,β-Carotene-4,4'-dione
α-Carotene	β,ε-Carotene
β-Carotene	β,β-Carotene
δ-Carotene	ε,ψ-Carotene
ε-Carotene	ε,ε-Carotene
γ-Carotene	β,ψ-Carotene
ζ-Carotene	7,8,7',8'-Tetrahydro-ψ,ψ-carotene
Chlorobactene	ϕ,ψ-Carotene
β-Cryptoxanthin	β,β-Caroten-3-ol
3,4-Dehydrorhodopin	3,4-Didehydro-1,2-dihydro-ψ,ψ-caroten-1-ol
Demethylspheroidene/(7,8-Dihydro-3',4'-dehydrorhodopin)	3,4-Didehydro-1,2,7',8'-tetrahydro-ψ,ψ-caroten-1-ol
4,4'-Diaponeurosporene	7,8-Dihydro-4,4'-diapocarotene
1'-Hydroxy-1',2'-dihydro-torulene	3',4'-Didehydro-1',2'-dihydro,-β,ψ-caroten-1'-ol
3,3'-Dihydroxyisorenieratene	ϕ,ϕ-Carotene-3,3'-diol
Echinenone	β,β-Caroten-4-one
Hydroxyneurosporene/ (Chloroxanthin)	1,2,7',8'-Tetrahydro-ψ,ψ-caroten-1-ol
Hydroxyspheroidene	1'-Methoxy-3',4'-didehydro-1,2,7,8,1',2'-hexahydro-ψ,ψ-caroten-1-ol
Hydroxyspheroidenone	1'-Hydroxy-1-methoxy-3,4-didehydro-1,2,1',2',7',8'-hexahydro-ψ,ψ-caroten-2-one
4-Keto-γ-carotene	β,ψ-Caroten-4-one

Table 1.1 (continued)

2-Ketospirilloxanthin	1,1'-Dimethoxy-3,4,3',4'-tetradehydro-1,2,1',2'-tetrahydro-ψ,ψ-carotene-2-one
4-Ketotorulene	3',4'-Didehydro-β,ψ-caroten-4-one
Leprotene	ϕ,ϕ-Carotene
Lutein	β,ε-Carotene-3,3'-diol
Lycopene	ψ,ψ-Carotene
Methoxyspheroidene	1,1'-Dimethoxy-3,4-didehydro-1,2,1',2',7',8'-hexahydro-ψ,ψ-carotene
Myxobactin	1'-Glucosyloxy-3,4,3',4'-tetradehydro-1',2'-dihydro-β,ψ-carotene
Myxobacton	1'-Glucosyloxy-3',4'-didehydro-1',2'-dihydro-β,ψ-caroten-4-one
Myxoxanthophyll	2'(β-L-Rhamnopyranosyloxy)-3',4'-didehydro-1',2'-dihydro-β,ψ-carotene-3,1'-diol
Neurosporene	7.8 Dihydro-ψ,ψ-carotene
Okenone	1'-Methoxy-1',2'-dihydro-χ,ψ-caroten-4'-one
Phytoene	7,8,11,12,7',8',11',12'-Octahydro-ψ,ψ-carotene
Phytofluene	7,8,11,12,7',8'-Hexahydro-ψ,ψ-carotene
Rhodopin	1,2-Dihydro-ψ,ψ-caroten-1-ol
Sarcinaxanthin	2,2'-Bis(4-hydroxy-3-methylbut-2-enyl)-γ,γ-carotene
Spheroidene	1-Methoxy-3,4-didehydro-1,2,7',8'-tetrahydro-ψ,ψ-carotene
Spheroidenone	1-Methoxy-3,4-didehydro-1,2,7',8-tetrahydro-ψ,ψ-caroten-2-one
Spirilloxanthin	1,1'-Dimethoxy-3,3,3',4'-tetradehydro-1,2,1',2'-tetrahydro-ψ,ψ-carotene

Table 1.1 (continued)

α-Zeacarotene	7',8'-Dihydro-ε,ψ-carotene
β-Zeacarotene	7',8'-Dihydro-β,ψ-carotene
Zeaxanthin	β,β-Carotene-3,3'-diol

are attached to a fully unsaturated all-*trans* polyene chain. This gives a tetraterpene carotene stem name which is prefixed by the Greek letters characteristic of the end groups. These names are then modified by prefixes and suffixes denoting changes in the carbon skeleton by nor, seco and apo, the saturation level by hydro, dehydro, oxygenation by methoxy, epoxy, -ol, -one, and -oic acid and lastly configuration by, *cis*-, *trans*-, (or *Z*, *E*) *R* and *S*. These terms define the structures of all carotenoids. The rules for nomenclature are published by IUPAC Commission on Nomenclature of Organic Chemistry and IUPAC Commission on Biochemical Nomenclature, 1971.

1.1.2 Distribution

As mentioned above carotenoids are widely distributed throughout Nature. They are ubiquitous in photosynthetic organisms. In higher plants they are found in chloroplasts where they form complexes with proteins by non-covalent bonding. Many yellow flowers and orange/red fruits are coloured by carotenoids, usually located in chromoplasts. Algae tend to accumulate a variety of unusual carotenoids in their chloroplasts although some ('secondary carotenoids') are found outside this organelle. Carotenoids are also to be found in many fungi. Some accumulate β-carotene in such high quantities that commercial production is feasible. Many bacteria contain a wide variety of carotenoids. Carotenoids are universally found in phototrophic bacteria in complexes in the photosynthetic apparatus. Many non-phototrophic bacteria also contain carotenoids, mainly located in their cell wall or membranes. Glycosidic structures are frequently present in these bacteria.

The yellow colour of bird eggs is due to the presence of carotenoids. Also the colour of the skin and feathers of many birds is due to carotenoids. Fish, amphibians reptiles and insects may also be coloured by carotenoids. Many marine invertebrates contain keto-carotenoids present as either the free compounds or as carotenoid-protein complexes (carotenoproteins). The best example of a carotenoprotein is the blue α-crustacyanin found in the carapace of the lobster *Homarus gammarus*. Carotenoids are not important in mammalion colouration, although small amounts are found in fatty tissues and liver.

1.1.3 Properties of carotenoids

(i) General

Carotenoids are lipid molecules and as such are soluble in organic solvents. Because of their hydrophobic nature carotenoids are usually located in lipophilic, hydrophobic regions of the cell. Glycosylation or complexing with proteins renders them slightly or completely water-soluble. Carotenoids are very sensitive to oxygen, heat and light either in solution or as solids, and also to acid and sometimes base. Carotenoids are generally part of the integral structures of the cell and as such are protected from the above hazards *in vivo*.

(ii) Light absorption

Carotenoids possess a conjugated double bond system. This gives rise to a delocalised π electron system which confers the ability to absorb light in the uv-visible region of the spectrum. When a photon of light is absorbed a transition occurs between two low lying singlet states the energy of which is dependent on the π electron delocalization of the chromophore. Increasing delocalization, by increasing the length of the chromophore, confers a greater stability on the excited state so that excitation is easier and is achieved by light of lower energy, i.e. longer wavelength. This effect is manifested in the absorption spectra of acyclic carotenoids, increasing from λ_{max} of 275, 285, 296nm for phytoene with 3 conjugated double bonds through the biosynthetic pathway to 440, 470, and 502nm for lycopene with 11 conjugated double bonds (Britton, 1983). *cis*-Isomers and carotenoids with cyclic end groups exhibit steric hindrance resulting in a decrease of π electron overlap thus lowering the λ_{max} and decreasing the spectral fine structure. Some oxygen functions can also alter spectral properties. Conjugated carbonyl groups can increase the length of conjugation giving a higher λ max but much reduced spectral fine structure.

1.1.4 Functions of carotenoids

Many of the functions of carotenoids are directly related to their ability to absorb light.

(i) Photoprotection

Carotenoids have an important protective role in all phototrophic organisms and in many non-phototrophic organisms. This is demonstrated by the fact that green plants lacking carotenoids are unable to survive in the presence of oxygen. This is because triplet excited state chlorophyll donates energy to molecular oxygen to produce the powerful oxidant, singlet oxygen, which rapidly oxidises chlorophyll, lipid, protein and nucleic acid etc., resulting in cell death. Carotenoids quench both triplet sensitizers and singlet oxygen by transferring the excess energy to give carotenoid triplets which decay harmlessly (Goodwin, 1980). The overall scheme is shown in Figure 1.3. The ability to quench is related to the number of conjugated double bonds and the structure of the carotenoid. Carotenoids whose triplet state energy is less than that of singlet oxygen are effective quenchers and so 7 conjugated double bonds are effective with 9 or more offering maximum protection (Cogdell, 1988).

(ii) Light harvesting

Carotenoids have another function in phototrophic organisms, as accessory light absorbers. Carotenoids absorb light in the 450-570nm region of the spectrum where chlorophyll and bacteriochlorophyll do not. This excess energy is then passed on to chlorophyll and bacteriochlorophyll by singlet energy transfer (Cogdell and Frank, 1987). To allow this to occur the carotenoid and chlorophyll need to be in close proximity. Carotenoids absorbing light energy are excited from the ground to the S2 state. This is very short lived and relaxation to S1 occurs before energy transfer to chlorophyll. The whole scheme is depicted in Figure 1.4.

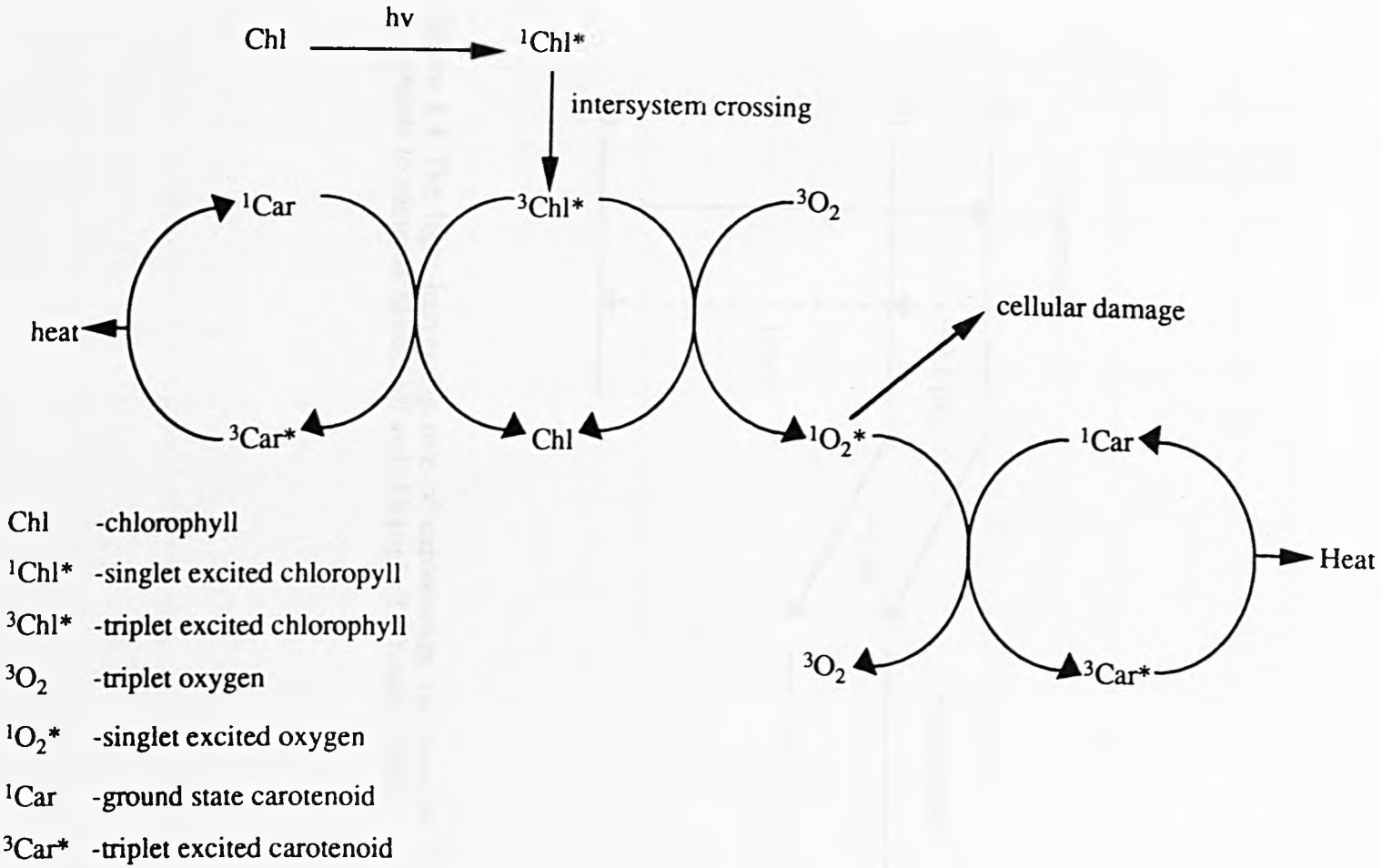


Figure 1.3 The photoprotective role of carotenoids in photosynthetic tissue, (Woodall, 1994).

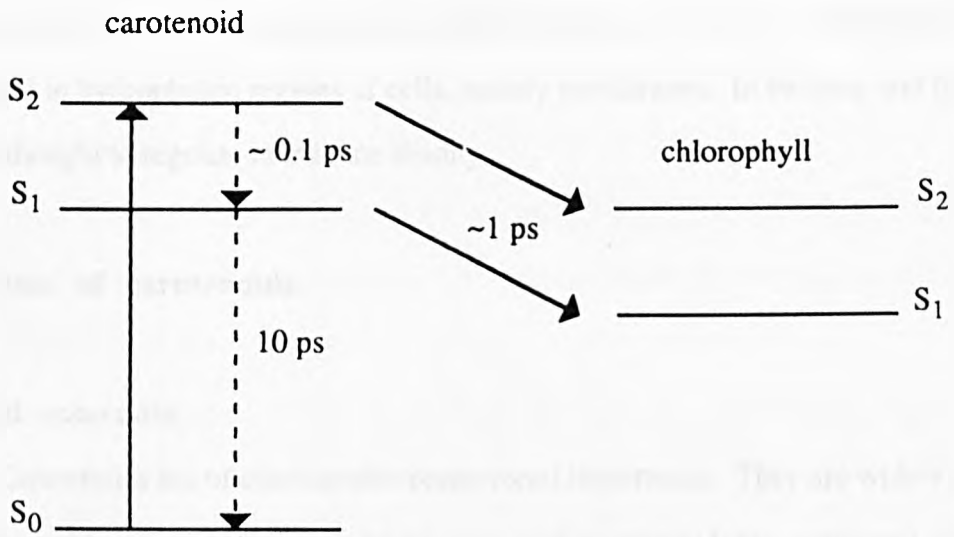


Figure 1.4 The light harvesting role of carotenoids, the transfer of energy from carotenoids to bacteriochlorophyll and chlorophyll (Owens, 1992).

(iii) Other functions

Carotenoids absorbed from the diet by a variety of animals have many functions. Carotenoids, especially β-carotene, are vitamin A precursors. One of the primary functions of vitamin A is in the visual process, but it is also involved in reproduction and growth. Other carotenoids accumulate in the reproductive organs and in the eggs of animals. This suggests a role in reproduction but this has yet to be proven.

The pigmentation of various animals is due to carotenoids and this can be used either as camouflage or as a display to potential predators or mates. Carotenoids are distributed in hydrophobic regions of cells, mainly membranes. In bacteria and plants they are thought to regulate membrane fluidity.

1.1.5 Uses of carotenoids

(i) Food colorants

Carotenoids are of considerable commercial importance. They are widely used in the food industry as colorants, and are now sold as antioxidants, sunscreen agents and vitamin precursors.

(ii) Anti-disease agents

(a) Cancer

It has become widely accepted that oxygen and organic free radicals can participate in the initiation, promotion and proliferation of carcinogenesis. Epidemiological studies have shown that carotenoids may offer some protection. Individuals with a high intake of carotenoids in food have the lowest risk of various cancers and serum β-carotene levels are also inversely associated with cancer risk (Zeigler, 1989; 1993). Most studies have been made on the possible actions of β-carotene and no discrimination between it and other dietary carotenoids, also possessing similar properties, has been attempted (Zeigler *et al.*, 1992). Several anti-

cancer actions of carotenoids have been proposed. Retinoids influence the promotional stages by increasing the immune response and influencing cell-cell interaction (Sporn and Roberts, 1983). The anti-oxidant effects of carotenoids may have an effect on quenching free radicals which cause lipid peroxidation, protein inactivation and nucleic acid damage (Sies *et al.*, 1992; Burton, 1989). The enhancement of gap junction communication promoted by carotenoids *in vitro* may cause an increase of cell to cell communication *in vivo*, restricting clonal expansion of initiated cells (Zhang *et al.*, 1991; 1992; Acevedo and Bertram, 1995). β-Carotene has also been shown to enhance the immune response and in so doing, to decrease tumour size (Bendich, 1989).

(b) Heart disease

Epidemiological studies have shown an inverse correlation between carotenoid intake and cardiovascular disease (Gey, 1993). These studies have been mainly done with β-carotene. It is proposed that the oxidation of low density lipoprotein (LDL) may be prevented (Esterbauer *et al.*, 1993). This is important since this is one of the primary factors in the pathogenesis of atherosclerosis.

(c) Cataract formation and macular degeneration

There are two major causes of impaired sight in the richer countries of the world, these being cataract formation and macular degeneration. The risks are increased by reducing carotenoid intake (Jacques *et al.*, 1988). The eye tissue is protected from oxidative damage by the carotenoids lutein (Appendix I) and zeaxanthin (found in the macula) which may act by quenching singlet oxygen and free radicals (Bendich, 1994).

(d) Erythropoietic protoporphyria

Observations that carotenoids found in plants and bacteria offered protection against photosensitization led to the proposal that they may be of use in the treatment of photosensitivity diseases. β-Carotene has been shown to ameliorate photosensitivity in the majority of patients suffering from erythropoietic protoporphyria (EPP). This is a

disorder of porphyrin metabolism and is characterized by increased levels of protoporphyrin IX deposited in the skin, and sensitivity to blue light within the range 380-560nm. The protoporphyrin acts as a photosensitizer and causes the formation of singlet oxygen. It is thought that β-carotene offers protection by the quenching of photosensitizer triplet states or singlet oxygen (Mathews-Roth, 1986).

1.2 Biosynthesis of carotenoids

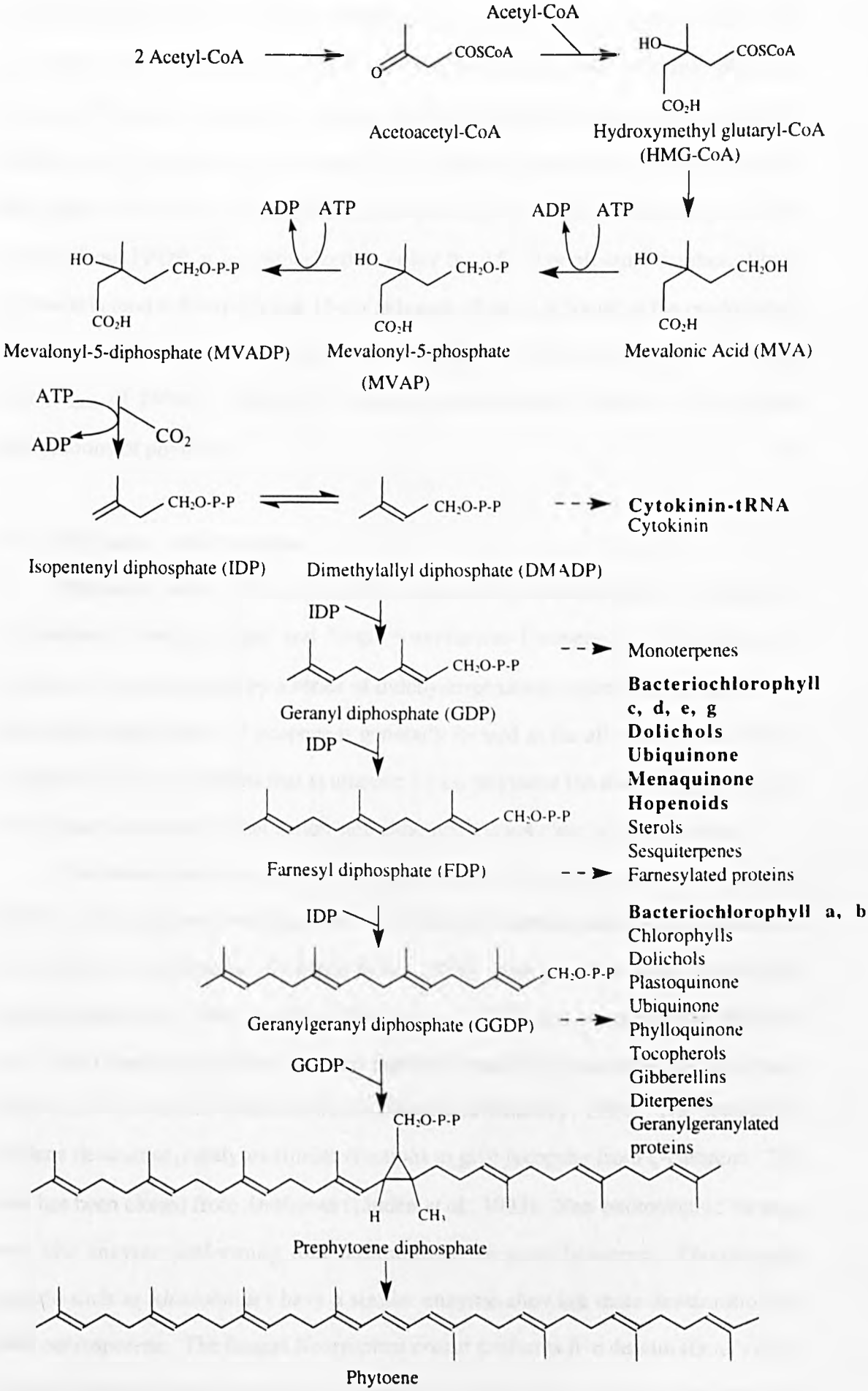
1.2.1 The formation of phytoene

The early steps of carotenoid biosynthesis follow the general isoprenoid pathway common to all isoprenoid compounds. Carotenoid biosynthesis follows a common pathway leading to the formation of phytoene in eukaryotes and prokaryotes. Extensive reviews of biosynthesis exist and can be found elsewhere (Britton, 1988; Bramley, 1985; Goodwin, 1980; Britton, 1976).

The first precursor of the pathway is mevalonic acid (MVA) which is decarboxylated to the ubiquitous C_5 isoprenoid precursor isopentenyl diphosphate (IDP). IDP then isomerizes to form dimethylallyl diphosphate (DMADP). The condensation of one molecule each of IDP and DMADP gives rise to geranyl diphosphate (GDP). Two further condensation reactions then proceed in which 2 molecules of IDP join to GDP to form the C_{20} geranylgeranyl diphosphate (GGDP). This is the first compound committed to carotenoid biosynthesis and its formation is depicted in Figure 1.5. An alternative pathway of isoprenoid biosynthesis has recently been discovered in the green alga *Scenedesmus obliquus* (Schwender *et al.*, 1996). The carotenoids β-carotene and lutein are not synthesized by the acetate/mevalonate pathway but via a novel glyceraldehyde 3-phosphate (G-3-P)/pyruvate route. IDP is formed by the condensation of a C_2 unit (derived from pyruvate carboxylation) with G-3-P.

Two molecules of GGDP condense head to head to form prephytoene diphosphate (PPDP) which is rapidly converted to yield phytoene. It was originally thought that the product of *crtE* catalyses the first reaction while the *crtB* product is

Figure 1.5 The formation of phytoene. Branches of the common isoprenoid pathway are given to the right. Important compounds found in some or all eubacteria are in bold, while eukaryotic substances are in normal type face (adapted from Armstrong, 1994; Britton, 1983).


2 Acetyl-CoA
Acetyl-CoA
COSCoA
O
HO
COSCoA
CO₂H
Acetoacetyl-CoA
Hydroxymethyl glutaryl-CoA
(HMG-CoA)
ADP
ATP
ADP
ATP
HO
CH₂O-P-P
CO₂H
HO
CH₂O-P
CO₂H
HO
CH₂OH
CO₂H
Mevalonyl-5-diphosphate (MVADP)
Mevalonyl-5-phosphate
(MVAP)
Mevalonic Acid (MVA)
ATP
ADP
CO₂
CH₂O-P-P
CH₂O-P-P
Cytokinin-tRNA
Cytokinin
Isopentenyl diphosphate (IDP)
Dimethylallyl diphosphate (DMADP)
IDP
CH₂O-P-P
Monoterpenes
Bacteriochlorophyll c, d, e, g
Dolichols
Ubiquinone
Menaquinone
Hopenoids
Sterols
Sesquiterpenes
Geranyl diphosphate (GDP)
IDP
CH₂O-P-P
Farnesyl diphosphate (FDP)
Farnesylated proteins
Bacteriochlorophyll a, b
Chlorophylls
Dolichols
Plastoquinone
Ubiquinone
Phylloquinone
Tocopherols
Gibberellins
Diterpenes
Geranylgeranylated proteins
IDP
CH₂O-P-P
Geranylgeranyl diphosphate (GGDP)
GGDP
CH₂O-P-P
H
CH₃
Prephytoene diphosphate
Phytoene

responsible for the second. Recent studies have now shown that the gene product CrtE is in fact GGDP synthase and CrtB encodes the bi-functional enzyme, phytoene synthase. Phytoene synthase has been purified from *Capsicum annuum* (Dogbo *et al.*, 1988) and its activity has also been demonstrated in numerous cell-free systems. Phytoene is the first C_{40} carotenoid and, depending on the stereochemistry of the H removal from PPDP, it is synthesized as either the 15-*cis* or all-*trans* isomer. Plants and bacteria tend to biosynthesise 15-*cis* although all-*trans* is found as the predominant isomer in some bacteria. Phytoene has three conjugated double bonds and is colourless with a λ_{max} of 286nm. Subsequent coloured carotenoids are formed by consecutive desaturations of phytoene.

1.2.2 Phytoene desaturation

Phytoene desaturation is the most intensively studied area of carotenoid biosynthesis. Bacteria, algae and fungi biosynthesize lycopene via phytofluene, ζ-carotene and neurosporene by a series of didehydrogenations alternately to the left and right of the chromophore. Lycopene is generally formed as the all-*trans* isomer shown in Figure 1.6. For organisms that synthesise 15-*cis* phytoene the above sequence must also contain a *cis-trans* isomerisation step although it is not clear when this occurs.

Two desaturation enzymes have been isolated from plants and cyanobacteria. The first, Pds, catalyses two desaturations yielding ζ-carotene and was first identified in *Rhodobacter capsulatus* (Guiliano *et al.*, 1986). The gene has been cloned from tomato (Mann *et al.*, 1994), soybean (Bartley *et al.*, 1991) and cyanobacteria (Misawa *et al.*, 1993) and the enzyme has been purified from *Synechococcus* (Fraser *et al.*, 1993) and *Phycomyces blakesleeanus* (Fraser and Bramley, 1994) The second, ζ-carotene desaturase, catalyses similar reactions to give lycopene from ζ-carotene. The gene has been cloned from *Anabaena* (Linden *et al.*, 1993). Non-phototrophic bacteria have one enzyme performing four desaturations to yield lycopene. Phototrophic bacteria such as *Rhodobacter* have a similar enzyme allowing three desaturations to yield neurosporene. The fungus *Neurospora crassa* performs five desaturations to give 3,4-didehydrolycopene (Bartley *et al.*, 1990). The 'pds' type enzyme is common to all

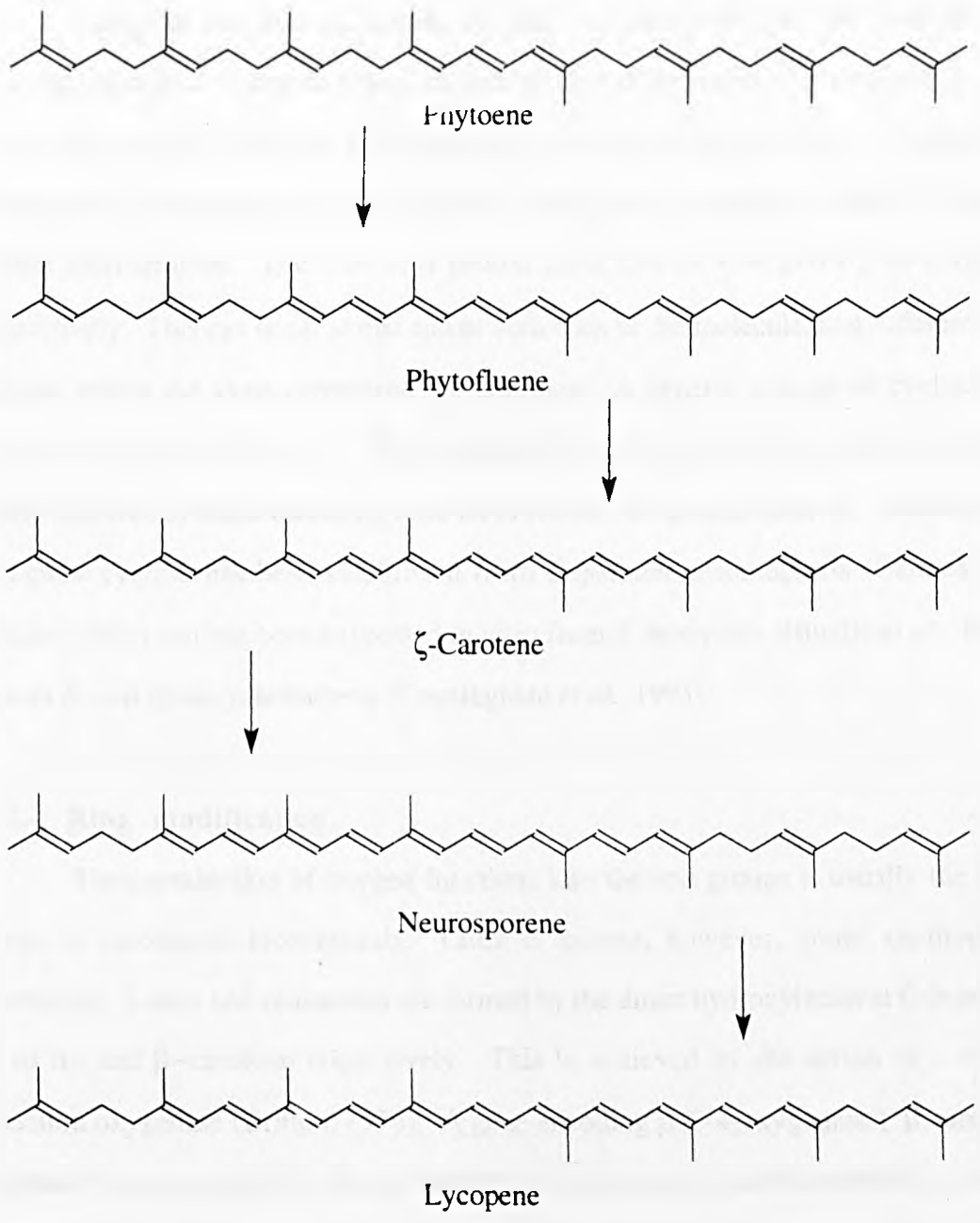


Figure 1.6 The desaturations of phytoene to form lycopene (Bramley and Mackenzie, 1988).

green phototrophic organisms and the 'CrtI' type performs 3, 4 or 5 desaturations depending on the "non-green" host organism.

1.2.3 Cyclization reactions

Lycopene can undergo various cyclization reactions to produce 6-membered ring end groups. Lycopene was shown to be the substrate for cyclization, by experiments using nicotine as an inhibitor of the reaction (Britton, 1988). Cyclization is initiated by proton attack at C-2 of acyclic precursors, producing a stable C-5 carbo cation intermediate. The loss of a proton from C-6 or C-4 gives β or ε rings, respectively. This can occur at one end or both ends of the molecule, and different end groups within the same carotenoid are common. A general scheme of cyclization products is given in Figure 1.7. The cyclization of lycopene has been demonstrated in many cell-free systems including ones from tomato, fungi and bacteria. The enzyme lycopene cyclase has been solubilized from *Capsicum* chromoplasts (Camara and Dogbo, 1986) and has been expressed *in vitro* from *E. herbicola* (Hundle *et al.*, 1993) and in *E. coli* from cyanobacteria (Cunningham *et al.*, 1993).

1.2.4 Ring modification

The introduction of oxygen functions into the end groups is usually the final stage in carotenoid biosynthesis. Little is known, however, about xanthophyll formation. Lutein and zeaxanthin are formed by the direct hydroxylation at C-3 and C-3' of α– and β–carotene respectively. This is achieved by the action of a mixed function oxygenase (Britton, 1976). A gene encoding β-C-4-oxygenase ("β-carotene ketolase"), responsible for the conversion of β-carotene to canthaxanthin, has been cloned from *Haematococcus pluvialis*, and expressed in *E. coli* (Lotan and Hirschberg, 1995; Kajiwara *et al.*, 1995). Another "ketolase" gene, designated *crtW* has been isolated from the marine bacteria *Agrobacterium aurantiacum* and *Alcaligenes* PC-1 (Misawa *et al.*, 1995). The enzyme converts the methylene groups of β-carotene to keto groups to eventually synthesize canthaxanthin. The hydroxylation of β-carotene to produce β-cryptoxanthin (Appendix I) carried out by *Aphanocapsa* is another reaction

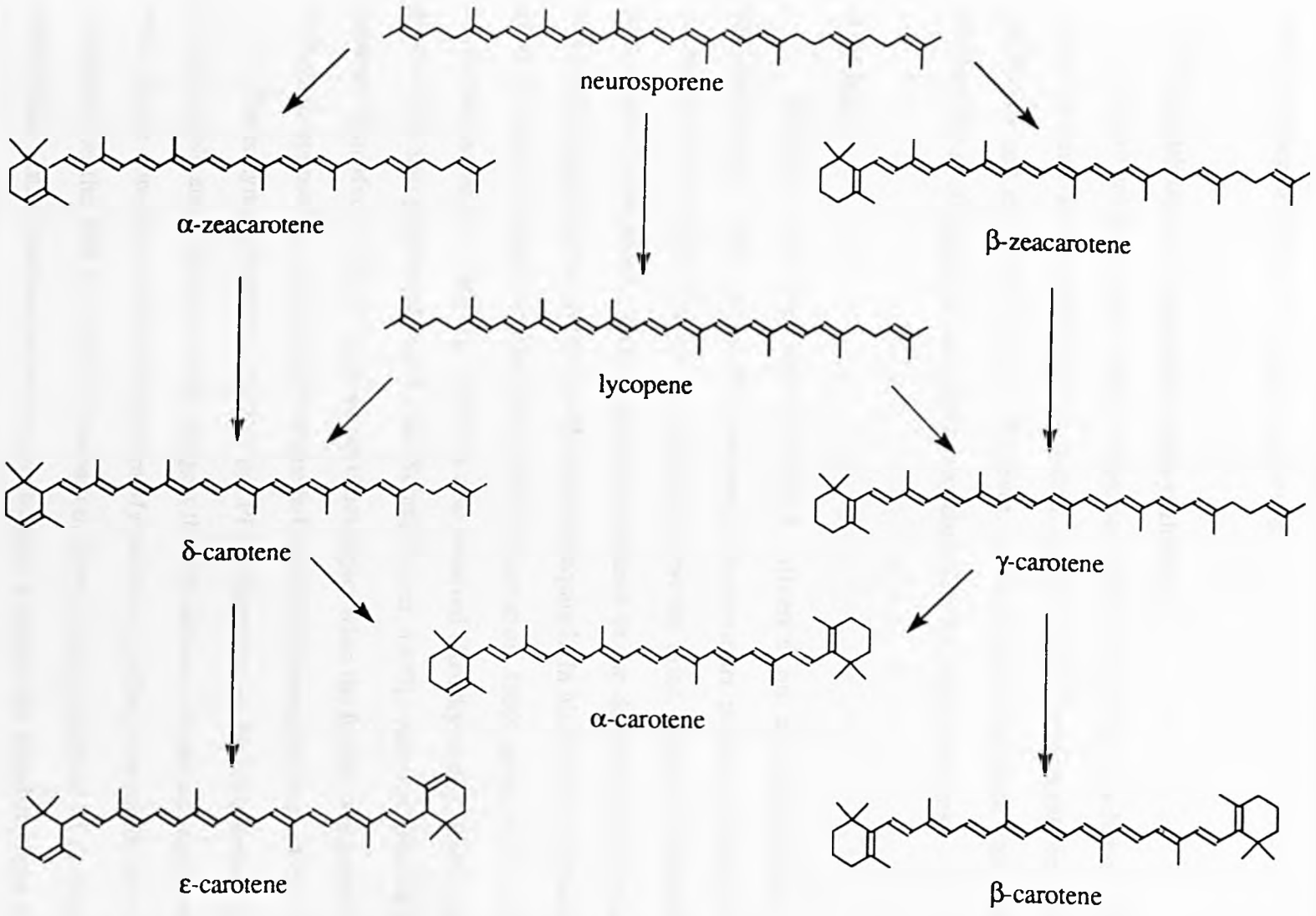


Figure 1.7 An overall scheme for the biosynthesis of cyclic carotenoids (Bramley, 1985).

studied in detail (Sandmann and Bramley, 1985). The mechanisms that are involved in the introduction of other groups remain unclear, and only through the cloning and characterization of these 'ring-modification' genes and their products will a better understanding of these mechanisms be achieved.

1.2.5 Regulation of carotenoid biosynthesis

There are many factors that regulate carotenoid biosynthesis, including light, transcriptional and translational regulation, growth and development, feedback inhibition and chemical control. However, not all factors influence carotenoid biosynthesis in all organisms and effects may also differ between organisms.

(i) Light

Probably the most widely studied influence on carotenogenesis is photoinduction. The fungus *Phycomyces blakesleeanus* produces little or no carotenoid in the dark but on illumination, however brief, massive carotenoid production is stimulated. This production continues in the dark but does decline eventually suggesting that synthesis occurs when required. In *Neurospora crassa* the genes *al-1* and *al-3* are under blue-light control (Nelson *et al.*, 1989; Baima *et al.*, 1991; Schmidhauser *et al.*, 1990) and cAMP is also involved (Kritsky *et al.*, 1982). The gene *al-2* is also photoregulated (Li and Schmidhauser, 1995). Accumulation of β-carotene from *Mucor rouxii* was about ten times higher when the fungus was grown in the light as opposed to dark growth (Mosqueda-Cano and Gutierrez-Corona, 1995).

The enzymes involved in GGDP and PPDP formation in *Mycobacterium* are photoinducible and inhibition studies suggest that regulation is at the transcriptional level. In the non-photosynthetic bacterium *Myxococcus xanthus* carotenoids are not synthesized in the dark but, upon illumination, there is huge synthesis of coloured carotenoids. Tn5 insertions led to the production of strains not requiring light for carotenogenesis, which is an absolute requirement in the wild-type (Martinez-Laborda *et al.*, 1986). The product of the gene *carR* is thought to be a negative regulator of synthesis in the dark (McGowan *et al.*, 1993; Hodgson, 1993). *CarC* is also tightly

regulated by blue light (Fontes *et al.*, 1993; Ruiz-Vazquez *et al.*, 1993). This will be discussed further in Section 1.3.2 (ii).

Carotenoids located in plant chloroplasts are most probably synthesized there. Plants grown in the dark do not form chloroplasts but etioplasts that are devoid of chlorophyll but contain large amounts of xanthophylls. Upon exposure to light normal carotenoid synthesis resumes along with the formation of functional chlorophyll (Britton, 1988). This suggests that carotenoid and chlorophyll biosynthesis are closely co-ordinated and that light is the main requirement. Mutants of the alga *Euglena* do not require light for carotenoid biosynthesis, but *Chlorella* and *Scenedesmus* (Sandmann, 1991) do, producing cyclic carotenoids from ζ-carotene. Astaxanthin production by the yeast *Phaffia rhodozyma* has been shown to be photo-inducible (Meyer and Du Preez, 1994) and was maximal under constant illumination.

(ii) Genetic

Colour mutants of *P. blakesleeanus* and *N. crassa* have been subjected to genetic complementation studies to discover regulatory elements. *P. blakesleeanus* contains two genes, *carB* and *carR*, that have 4 and 2 copies of the genes respectively. The products of these genes encode enzymes involved in desaturation (*carB*) and cyclization (*carR*) and form a multi-enzyme complex to synthesize β-carotene from phytoene (Aragon *et al.*, 1976; Candau *et al.*, 1991). A third protein, CarA, is thought to transfer intermediate substrates to the cyclase enzyme. The gene *carS* regulates the amount of carotenoid produced. Similar results were observed in *N. crassa*. The wild type and an *al-2* mutant of this fungus differ in the presence of one protein. This protein is required for complex formation and carotenogenesis.

In the green tissues of plants the desaturation and cyclization of carotenoids is under direct nuclear control. Tomato plants have been widely studied in the genetic control of carotenoid biosynthesis. Several of the genes act before phytoene synthesis and control the flow of precursors into the pathway.

(iii) Growth and development

By variations in culture conditions the carotenoid content of many bacteria and fungi can be altered. Factors that affect quality and quantity of carotenoid synthesis are the carbon and nitrogen sources and mineral, vitamin and growth factor availability. Environmental factors such as aeration, pH and temperature also play a role.

The ripening of fruit causes an increase in carotenoid synthesis as chromoplasts develop. An increase in expression and activity of GGDP synthase in *Capsicum annuum* has been observed in the chloroplast to chromoplast transition of the developing fruit (Kuntz *et al.*, 1992). Carotenoid content in fruits that are removed from the plant are affected by temperature but not light.

(iv) Chemical control

Trisporic acid has been identified as a stimulatory factor in increasing β-carotene production when the + and - strains of *Blakeslea trispora* are grown together. Trisporic acid itself is synthesized from β-carotene and stimulates sporulation and reproduction as well as increasing carotenogenesis and is therefore responsible in part for its own increased production. Abscisic acid, β-ionone and retinol are all structurally similar and produce the same effect (Britton, 1988).

(v) Herbicides

Chemicals which block carotenoid biosynthesis, especially phytoene desaturation remove the protective role of carotenoids in plants so that the plants die on exposure to light and oxygen. These compounds are in use as bleaching herbicides. The most effective are those causing accumulation of phytoene or ζ-carotene. The first compounds used were metflurazon and norflurazon, but many new ones have been developed. Norflurazon is a reversible non-competitive inhibitor of phytoene desaturase. CrtI and 'Pds' type enzymes react differently to norflurazon and CrtI is not always affected (Windhovel *et al.*, 1994a; Sandmann and Fraser, 1993). Point mutations in *Synechococcus* sp. and *Synechocystis* sp. phytoene desaturase enzymes allowing the replacement of one amino acid by another, confer resistance to norflurazon

(Chamovitz *et al.*, 1993). There are many other possible mechanisms of herbicide action on carotenogenesis. The direct blocking of an enzyme active site or the prevention of enzyme synthesis or complex formation could all be important factors. Compounds affecting lycopene cyclization can also be lethal. Although in principle lycopene should be able to offer photoprotection it cannot physically occupy the correct structural sites in the thylakoid and so cannot prevent cell death.

(vi) Feedback inhibition

In *P. blakesleeanus* phytoene desaturation was shown to be inhibited by neurosporene, lycopene, β-zeacarotene and γ-carotene. β-Carotene was also shown to inhibit its own formation. Feedback inhibition of phytoene desaturation has also been demonstrated in *M. xanthus* (Martinez-Laborda *et al.*, 1990) and *Anacystis* (Sandmann and Kowalczyk, 1989).

1.2.6 Carotenoids biosynthesized in phototrophic bacteria

The purple and green sulphur bacteria of the *Chromatiaceae* and *Chlorobiaceae* synthesize aromatic compounds, characteristically chlorobactene (Appendix I) and okenone (Appendix I). Purple non-sulphur phototrophic bacteria are characterized by the production of acyclic carotenoids. Water or hydrogen may be added to the C-1,2 double bond to give either 1-hydroxy, 1-methoxy or 1,2-dihydrocarotenoids. These reactions are thought to be analogous to the first stages of cyclization and are also blocked by cyclization inhibitors such as nicotine. The proposed pathway is shown in Figure 1.8. Gene function in the pathway has been assigned on the basis of Tn5 insertion phenotypes in *Rhodobacter sphaeroides* and assignment of genes in *Rhodobacter capsulatus* by chemical inhibitors and mutants. The biosynthetic pathway follows that of higher plants, algae and cyanobacteria up to the production of neurosporene. In 1980 Scolnik *et al.* proposed a continuation of the pathway on the basis of genetic and chemical analyses of pigments produced by mutants and wild type in the presence of inhibitors. The C-1,2 double bond is hydrated to yield hydroxyneurosporene. Spheroidene is formed by the desaturation of the C-3,4 bond

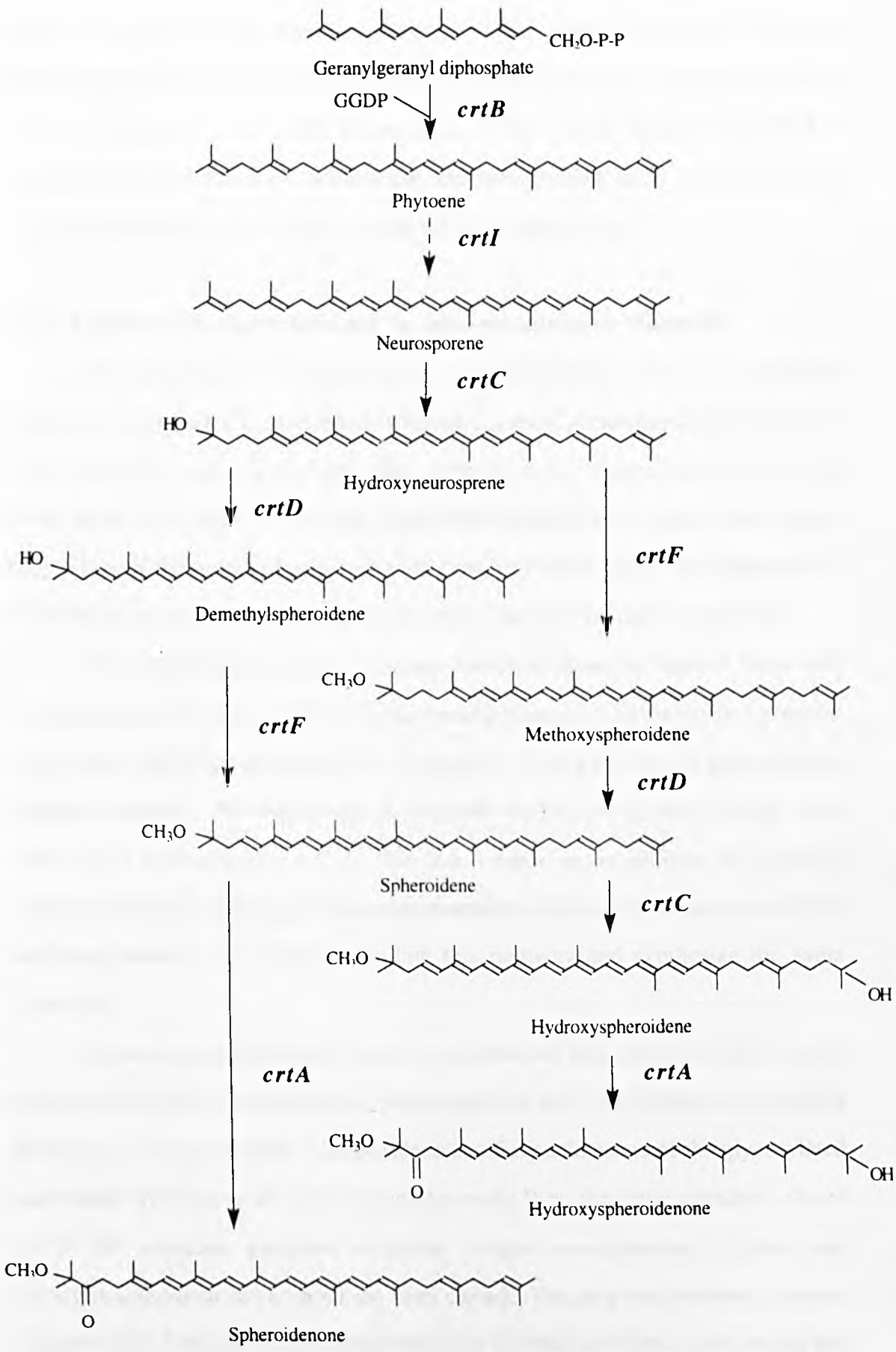


Figure 1.8 The carotenoid biosynthetic pathway of *R. sphaeroides* and *R. capsulatus*. The genes involved are indicated in bold (adapted from Armstrong *et al.*, 1990).

and methylation of the hydroxyl group. The reactions do not proceed in any particular order. No stable mutants from this region grow in the presence of oxygen, implying that demethylspheroidene accumulation and the inability to produce spheroidenone may be toxic (Coomber *et al.*, 1990; Taylor *et al.*, 1983). When oxygen is present the brownish colour of cells due to spheroidene and hydroxyspheroidene changes to red as C-2 oxygenation yields spheroidenone and hydroxyspheroidenone.

1.2.7 Carotenoids biosynthesized in non-phototrophic bacteria

The appearance of carotenoids in non-phototrophic bacteria is sporadic. *Staphylococci* contain C_{30} carotenoids, whereas C_{45} and C_{50} carotenoids are found in a wide variety of bacteria and are often glycosylated. Many pathways for the biosynthesis of carotenoids have been suggested based mainly on carotenoid content. Only a few of the proposed pathways have been verified through the assignment of carotenoid biosynthesis gene function by nucleotide sequencing and mutagenesis.

The biosynthetic pathway of *Erwinia herbicola* shown in Figure 1.9 was first proposed by Hundle *et al.* in 1991. The universal pathway is followed to the formation of phytoene which is subsequently desaturated four times by the crtI type enzyme, yielding lycopene. The end groups of lycopene undergo cyclization and are then modified by hydroxylation at C-3. The last reaction is the addition of glycoside moieties to the C-3 hydroxy groups to give zeaxanthin diglucoside. *Erwinia uredovora* and other strains of *E. herbicola* follow this pathway and synthesize the same carotenoids.

Myxococcus fulvus and *Myxococcus xanthus* are two other non-phototrophic bacteria in which the carotenoid biosynthetic pathway has been partially characterized genetically. The carotenoids 4-ketotorulene and fatty acid esters of the glycosylated carotenoids myxobacton and myxobactin (Appendix I) are the major pigments. Genes for GGDP synthase, phytoene synthase, hydroxyneurosporene synthase and hydroxyneurosporene desaturase have been cloned. The proposed pathway is shown in Figure 1.10. Until all the genes/enzymes of the pathway have been characterized it is not possible to say which pathway/product, if any, is favoured.

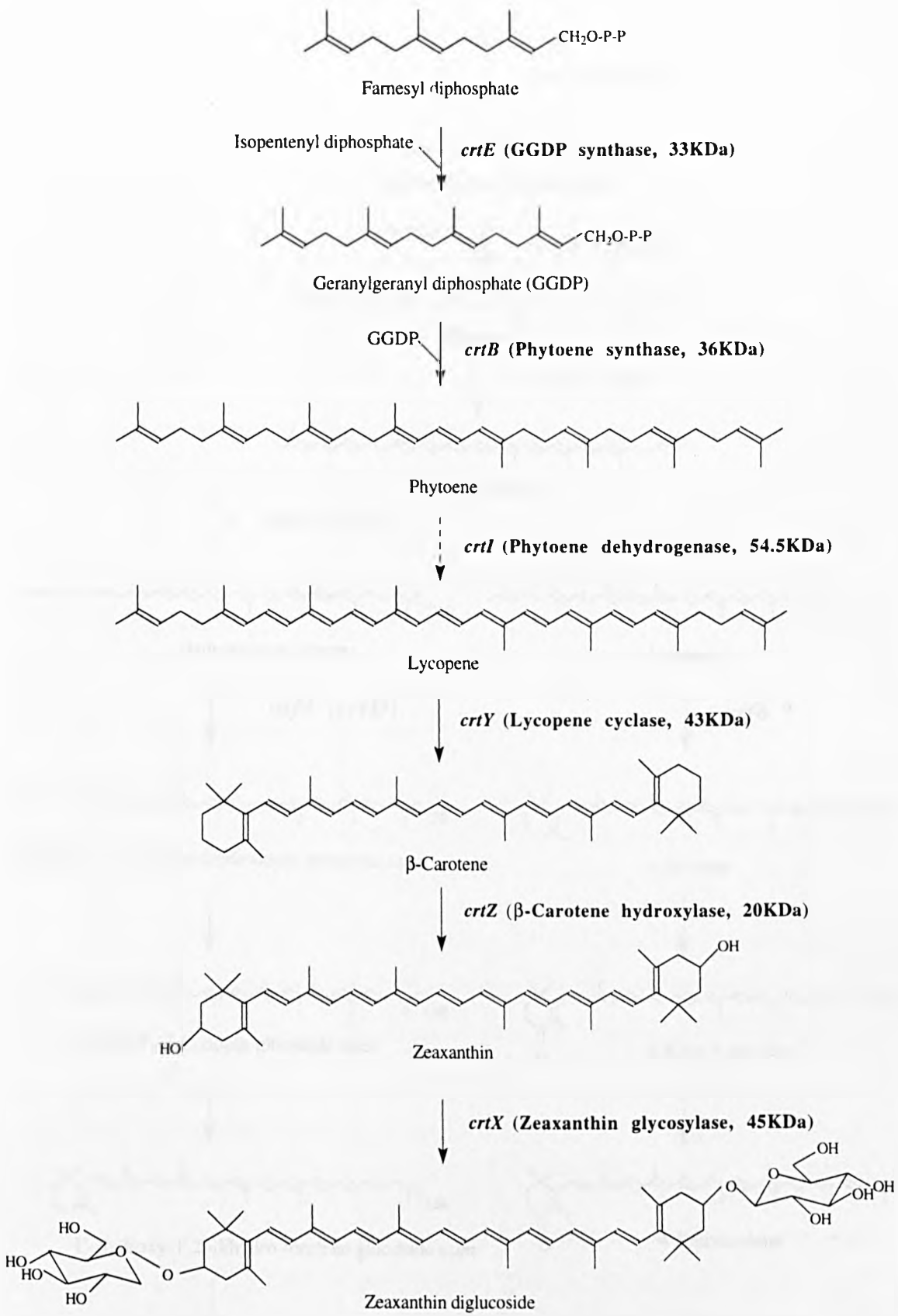


Figure 1.9 The carotenoid biosynthetic pathway of *E. herbicola* and *E. uredovora*. The genes and corresponding enzymes of the pathway are indicated in bold (Hundle *et al.*, 1991).

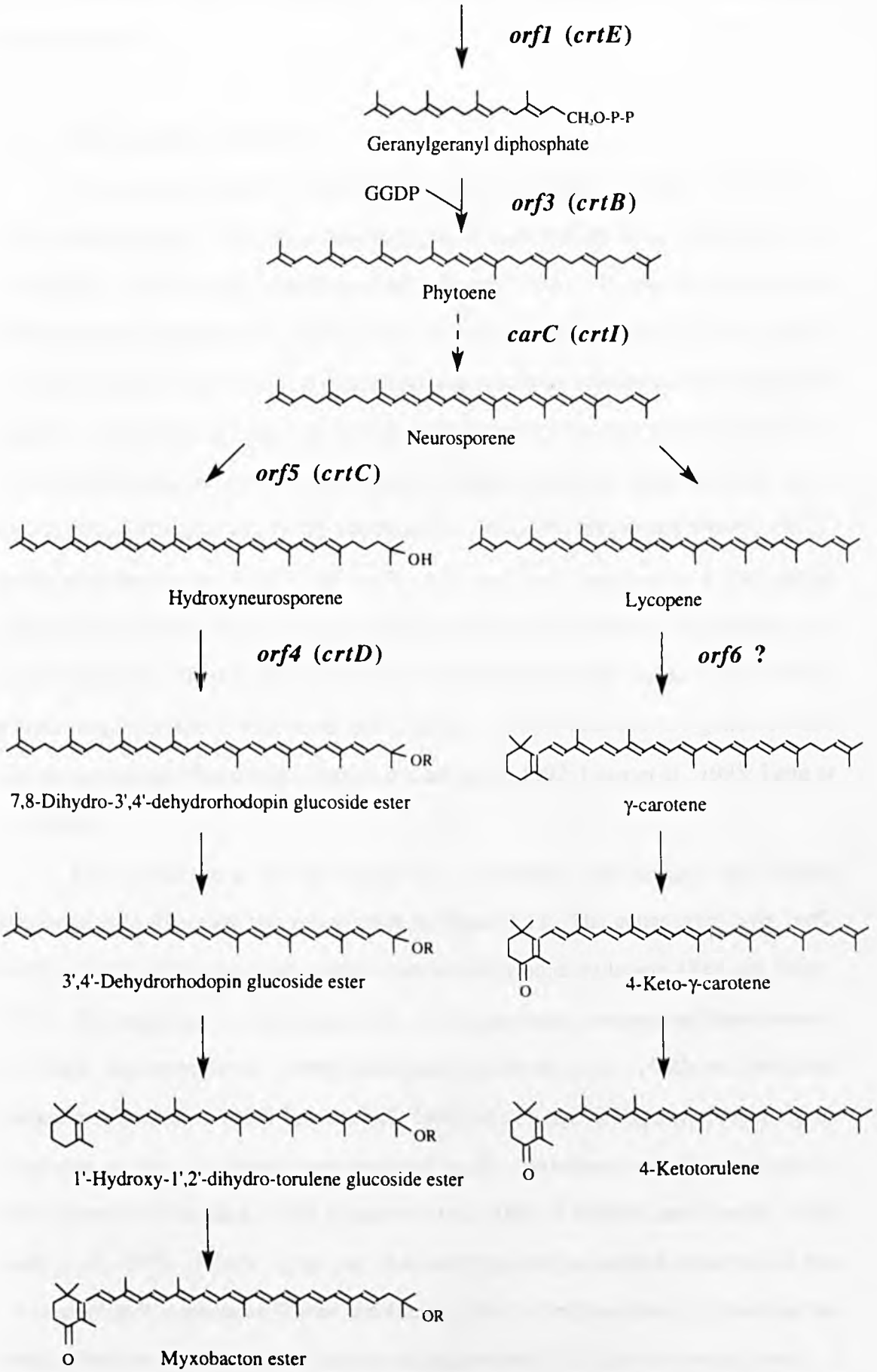


Figure 1.10 Carotenoid biosynthesis pathway of *M. fulvus* and *M. xanthus*. Where possible, genes involved in the pathway are given in bold (Botella *et al.*, 1995).

1.3 Structure and expression of genes involved in carotenoid biosynthesis

1.3.1 Phototrophic bacteria

R. capsulatus and *R. sphaeroides* are two related purple, non-sulphur, phototrophic bacteria. The first carotenoid genetic experiments were performed in the 1950s on *R. sphaeroides* (Griffiths and Stainier, 1956). It was discovered that photosynthetic membranes form and carotenoids accumulate if either the light intensity or oxygen tension is decreased. By genetic manipulation and mapping of photopigment mutants, it was observed that carotenoid and bchl (bacteriochlorophyll) loci were in very close proximity and in some cases overlap, implying superoperonal gene organization (Young *et al.*, 1989; Young *et al.*, 1992; Wellington and Beatty, 1991). Seven clustered genes *crtA*, *crtB*, *crtC*, *crtD*, *crtE*, *crtF* and *crtI* in a 9kb region upstream of the *bchC* gene are responsible for the biosynthesis of spheroidene, or spheroidenone in the presence of oxygen, from FDP. *R. sphaeroides* has a similar genetic organization (Pemberton and Harding, 1986), and the *crt* genes of this organism have now been fully sequenced (Gari *et al.*, 1992; Lang *et al.*, 1995; Lang *et al.*, 1994).

The organization of the genes for carotenoid and bacteriochlorophyll biosynthesis in *Rhodobacter* sp. are given in Figure 1.11. The genes *crtA*, *crtB*, *crtC*, *crtD*, and *crtE* of *R. capsulatus* were first identified by point mutations (Yen and Marrs, 1976). By sequencing (Armstrong *et al.*, 1989) alongside, transposon (Kaufmann *et al.*, 1984; Armstrong *et al.*, 1990), interposon (Guiliano *et al.*, 1988) and mediated marker rescue (Marrs, 1981; Taylor *et al.*, 1983) of a cloned 46kb photosynthetic gene supercluster, the *crt* genes, were mapped in *R. capsulatus* and *R. sphaeroides* (Pemberton and Harding, 1986; Coomber *et al.*, 1990; Coomber and Hunter, 1989; Lang *et al.*, 1995). Another gene *crtJ* that was physically separated from the rest was discovered in *R. capsulatus* (Zsebo and Hearst, 1984). *CrtJ* encodes a product that has similar features to transcription factors and suppresses Bchl and carotenoid levels. A *R. sphaeroides* gene *pps* located 11Kb from *crtA* induced *trans* suppression of

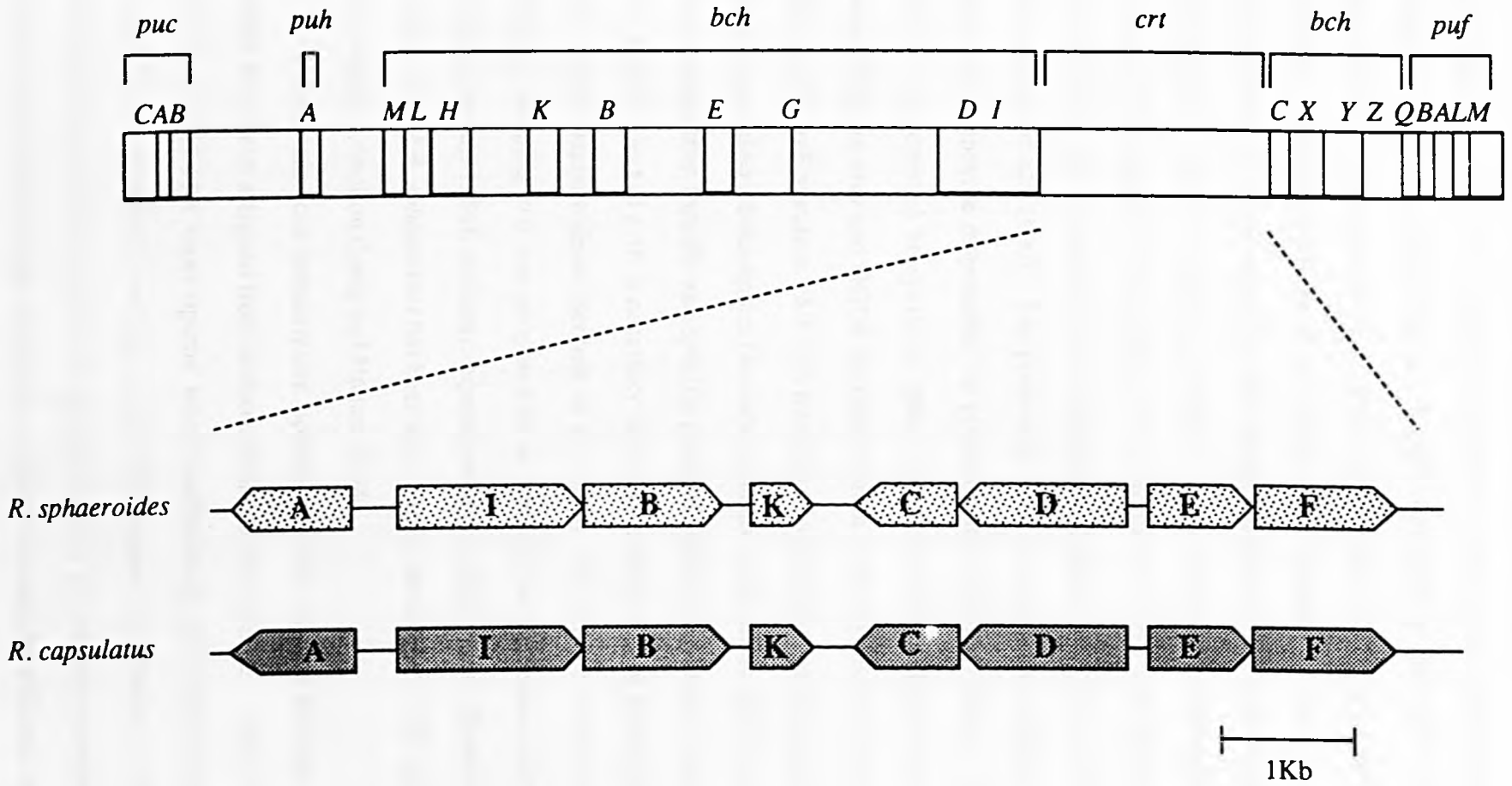


Figure 1.11 The organization of carotenoid biosynthesis genes within the photosynthetic gene cluster of *Rhodobacter* species. The two clusters are identical in terms of organization and direction of transcription (adapted from Lang *et al.*, 1995; Armstrong *et al.*, 1990).

bacteriochlorophyll and carotenoid levels in both *R. sphaeroides* and *R. capsulatus* (Penfold and Pemberton, 1991). Later characterization of the gene showed that it was homologous to *crtJ* of *R. capsulatus* and it was proposed to act as a transcriptional repressor (Penfold and Pemberton, 1994). Sequencing identified another ORF, designated *crtK* (Armstrong *et al.*, 1989), which had also been subjected to earlier mutation analysis (Guiliano *et al.*, 1988). The protein product of *crtK* is very hydrophobic and, along with CrtC, was thought to be involved in the hydroxylation of neurosporene. CrtK is thought possibly to be a membrane attachment point for the enzymes of carotenoid biosynthesis. The products of *crtB* and possibly *crtJ* were originally identified as prephytoene diphosphate synthase and *crtE* as GGDP synthase (Armstrong *et al.*, 1990). The product of *crtB* has now been shown to encode a bifunctional enzyme responsible for phytoene and GGDP synthases, so *crtE* is not strictly a carotenoid biosynthesis gene. Mutations in *crtB* and *crtE* allow Bchl biosynthesis *in vivo* and GGDP accumulation *in vitro* suggesting an active pathway through the *crtE* mutation. All *crtE* mutations have Bchl-containing pigment-protein complexes and so a branchpoint between carotenoid and chlorophyll biosynthesis may occur earlier than thought and specific pools of GGDP may exist (Armstrong, 1994). The gene product of *crtF* is necessary for the O-methylation of hydroxyneurosporene and demethylspheroidene (Scolnik *et al.*, 1980). By genetic complementation and deletion mapping *crtI* was proposed to be involved in the desaturation of phytoene (Guiliano *et al.*, 1986), and later sequence analysis verified this (Bartley and Scolnik, 1989). CrtI of *R. sphaeroides* has been shown to be involved in LHII (light-harvesting II) complex formation (Lang and Hunter, 1994).

There are four transcription operons, A, IB, DC and EF, the directions of which have been assigned from deduced amino acid sequences. The last of these, EF, also forms part of a 'super-operon' where read-through transcription can occur, from *crtEF* through *bchCA* and *puf* genes (Wellington and Beatty, 1991). Further characterization of this operon, by growth studies of interposon mutants, led to the proposal that read-through transcription was necessary for efficient transition from respiratory to photosynthetic growth (Wellington *et al.*, 1991). A moderate increase in

promoter activity of A and EF and levels of mRNA for the genes *crtA, crtC, crtD, crtE* and *crtF* are observed in anaerobiosis, but IB is unaffected, reflecting an increased demand for carotenoid biosynthesis (Armstrong *et al.*, 1993; Armstrong, 1994). The expression of *crt* genes of *R. capsulatus* was increased following an increase in light intensity in the presence of oxygen (Zhu and Hearst, 1986). A shift from anaerobic to aerobic growth conditions resulted in an increase in *crt* gene expression, particularly *crtA* responsible for the formation of spheroidenone (Zhu *et al.*, 1986). Upstream sequence analysis of the *crt* operons identified promoter and regulatory sequences. These were suggested to be responsible for regulating oxygen-dependent pigment biosynthesis (Ma *et al.*, 1993). Recently an ORF has been characterized and shown to encode an aerobic repressor that coordinately regulates bacteriochlorophyll and LHII expression as well as carotenoid (Ponnampalam *et al.*, 1995).

R. capsulatus genes transformed into *E. coli* do not allow *E. coli* to produce carotenoids. However when the genes of *R. sphaeroides* were transformed into the un-related, non-carotenogenic bacteria *Paracoccus denitrificans, Agrobacterium tumefaciens, Agrobacterium radiobacter* and *Azotomon insolita* carotenoid production was observed (Pemberton and Harding, 1987).

1.3.2 Non-phototrophic bacteria

(i) *Erwinia herbicola* and *Erwinia uredovora*

E. herbicola is a non-photosynthetic, Gram-negative bacterium of the family Enterobacteriaceae. These bacteria are often yellow and are isolated from a wide variety of organisms as non-pathogenic epiphytes. These include the leaves and buds of plants, soil and water (Starr, 1981; Billing and Baker, 1963).

The yellow pigments of *E. herbicola* were shown to be carotenoids (Perry *et al.*, 1986; Starr, 1981), the main ones being β-cryptoxanthin and mono- and di-glucosides of zeaxanthin. The carotenoids of *E. herbicola* and the transformed *E. coli* strain carrying the pigment genes were analysed. Tuveson *et al.* (1988) provided the first evidence that the pigments were polar carotenoids. *E. coli* was transformed with

the plasmid pPL376 carrying the carotenoid biosynthesis genes of *E. herbicola*. The carotenoids present offered a protective role against near-uv irradiation and phototoxic molecules where the membrane is a lethal target (Tuveson *et al.*, 1988) Chemical structures of the carotenoids and their glycosylated derivatives were elucidated (Sandmann *et al.*, 1990; Hundle *et al.*, 1991) and the biosynthetic pathway shown in Figure 1.9 was also proposed (Hundle *et al.*, 1991).

In some *E. herbicola* strains carotenoid production is plasmid encoded. By growing *E. herbicola* Y46 at elevated temperatures non-pigmented variants were induced (Chatterjee and Gibbins, 1971). This effect was irreversible and was attributed to the loss of endogenous plasmids. However the carotenoid biosynthesis genes of *E. herbicola* Eho10 and Eho13 (Perry *et al.*, 1986; Lee and Liu, 1991) and *Erwinia uredovora* (Misawa *et al.*, 1990) are chromosomally encoded. The carotenoid biosynthesis genes of *E. herbicola* Eho10 were the first of a non-photosynthetic bacterium to be cloned (Perry *et al.*, 1986). The genes for pigment production lay within a 12.4Kb chromosomal fragment. Carotenoid production was repressed by glucose in *E. herbicola* and the pigmented *E. coli* clone and cyclic AMP (cAMP) was suggested to play a role in the regulation of carotenoid gene expression.

By sequencing, mutagenesis and the identification of carotenoid intermediates carotenoid biosynthesis genes were identified. Expression of the *crt* genes of *E. herbicola* in *E. coli* and *Agrobacterium tumefaciens* has also allowed the assignment of gene function. The nucleotide sequences of three genes from the carotenoid biosynthetic pathway of *E. herbicola* were reported in 1990 (Armstrong *et al.*, 1990). They encoded homologues of the CrtB (prephytoene diphosphate synthase), CrtE (phytoene synthase) and CrtI (phytoene dehydrogenase) enzymes of *R. capsulatus*. However the gene functions of *crtB* and *crtE* have been reassigned. The *crtE* gene encodes GGDP synthase and *crtB* encodes phytoene synthase (Math *et al.*, 1992; Sandmann and Misawa, 1992).

The functional assignment of all the carotenoid biosynthesis gene cluster has now been reported (Hundle *et al.*, 1994). The six carotenoid genes are arranged in at least two operons (Figure 1.12). The gene *crtZ* is transcribed in the opposite direction

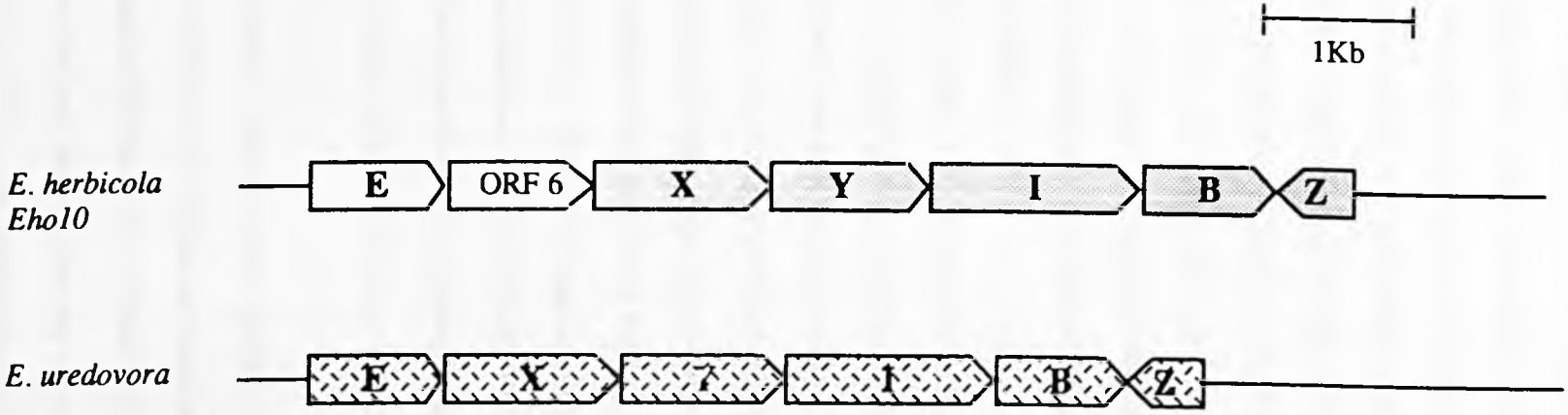


Figure 1.12 The organization of carotenoid biosynthesis genes in *Erwinia* species. The two clusters are identical with the exception of an ORF6 in *E. hebicola* not present in *E. uredovora*. All genes are transcribed in the same direction apart from *crtZ* (Armstrong, 1994).

to the other five and overlaps *crtB* by 60bp. An almost identical cluster is found in *E. uredovora* with the exception that ORF6 occurs only in *E. herbicola* (Misawa *et al.*, 1990) and in another strain of *E. herbicola* Eho13 (Liu, 1993; To *et al.*, 1994). Eho13 showed a higher percentage homology with *E. uredovora* than Eho10. The cluster of Eho13 has been suggested for use as a genetic marker in gene cloning. It is proposed to function in a similar way to blue-white screening, in that recombinant clones will appear white when grown on agar plates (Liu, 1993).

The expression of *E. herbicola crtI* in a *R. sphaeroides crtI* mutant results in novel xanthopyhll production (Hunter *et al.*, 1994). *E. uredovora* carotenoid genes expressed in *Zymomonas mobilis* and *A. tumefaciens* results in β-carotene accumulation and zeaxanthin diglucoside production (Misawa *et al.*, 1991; Nakagawa and Misawa, 1991).

The *crtI* gene of *E. uredovora* has been cloned and over-expressed in *E. coli*. The protein was purified from inclusion bodies and activity was restored upon removal of urea and FAD was required as a cofactor for activity (Fraser *et al.*, 1992). In the same way CrtE has been purified and characterized (Weidemann *et al.*, 1993). The gene *crtI* expressed in transgenic plants conferred resistance to herbicides and increased β-carotene biosynthesis as well as altering xanthophyll metabolism (Misawa *et al.*, 1993; Misawa *et al.*, 1994). Resistance to herbicides was also apparent when *crtI* was expressed in *Synechococcus* PCC7942 (Windhovel *et al.*, 1994b). Recently lycopene cyclase has been purified and characterized after over-expression in *E. coli*. NADH or NADPH was required as a cofactor for enzyme activity (Schnurr *et al.*, 1996).

(ii) *Myxococcus xanthus* and *Myxococcus fulvus*

M. xanthus and *M. fulvus* are Gram negative, non-photosynthetic bacteria and are dependent on illumination with blue light for carotenoid synthesis. The carotenoids myxobacton ester and ketotorulene are the main products of a branched pathway in *M. fulvus* (Kleinig, 1975). The same pathway has been confirmed in *M. xanthus* by the analysis of carotenoids accumulated by various mutant strains (Ruiz-Vazquez, 1993). The pathway is shown in Figure 1.10 with the various gene assignments indicated. A

cluster of genes responsible for early and late stages of carotenoid synthesis lies between the loci *carA* and *carB*. The assignment of gene function within this cluster has tentatively been made (Botella *et al.*, 1995). The cluster is shown in Figure 1.13. The clustered genes are transcribed from a single light-inducible promoter and are likely to form a single operon. The gene product of *carB* is responsible for phytoene synthesis and is under the control of the *cis*-acting gene *carA* which represses the promoter in the dark (Martinez-laborda 1990). *CarC* is not linked to the *carA-carB* region and encodes phytoene dehydrogenase. It is also regulated by light, but the promoter only becomes activated in stationary-phase growth conditions or carbon starvation (Fontes, 1993). The *carC* promoter can also become stimulated by CarA in the light. Two other unlinked loci *carR* and *carD* are involved in the regulation of carotenoid biosynthesis. The *carR* locus involves three translationally coupled genes *carQ*, *carR*, and *carS* encoding a positive regulator, negative regulator and *trans*-acting activator respectively (McGowan, 1993). Mutations in *carR* result in the production of carotenoids in the absence of light, due to *carB* becoming light independent. *carQ* mutations result in the blocking of *carB* and consequently no carotenoids are produced either in the light or dark. The mutations have the same effect on *carC*. The other locus *carD* is required for the activation of the *carQRS* and *carC* promoters by blue light (Nicolas, 1994). CarD is also thought to be involved in the expression of development-specific genes. The control of carotenoid biosynthesis in *M. xanthus* is complicated (Figure 1.14) and may also involve the interaction of singlet oxygen with *carR* to initiate a cascade involving *carQ*, *S* and *D*. Recently *carQ* was proposed to encode a sigma factor and *carR* a membrane associated anti-sigma factor that regulate the *carQRS* operon (Gorham *et al.*, 1996). CarR is thought to sequester CarQ to the membrane in the dark, but upon illumination loss of CarR results in CarQ release and subsequent carotenogenesis.

(iii) Others

A cluster of five genes from *Agrobacterium aurantiacum* responsible for the biosynthesis of astaxanthin have recently been isolated and characterized (Misawa *et*

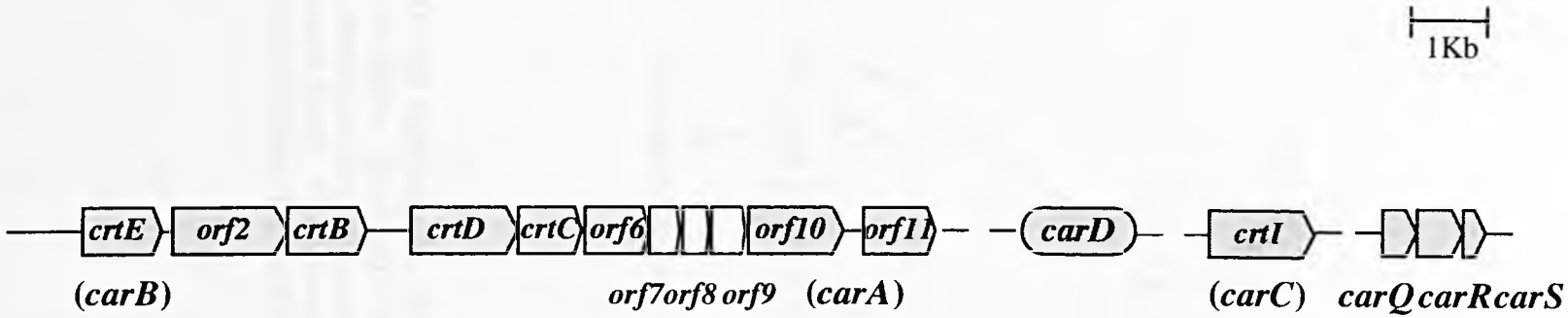


Figure 1.13 The organization of the carotenoid biosynthesis genes of *M. xanthus*. Where known the direction of transcription is indicated. Genes involved in carotenoid biosynthesis are indicated with shaded boxes. Previous nomenclature of the *crt* genes is given in parentheses. Oval boxes of arbitrary size indicate cloned genes of unpublished sequence and hatched lines indicate unknown distances between genes. *orf2* and *orf6* are thought to be involved in desaturation and cyclization reactions respectively, and *orf10* and *orf11* (previously *carA*) encode regulatory proteins (adapted from Botella *et al.*, 1995; Armstrong, 1994).

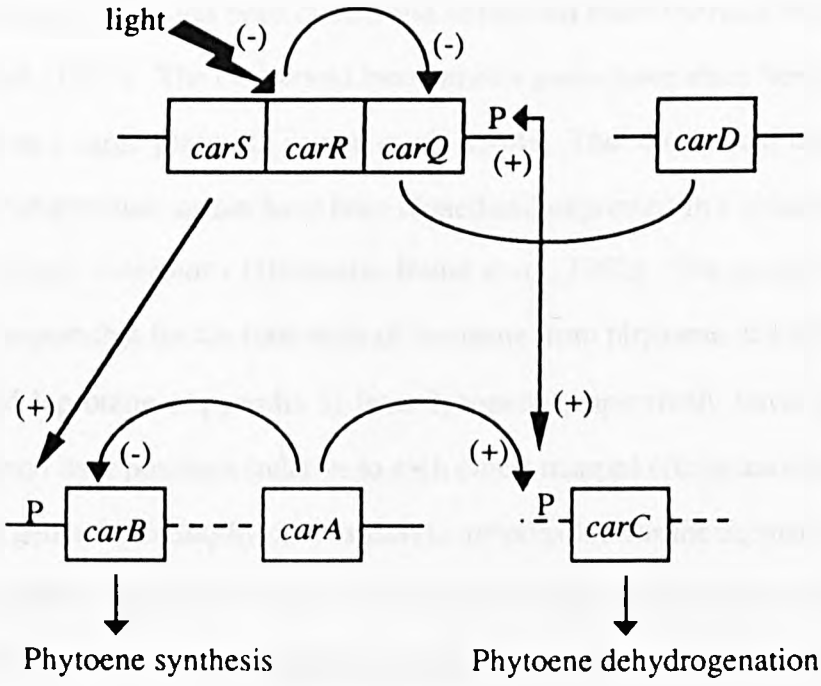


Figure 1.14 The light regulon of *M. xanthus*. + and - indicate effects of light on promoters or genes. Blue light has an inhibitory effect on the protein product of *carR*, This results in CarQ stimulating the promoter of *carQRS* and *carC* through its interaction with CarD. The activation of the *carQRS* operon causes high expression of CarS which activates the *carBA* operon promoter, (adapted from Nicolas *et. al.*, 1995).

al., 1995). The biosynthetic pathway of astaxanthin production has also been proposed (Yokoyama and Miki, 1995). The genes *crtY*, *crtI* and *crtB* are organized in the same way as in *Erwinia* sp. and *Rhodobacter* sp. However *crtZ* is transcribed before *crtY*, and *crtW* before *crtZ*. The gene products of *crtZ* and *crtW* have low substrate specificity, and allow the production of numerous pathway intermediates. Various other carotenoid biosynthesis genes have been characterized from isolated bacteria. Phytoene synthase (*crtB*) has been cloned and sequenced from *Thermus thermophilus* (Hoshino *et al.*, 1993). The carotenoid biosynthesis genes have since been shown to be clustered on a large plasmid (Tabata *et al.*, 1994). The carotenoid biosynthesis genes of *Mycobacterium aurum* have been cloned and expressed in a colourless strain of *Mycobacterium smegmatis* (Houssaini-Iraqui *et al.*, 1992). The genes *crtI*, *carα*, and *carLep* responsible for the formation of lycopene from phytoene, α-carotene from lycopene and leprotene (Appendix I) from lycopene respectively have since been charcterized and their positions (relative to each other) mapped (Houssaini-Iraqui *et al.*, 1993). Two genes from *Staphylococus aureus* responsible for the formation of 4,4'-diaponeurosporene (Appendix I) have been cloned and analysed (Wieland *et al.*, 1994). The gene *crtM* encodes dehydrosqualene synthase and *crtN* encodes dehydrosqualene desaturase. Carotenoid biosynthesis in *Streptomyces setonii* can be induced by the presence of a gene *crtS*. This gene has been characterized and encodes a sigma factor which is necessary for carotenoid synthesis (Kato *et al.*, 1995).

1.3.3 Cyanobacteria

Cyanobacteria, commonly known as blue-green algae, are a species of bacteria that have an oxygen-evolving photosynthetic system that is similar to that of higher plants. Carotenoids occur universally, the major pigments synthesized being β-carotene, zeaxanthin, echinenone (Appendix I) and myxoxanthophyll (Appendix I). By using various genetic and biochemical experiments involving herbicides, genes responsible for the biosynthesis of β-carotene from GGDP have been isolated. The gene encoding phytoene desaturase has been isolated and sequenced from *Aphanocapsa* PCC6714 (Schmidt and Sandmann, 1990), *Synechococcus* PCC7942 (Chamovitz *et*

al., 1991) and *Synechocystis* sp. PCC6803 (Martinez-Ferez and Vioque, 1992). The enzyme from *Synechococcus* PCC7942 has been overexpressed and characterized in *E. coli* and was shown to catalyse the conversion of phytoene into ζ-carotene (Fraser *et al.*, 1993). The gene *psy* encoding phytoene synthase has also been cloned and expressed in *E. coli* from *Synechococcus* PCC7942 (Chamovitz *et al.*, 1992) and recently from *Synechocystis* PCC6803 (Martinez-Ferez *et al.*, 1994). The sequences of the enzymes showed a high degree of homology to each other and to the tomato cDNA sequence, pTOM5. They also showed conserved domains to other bacterial phytoene synthase enzymes. Heterologous complementation experiments allowed the cloning of a gene from *Anabaena* PCC7120 encoding ζ-carotene desaturase (Linden *et al.*, 1993). Subsequent characterization, by overexpressing the protein in *E. coli*, found high sequence similarity with phytoene desaturase genes from bacteria and *al-1* from *N. crassa*. Much lower similarity was observed with phytoene desaturases from other cyanobacteria and higher plants (Linden *et al.*, 1994). Chemically induced mutants resistant to the herbicide MPTA, which inhibits the cyclization of lycopene, were isolated from *Synechococcus* PCC7942. The resistance was mapped to a gene designated *crtL* (*lcy*) encoding lycopene cyclase (Cunningham *et al.*, 1993). Sequence analysis showed little similarity to the other bacterial lycopene cyclases but was highly homologous to the corresponding enzyme in plants and algae (Cunningham *et al.*, 1994). The genes *psy* and *crtL* are organized in the same way as their bacterial counterparts *crtB* and *crtY* (Armstrong, 1994).

1.3.4 Plants

(i) *Lycopersicon esculentum* (tomato)

The enzyme phytoene synthase was originally isolated as a cDNA clone TOM5 encoding a protein of 47KDa that accumulated during fruit ripening (Ray *et al.*, 1987). The sequence showed homology to the *crtB* genes of *E. herbicola* and *R. sphaeroides* (Armstrong *et al.*, 1990). The protein product of TOM5 has an N-terminal domain of 120 amino acids that is not present in the bacterial equivalents (Bartley *et al.*, 1992).

Transgenic plants expressing an antisense RNA to TOM5 showed significant reduction in carotenoid levels in fruit and flowers (Bird *et al.*, 1991). The block in carotenoid biosynthesis was between GGDP and phytoene. Leaf carotenoid levels were not affected by this block, suggesting that another enzyme is involved in leaf carotenoid biosynthesis (Bramley *et al.*, 1992). The cDNA complemented a *crtB* mutation in *R. sphaeroides* and was proposed to encode phytoene synthase (Bartley *et al.*, 1992). The cDNA of phytoene synthase, expressed in *E. coli* with a plasmid containing an *E. uredovora crtB* mutation, compensated for the mutation resulting in phytoene synthesis (Misawa *et al.*, 1994). A second phytoene synthase cDNA designated *psy2* has been cloned and shown to be abundant in leaves (Bartley and Scolnik, 1993). The genomic sequences have been cloned for both synthases. GTOM5 (*psy1*) contains 5 introns and 6 exons spanning 3.6Kb. In the clone GTOMF (*psy2*) the first exon of *psy1* is split into 3 exons (Ray *et al.*, 1992). The cDNA of the enzyme phytoene desaturase has been cloned and sequenced. The deduced amino acid sequence showed conservation with the equivalent enzymes of cyanobacteria and algae, and appeared unrelated to the purple bacteria and fungal equivalents (Pecker *et al.*, 1992). The genomic sequence has also been characterized and has been shown to contain 15 exons and 14 introns that span 7.7Kb (Aracri *et al.*, 1994; Mann *et al.*, 1994), and was mapped to chromosome 3 (Guiliano *et al.*, 1993; Scolnik and Guiliano, 1993).

(ii) *Zea mays* (corn)

A putative carotenoid regulatory locus *Yellow1* (*y1*) from maize was tagged by using the *Mu* transposable element (Buckner *et al.*, 1990; Buckner and Robertson, 1993). A cloned *y1* cDNA showed high homology to tomato and bacterial *psy* genes indicating that *y1* encodes the equivalent enzyme in maize, reviewed in (Bartley *et al.*, 1994). Recently a cDNA encoding phytoene desaturase (*pds*) has been cloned and characterized from maize (Li *et al.*, 1996). Analysis of *pds* transcripts in developing endosperm indicated that it was constitutively expressed, and led to the conclusion that carotenogenesis was not regulated by increasing levels of *pds* transcript.

(iii) *Capsicum annuum* (pepper)

The cDNA encoding GGDP synthase was cloned by use of specific antibodies against the enzyme (Kuntz *et al.*, 1992). The protein contains 369 amino acids, the first 60 of which are a transit peptide for plastid location. Localization to the plastids has also been determined by immunogold cytochemistry (Cheniclet *et al.*, 1992). The cDNA sequence showed similarities to *crtE* of photosynthetic and non-photosynthetic bacteria.

Phytoene synthase has been purified from *Capsicum* chromoplast stroma and shown to catalyse two reactions to yield phytoene from GGDP (Dogbo *et al.*, 1988). The cDNA of pepper phytoene synthase has been cloned from a cDNA expression library by the use of a tomato cDNA as a probe (Romer *et al.*, 1993). Comparison of the two sequences showed high degrees of homology. The first 127 amino acids are involved in transport of the peptide to the plastid.

The phytoene desaturase enzyme has been purified and used to raise antibodies, which were subsequently used to isolate the cDNA encoding the enzyme (Hugueney *et al.*, 1992). The sequence showed high homology to tomato and soybean *pds* and overexpression of the enzyme in *E. coli* resulted in a fully active enzyme. The cDNA encoding ζ-carotene desaturase has been cloned and expressed in *E. coli* (Albrecht *et al.*, 1995). The enzyme showed little sequence similarity to the ζ-carotene desaturase of *Anabaena* compared to 33-35% similarity with plant or cyanobacterial phytoene desaturases.

A cDNA encoding the bifunctional enzyme capsanthin-capsorubin synthase responsible for ketocarotenoid synthesis has recently been cloned using the same antibody techniques as previously (Bouvier *et al.*, 1994). The gene is specifically expressed during chromoplast development and the protein contains a transit peptide for plastid location.

(iv) *Arabidopsis thaliana*

A genomic fragment of the gene encoding GGDP synthase of *A. thaliana* has been cloned and reported to Genbank. A cDNA sequence has also been reported.

Sequence divergence between the cDNA and the genomic fragment suggest that the cDNA is derived from a second copy of the gene (Bartley *et al.*, 1994). The cDNA encoding phytoene desaturase has also been cloned and sequenced (Scolnik and Bartley, 1993). An *immutans* (*im*) variegation mutant of *A. thaliana* containing green and white sectored leaves accumulates phytoene in the white sectors. This suggests that *im* can control phytoene desaturase (Wetzel *et al.*, 1994). Two mutants have been isolated and characterized that affect phytoene desaturation and are defective in plastoquinone and tocopherol biosynthesis. Neither of the non allelic mutations (*pds1* and *pds2*) are in the phytoene desaturase gene (Norris *et al.*, 1995). Plastoquinone was also shown to be an essential component of carotenoid biosynthesis.

(v) Others

A soybean cDNA was cloned and used in complementation experiments with *R. capsulatus*. It encoded phytoene desaturase and like other plant desaturases, is synthesized as a precursor that is processed to the mature form upon import into chloroplasts. It also contained a dinucleotide binding motif characteristic of other desaturase enzymes (Bartley *et al.*, 1991).

A cDNA clone has recently been isolated from a melon fruit cDNA library by use of the tomato cDNA clone TOM5 as a heterologous probe. The cDNA encoded a protein homologous to other plant phytoene synthases with a high degree of conservation (Karvouni *et al.*, 1995).

1.3.5 Fungi

Three gene sequences that are involved in carotenoid biosynthesis in the fungus *N. crassa*, have been cloned. These are *al-1*, *al-2* and *al-3* which encode phytoene dehydrogenase, phytoene synthase and GGDP synthase, respectively (Schmidhauser *et al.*, 1990; Carattoli *et al.*, 1992; Schmidhauser *et al.*, 1994; Nelson *et al.*, 1989). The *al-3* gene was cloned by Nelson *et al.* (1989) and mRNA was shown to increase 15-fold in light-induced (compared to that of dark grown) mycelia. Further analysis indicated that the gene was photo-regulated (Baima *et al.*, 1991). Photo-induction of

al-3 is also apparent in the conidia (vegetative spores) of *N. crassa* (Baima *et al.*, 1992). Comparison of the GGDP synthase with other prenyl transferases, from bacteria to higher eukaryotes, revealed the presence of three highly homologous regions probably indicative of the catalytic site of the enzyme (Carattoli *et al.*, 1992). The functional identity of *al-3* was confirmed by the use of *in vitro* assays and genetic complementation experiments with the *crt* genes of *E. uredovora* (Sandmann *et al.*, 1993). Recent manipulations of the upstream regions of the gene have demonstrated a sequence between positions -226 and -55 that is necessary for photoinduction. The sequence contained multiple regulatory elements, one termed APE which is also present in the promoter region of *al-1* (Carattoli *et al.*, 1994).

Characterization of *al-2* revealed homologous regions to other phytoene synthases and a photo-inducible expression. Levels of *al-2* mRNA increased 30-fold in photoinduced mycelia when compared to dark grown (Schmidhauser *et al.*, 1994). The gene *al-1* encoding phytoene dehydrogenase showed homology to other dehydrogenases and was photo-inducible with levels of mRNA increasing by 70-fold (Schmidhauser *et al.*, 1990). The levels of *al-1* and *al-2* mRNA accumulate in conidia initially independent of the light environment. It was observed that light induced additional accumulation of *al-1* and *al-2* (Li and Schmidhauser, 1995). These results indicate that *al* gene expression is under developmental control as well as photoregulation. Inactivation of *al-1* and *al-3* was demonstrated when the fungus was transformed with different portions of the *al-1* or *al-3* genes (Romano and Macino, 1992; Cogoni *et al.*, 1994). Expression of the endogenous *al-1* or *al-3* genes was impaired, resulting in albino phenotypes, but was found to be spontaneously and progressively reversible. This phenomenon was termed 'quelling' and, once reversed, could not occur again despite the continuing presence of exogenous *al-1* or *al-3*.

A gene encoding phytoene dehydrogenase has been cloned, sequenced and characterized from *Cercopsora nicotianae*. The protein showed 50% homology with the *N. crassa* equivalent. Northern analysis of mRNA accumulated in light and dark growth showed no differential expression of phytoene dehydrogenase in either environment (Ehrenshaft and Daub, 1994).

1.3.6 Algae

A gene from the alga *Cyanophora paradoxa* encoding a protein similar to CrtE of *R. capsulatus* has been cloned and sequenced. The gene is found in the plastid (cyanelle) DNA (Michalowski *et al.*, 1991). A high resolution genomic mapping of the chloroplast genome of the red alga *Porphyra purpurea* has been made and *crtE* was present (Reith and Munholland, 1993). More recently a cDNA responsible for the production of canthaxanthin form β-carotene in *Haematococcus pluvialis* has been cloned and designated β-C-4 oxygenase ("β-carotene ketolase") (Lotan and Hirschberg, 1995).

1.4 Aims of the work described in this thesis

The aims of this work were to contribute to the understanding of the molecular biology of carotenoid biosynthesis. Chapter 3 describes the genetic manipulation of the *crt* genes of *E. herbicola* to facilitate increased carotenoid production. Localization of the enzymes is described in Chapter 4. The genes of *E. herbicola* were used as heterologous probes to obtain other *crt* sequences in non-photosynthetic bacteria, the details of which are described in Chapters 5 and 6. The carotenoid contents of these bacteria (previously uncharacterized) were also analysed. Chapter 7 describes the screening of a genomic library of the photosynthetic bacterium *Rhodopseudomonas acidophila* for carotenoid biosynthesis genes.

CHAPTER 2

MATERIALS AND METHODS

2.1 Materials

Chemicals and materials were purchased from Sigma Chemical Company Ltd. or BDH Chemicals Ltd. and stored as recommended by the manufacturer. AnalaR® or equivalent grade products were used unless otherwise stated in the text.

2.1.1 Solutions, media and materials

All growth media and solutions were sterilized by autoclaving ($15 lbs in^{-2}$, 20 minutes) or micropore filtration (0.45µm-Sartorius). Equipment involved in RNA manipulations was treated with 0.1% DEPC water for 12 hours prior to sterilization.

2.1.2 Growth media

The following media were used in this work:-

(i) Nutrient medium:

(15g/l nutrient broth-Oxoid).

(ii) Nutrient agar:

Nutrient medium + Bacto-agar (Difco) [1.5% (w/v)].

(iii) Luria-Bertani (LB) medium:

1% (w/v) NaCl; 1% (w/v) Bacto-tryptone (Difco) and 0.5% (w/v) Bacto-yeast extract (Difco). The pH was adjusted to 7.5 with 5M NaOH.

(iv) M9 minimal medium:

6g/l Na_2HPO_4; 3g/l KH_2PO_4; 0.5g/l NaCl: 1g/l NH_4Cl. These components were dissolved in ddH_2O. Aliquots of 190ml were sterilized by autoclaving. To make

complete M9 medium the following previously sterilised solutions were added: 8ml of 20% (w/v) glucose; 0.2ml of 1M $MgSO_4$; 0.2ml of 1M $CaCl_2$ and 100ml of 0.1% (w/v) thiamine. Methionine assay medium (Difco) was added to a final concentration of 0.25% (w/v).

(v) Tryptone (TB) top agar:

1% (w/v) Bacto-tryptone (Difco); 0.5% (w/v) NaCl and 0.8% Bacto-agar (Difco). After sterilization, 1M $MgSO_4$ was added to a final concentration of 1%.

2.1.3 Antibiotics

Stock solutions of the following antibiotics were made up in sterile ddH_2O (unless otherwise indicated) and stored at 4°C for up to 2 weeks: ampicillin (10mg/ml), kanamycin (20mg/ml), chloramphenicol (20mg/ml in 50% ethanol) and tetracycline (20mg/ml in 50% ethanol). All antibiotics were used at a final concentration of 25µg/ml. When required antibiotics were added to media (55°C) to ensure selective growth and maintenance of plasmids.

2.1.4 Bacterial strains

The bacterial strains used in this study are listed in Table 2.1. Strains were propagated on agar plates containing appropriate antibiotics and maintained at 4°C. Subcultures were made every 4 to 6 weeks. For long term storage, cultures in nutrient broth containing 20% (v/v) glycerol were frozen at -70°C.

2.1.5 Plasmids

The plasmids used and constructed in this work are listed in Table 2.2.

Table 2.1 Bacterial strains

Srain	Genotype/Characteristics	Source/Reference
Escherichia coli::		
DH5α	*sup44 ΔlacU169 (ϕ80 lacZ ΔM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1*	Hanahan, 1983
BL21(DE3)	*hsdS gal (λcI^{s}857 ind1 S$_{am}$7 nin5) lacUV5-T7gene1*	Studier and Moffatt, 1986
LE392	*supE44 supF58 hsdR514 galK2 galT22 metB1 trpR55 lacY1*	Promega
KW251	F^{-} *supE44 galK2 gal22 metB1 hsdR2 mcrB1 mcrA^{-} argA81*:Tn10 *rec Δ1014*	Promega
NM621	F^{-} *hsdR* (r$_K^{-}$m$_K^{+}$) *mcrA mcrB supE44 recD1009*	R. Cogdell
ZK126	Parent K12 strain of ZK1000, *W3110 ΔlacU169 tna2*	J. Pratt
ZK1000	*W3110 ΔlacU169 tna2*, rpoS::kan	J. Pratt
Brevibacterium linens		NCIMB 8546
Micrococcus luteus		J. R. Saunders
Rhodococcus ruber		NCIMB 11149
Sphingobacterium multivorum	(formerly *Flavobacterium*)	NCIMB 12557

Table 2.2 Plasmids

Plasmid	Antibiotic	Characteristics	Source/Reference
pAPU211	Amp^r	10kbp Sph1-EcoR1 fragment carrying crt genes from pPL376 (Perry *et al*, 1986) in pUC19	B. S. Hundle, 1994
pAPU211Z	Amp^r	as above with *crtZ* deleted	B. S. Hundle, 1994
pUC19	Amp^r	pBR322/M13mp19 derivative	Yanisch-Perron *et al.*, 1985
pRG1	Kan^r	Derived from pACYC177 carrying the $lacI^q$ gene	Griffin and Kolodner, 1990
pT7-4	Amp^r	Derived from pBR325 carrying the T7 promoter	Tabor and Richardson, 1985
pT7-5	Amp^r	Derived from pBR325 carrying the T7 promoter	Tabor and Richardson, 1985
pT7-6	Amp^r	Derived from pBR325 carrying the T7 promoter	Tabor and Richardson, 1985
pT7-CH2	Amp^r	*Hin*dIII-*Eco*RI fragment from pAPU211 with *crtZ* under T7 promoter, in pT7-5	This study
pT7-CH3	Amp^r	*Hin*dIII-*Eco*RI fragment from pAPU211 with 5 *crt* genes under T7 promoter, in pT7-5	This study
pCH4	Amp^r	*crtZ* in pUC19	This study
pT7-CH5	Amp^r	*crt Z* from pCH4 under T7 in pT7-6	This study
pCH6	Amp^r	*crtZ* cloned in pAPU211Z such that all *crt* genes are transcribed in the	This study

		same direction	
pT7-CH7	Amp^r	*crt* cluster from pCH6 under T7 in pT7-4	This study
pSCN6-20	Amp^r, Cm^r Tet^r	*crt* and *Bchl* genes from *R. sphaeroides* in pSUP202	Coomber and Hunter, 1989

2.2 Methods

2.2.1 Rapid $CaCl_2$ transformation procedure

Cultures to be transformed were propagated at 37°C until an A_{550} of 0.4 to 0.5 was reached. The culture was chilled on ice for 20 minutes, transferred to a centrifuge tube and pelleted by centrifugation for 5 minutes at 2500g (4°C). The cell pellet was resuspended in 0.5 volumes of ice-cold 0.1M $CaCl_2$ and immediately pelleted by centrifugation for 10 minutes at 2500g (4°C). The cell pellet was resuspended in $^1/_{20}$th the volume of ice-cold 0.1M $CaCl_2$ and kept on ice for at least one hour to allow the cells to become competent.

For each transformation, an aliquot of cells (100µl) was placed in an ice-cold 15ml Falcon polypropylene tube, 2-20ng of DNA in 5-10µl of TE [10mM Tris-HCl; 1mM EDTA; pH7.5] was added to each aliquot of cells and mixed by trituration. One aliquot of cells without added DNA was used as a negative control, whilst 10ng of plasmid DNA was used as a positive control.

The cells were held on ice for 40 minutes and then subjected to heat shock treatment at 42°C (2 minutes). The cells were placed on ice for 2 minutes and subsequently mixed with 900µl of SOC medium [2% (w/v) bacto-tryptone, 0.5% (w/v) bacto-yeast extract, 10mM NaCl and 2.5mM KCl, and after autoclaving, filter sterilized solutions of $MgCl_2$ and $MgSO_4$ added to a final concentration of 10mM. Prior to use a filter sterilized solution of glucose was added to a final concentration of 20mM]. The cells were shaken in an orbital incubator at 37°C for one hour and then spread on solid medium containing the appropriate antibiotic. The plates were incubated routinely at 37°C overnight.

2.2.2 Rapid preparation of plasmid DNA

The following protocol adapted from Holmes and Quigley (1981) was used to screen *E. coli* transformants for recombinants.

Putative transformant colonies were picked and streaked onto fresh selective agar plates and grown overnight at 37°C. A toothpick streak of colonies was

transferred to a microfuge tube containing 100µl of STET buffer [8% (w/v) sucrose; 5% (v/v) Triton X100; 50mM EDTA; 50mM Tris-HCl (pH 8.0)]. 10µl of lysozyme (10mg/ml) was added and the cells were incubated at room temperature for 5 minutes. The cells were lysed by boiling for 1 minute. The cell debris was pelleted by centrifugation at 12000g (15 minutes). The supernatant, containing the plasmid DNA was transferred to a fresh tube and mixed with 1 volume of isopropanol.

The microfuge tube was incubated at -20°C for 1 hour to precipitate DNA and then centrifuged at 12000g (20 minutes). The supernatant was discarded and the DNA pellet washed with 200µl of 80% ethanol. The plasmid DNA was dried under vacuum for 10 minutes and then resuspended in 15µl TE buffer. The plasmid samples were then analysed directly by agarose gel electrophoresis either uncut or pre-digested with restriction endonucleases.

2.2.3 Preparation of plasmid DNA by alkaline lysis

The following alkaline lysis method of DNA preparation is that based on Birnboim and Doly (1979).

Nutrient medium (50ml) containing the required antibiotic was inoculated with a single colony from a freshly grown bacterial plate. This was incubated in an orbital shaker (37°C) overnight until the culture reached stationary phase. The culture was distributed between two plastic Universal tubes (Sterilin) and centrifuged at 2500g (10 minutes) in a bench centrifuge to pellet the cells. The supernatant was discarded and each cell pellet was gently resuspended in 0.5ml of ice-cold TEG buffer [25mM Tris-HCl; 10mM EDTA; 50mM glucose; pH 8.0]. The resuspended cells were transferred to fresh microfuge tubes and pelleted at 12000g (30 seconds). The supernatants were discarded and the pellets resuspended in 0.5ml TEG + 4mg/ml lysozyme. Each suspension was then divided to give four tubes each containing approximately 0.3ml, and left on ice for 10 minutes.

To each tube 0.5ml of freshly prepared NaOH/SDS solution [0.2M NaOH; 1% SDS] was added, and the tubes were mixed by inversion and held on ice for 5 minutes. 375µl of ice-cold 3M sodium acetate (pH 5.0) was added and the tubes were again

mixed by inversion. The tubes were held on ice for a further 10 minutes and then centrifuged at 12000g (10 minutes).

The supernatants were recovered in 0.5ml volumes and to each, two volumes of ethanol at -20°C were mixed in. The tubes were held at -20°C for 30 minutes, before centrifugation at 12000g (20 minutes). The supernatant was discarded and the crude DNA pellets were resuspended in a total volume of 300µl of TE (50ml starting culture gave 4 x 300µl).

To each tube, 150µl of ice-cold 7.5M ammonium acetate was added and the tubes were held on ice for 20 minutes and subsequently centrifuged at 12000g (10 minutes). The supernatants were transferred to fresh tubes and two volumes of ethanol at -20°C were mixed in. The mixtures were held at -20°C for 30 minutes before centrifugation at 12000g (20 minutes). The supernatant was discarded and the plasmid DNA dried under vacuum for approximately 10 minutes.

2.2.4 Purification of plasmid DNA by means of NACS cartridges

This method was adapted from that described in the Gibco-BRL instruction manual. The NACS cartridges consist of ion-exchange resin that binds and releases DNA at low and high salt concentrations, respectively.

The plasmid DNA pellets prepared by the alkaline lysis procedure were dissolved in 0.5ml of 50mM NaCl in TE containing 200 units/ml of RNase T1. The tubes were incubated at 55°C for 15 minutes, after which an equal volume of 0.35M NaCl in TE was added to each sample.

A 10ml disposable plastic syringe was attached to the NACS cartridge and clamped in a retort stand. A 2ml volume of 2M NaCl in TE was pushed through the cartridge to hydrate the resin. The cartridge was then equilibrated by pushing 10ml of 0.2M NaCl in TE through it.

A hypodermic needle was attached to the other end of the NACS cartridge and the DNA solution was applied and allowed to flow through by gravity. The eluate was collected and passed through the column again, after which the resin was washed with 5ml of 0.2M NaCl in TE. The DNA was eluted with two 400µl portions of 0.7M NaCl

in TE, and collected in separate microfuge tubes.

The DNA was precipitated by adding 2 volumes of ethanol and holding at -20°C (30 minutes) before centrifuging at 12000g (20 minutes). The supernatant was discarded and the purified DNA pellets were dried under vacuum for approximately 10 minutes before being resuspended in 50µl TE. An aliquot was taken and run on an agarose gel to assess the yield.

2.2.5 Preparation of genomic DNA

A volume of bacterial culture containing 8 x 10^{10} cells was centrifuged at 5000g (10 minutes, 4°C). The supernatant was discarded and the cell pellet resuspended in 11ml of buffer B1 [50mM EDTA; 50mM Tris-HCl, 0.5% Tween®-20; 0.5% Triton® X-100, pH 8; RNase A 200µg/ml]. The cells were incubated at 37°C for 30 minutes with 300µl of lysozyme (100mg/ml) and 500µl of Proteinase K (20mg/ml). 4ml of buffer B2 [3M GuHCl; 20% Tween®-20, pH 5.5] was added and incubation continued for a further 30 minutes at 50°C.

An equal volume of phenol/chloroform/isoamyl alcohol was added and gently mixed in. The solution was centrifuged at 5000g (10 minutes), the upper aqueous layer was collected and an equal volume of phenol/chloroform/isoamyl alcohol was added to it and the above process was repeated until the upper layer became a clear solution. The upper layers were collected and 2 volumes of ethanol were added and mixed in. The visible genomic DNA was spooled around a glass rod, removed and placed in a vial. Further traces of ethanol were removed by air-drying and the DNA was resuspended in a minimum volume of TE.

To determine purity and yield an aliquot was taken and its absorbance at 260nm and 280nm was measured (CECIL CE303 spectrophotometer).

2.2.6 Agarose gel electrophoresis

Horizontal agarose gel electrophoresis was used to examine all DNA prepared in this work.

Agarose (Appligene) at a concentration of 0.4% to 1% (w/v) was boiled in

electrophoresis buffer [40mM Tris-acetate; 2mM EDTA; pH 8.0] until it had fully dissolved and the solution was left to cool before addition of ethidium bromide to a final concentration of 0.5µg/ml. All samples were mixed with 0.2 volumes of 6 x sample buffer [15% Ficoll 400; 0.25% bromophenol blue; 0.25% xylene cyanol FF] before loading. Electrophoresis was performed at 100 volts until there was sufficient resolution of the DNA bands. DNA was visualized by examination under short wave (260nm) ultraviolet light on a UV transilluminator (A-VIOLET Products Ltd.). A 1Kb DNA ladder purchased from Gibco-BRL was used as molecular weight markers.

2.2.7 Recovery of DNA from agarose gels by use of the Geneclean II® kit

The solutions and instructions provided with the Geneclean II® kit (BIO 101 Inc.) were used to purify DNA from agarose gels. The procedure was based on that of Vogelstein and Gillespie (1979).

2.2.8 Restriction endonuclease digestion of DNA (plasmid and genomic)

Restriction enzymes were purchased from Gibco-BRL, Boehringer Mannheim, Pharmacia and Promega and were used in the buffers recommended by the manufacturer. A typical digestion mix consisted of:

DNA solution	Xµl
10 x reaction buffer	0.1 volume
Restriction enzyme	Y units
H_2O	Zµl to give the required volume.

The number of units of enzyme used was dependent on the mass of DNA to be digested. Typically 10 units of enzyme were used to digest > 1µg. Plasmid digestion mixtures were incubated for 2-3 hours at the optimum temperature for the restriction enzyme.

Genomic DNA was equilibrated on ice in reaction buffer and water with

occasional stirring for 2 to 3 hours prior to digestion. An aliquot of enzyme was added and the digestion allowed to proceed for one hour at the optimum temperature before addition of a second aliquot and continued digestion overnight (Maniatis *et al.*, 1989).

2.2.9 Treatment of DNA with alkaline phosphatase

Calf intestinal phosphatase (CIP) was used to remove 5' phosphate groups from linearized vector DNA prior to cloning. This reduces the level of vector recircularisation upon ligation and ensures that a greater proportion of transformant clones are recombinants.

Following restriction endonuclease digestion the enzyme was inactivated by incubation at 75°C (10 minutes). The pH was adjusted to 8.0 by the addition of 10x CIP buffer to give double the volume of 1x. The mixture was subsequently made up to the required volume with ddH_2O. 1 unit of alkaline phosphatase per µg of DNA was added to dephosphorylate linearized vector DNA. The mixture was incubated at 37°C for 30 minutes. If the vector had 3' cohesive or blunt termini, incubation was carried out at 55°C (30 minutes). The phosphatase was inactivated by the addition of 1µl of 0.5M EDTA and incubation at 75°C (10 minutes).

2.2.10 Ligation of DNA by T4 DNA ligase

T4 DNA ligase was used to join the ends of vector and fragment DNA. Ligation mixes were prepared as follows:

vector DNA (10-20ng)	Xµl
fragment DNA (Yng)	Yµl
10 x T4 ligase buffer	1µl
DTT (100mM)	1µl
T4 ligase (0.5-1 Unit)	0.5-1µl
sterile ddH_2O	Zµl to give a total volume of 10µl.

The molar ratio of fragment to vector was between 3:1 and 5:1. 10 x T4 ligase buffer was composed of 700mM Tris-HCl; 70mM $MgCl_2$; 10mM ATP; pH 7.5. 0.5 units of ligase were used for cohesive termini and 1 unit for blunt-ended ligations.

Ligation mixtures were incubated overnight at 14°C.

2.2.11 PCR amplification

The polymerase chain reaction (PCR) was used to amplify specific DNA sequences bounded by two custom-synthesized oligonucleotide primers. Synthesis of primers by means of an APPLIED BIOSYSTEMS 381A was carried out on appropriate columns.

The reaction mixtures for PCR were prepared as follows:

template DNA (10ng)	Yµl
10 x *Taq* DNA Pol. buffer	10µl
dNTPs (1mM)	5µl
Primer 1 (2µM)	10µl
Primer 2 (2µM)	10µl
Taq polymerase	2.5U
ddH_2O	Xµl to give 100µl volume.

The composition of the 10 x *Taq* DNA polymerase buffer was 500mM KCl; 100mM Tris-HCl; 5, 15 or 25mM $MgCl_2$; 0.1% (w/v) gelatin; pH 8.3.

The final mixture was overlaid with 75µl of sterile light mineral oil to prevent evaporation during the repeated thermal cycle. The PCR reaction was carried out in a HYBAID OMNIGENE thermal cycler. Cycles varied depending on the sequence to be amplified, and the details are given in subsequent chapters. As a general rule, target DNA was denatured for 60 seconds at 94°C. This was followed by annealing of the primers to complementary sequences on the template for 30 seconds at an optimum temperature, before an extension period of 90 seconds at 72°C. This cycle was

repeated at least 25 times before a final extension period of 10 minutes at 72°C. Upon completion, the aqueous reaction mixture was taken from beneath the oil and transferred to a fresh tube.

2.2.12 Treatment of PCR fragments for blunt-ended cloning

Proteinase K digestion was performed on PCR products to remove any bound *Taq* polymerase, thus improving the efficiency of ligation and cloning of these products (Crowe *et al.*, 1991).

A volume of the PCR mixture containing approximately 2µg of DNA was treated with proteinase K as follows:

PCR mixture (2µg DNA)	Xµl
Proteinase K (0.5mg/ml)	10µl
0.5% SDS in TE (w/v)	Yµl to give a total volume of 100µl

The mixture was incubated at 37°C (30 minutes) and subsequently at 68°C for 10 minutes. The proteinase K treated DNA was purified (Section 2.2.7) and resuspended in 18µl ddH_2O. The tube was incubated at 72°C (2 minutes) to separate the ends of DNA duplex and allow access of the kinase to the 5' ends. The mixture was snap-cooled by placing on ice and the following additions were made:

PNK/Klenow buffer + 10mM ATP	3µl
dNTPs (200µM)	3µl
DTT (50mM)	3µl
T4 polynucleotide kinase (10U)	1µl

The PNK/Klenow buffer consisted of 0.5M Tris-HCl (pH7.6); 0.1M $MgCl_2$ and 6.05mg/ml ATP.

The tube was incubated at 37°C (30 minutes) to promote phosphorylation, after which 2µl (10U) of Klenow was added to remove 3' overhangs. Incubation was carried out at room temperature for a further 30 minutes. The reaction was stopped by addition of 70µl of TE. The PCR fragments were then purified (Section 2.2.7).

2.2.13 Analysis of genomic DNA by Southern hybridization

This method described by Southern (1975) facilitates the detection of sequences within genomic DNA via homology with labelled DNA probes.

(i) Transfer of DNA to membranes via capillary action (Southern blotting)

DNA was alkali-transferred overnight from agarose gels to Hybond-N+ membranes as described in the protocol provided by the supplier (Amersham).

(ii) Radiolabelling of DNA probes

Typically 20ng of probe was labelled to a specific activity of approximately 2 x 10^9 dpm/µg with 20µCi of (α^{32}P)dCTP. A random primed labelling kit (Boehringer Mannheim) was used for this procedure. The specific activity of probe DNA was calculated by use of trichloroacetic acid (TCA) precipitation and scintillation counting (Maniatis *et al.*, 1989).

(iii) Hybridization of DNA to radiolabelled probes

The membranes were pre-hybridized at 65°C for 4 hours, in 10 ml of pre-hybridization solution [5 x SSPE, (20 x SSPE, 3.6M NaCl; 0.2M sodium phosphate; 0.02M EDTA, pH7.7); 5 x Denhardts solution; 0.5% (w/v) SDS; 1% Marvel]. The labelled probe was denatured by heating to 100°C for 5 minutes before adding to the pre-hybridization solution. The membrane was incubated with the probe for at least 12 hours at 50°C.

NB. A probe concentration of 20ng/ml of hybridization solution was not exceeded.

Following hybridization, the membrane was washed in 2 x SSPE, 0.1% (w/v)

SDS at room temperature for 10 minutes (twice). This is a low stringency wash. If required the membrane was subsequently washed in 1 x SSPE, 0.1% (w/v) SDS and sometimes 0.1 x SSPE, 0.1% (w/v) SDS at 65°C for 10 minutes. In between the stringency washes the membrane was exposed to X-ray film (Fuji RX-100) at -70°C for the appropriate length of time. Care was taken that the membrane did not dry out between stringency washes to facilitate the successful removal of probes.

Radiolabelled probes were removed by immersing the membranes in a boiling solution of 0.1% (w/v) SDS that was then allowed to cool to room temperature. The membranes were autoradiographed to check that the probe was completely removed.

2.2.14 Construction of a genomic library

(i) Optimization of conditions for partial digestion with *Sau*3A1

In order to optimize the enzyme concentration required to generate DNA fragments within the size range of 15-23Kb, small-scale reactions were performed. The following method was adapted from the protocol supplied with the Promega λ–GEM™-11 *Bam*H1 arms cloning system.

An assay buffer was made up as follows:

genomic DNA (1µg/µl)	10µl
*Sau*3A1 buffer (10 x)	20µl
BSA (1mg/ml)	20µl
ddH_2O	100µl

The *Sau*3A1 buffer consisted of 100mM Tris-HCl, pH 7.5; 1M NaCl; 70mM $MgCl_2$. 15µl of the assay buffer was added to each of 10 microfuge tubes.

The enzyme dilution buffer was made up as follows:

*Sau*3A1 buffer (10 x)	0.15ml
BSA (1mg/ml)	0.15ml

ddH_2O	1.20ml

The enzyme *Sau*3A1 was serially diluted with the above buffer to give the following enzyme concentrations for partial digestion; 1, 0.1, 0.05, 0.025, 0.015, 0.0125, 0.01, 0.0085, 0.005, 0.0035 units/μg DNA. 5μl of each dilution was added to the 15μl of assay buffer and digestion was allowed to proceed for 30 minutes at 37°C. The digestion was then analysed by agarose gel electrophoresis.

(ii) Partial digestion and size fractionation of genomic DNA

Large-scale preparation of partially digested genomic DNA was carried out under the conditions optimized above. The DNA concentration, time and temperature of incubation were identical to those used in the small-scale reactions. The size distribution of the digestion products was determined by analysing a small aliquot of the DNA (0.5μg) by gel electrophoresis.

If the digestion was satisfactory the remainder of the digest was run on an agarose gel. DNA within the size range 9-23Kb was removed from the gel as described in Chapter 6. DNA was resuspended in TE to a final concentration of 0.5mg/ml.

(iii) Ligation of genomic DNA to *Bam*H1 half-site arms

Optimal ligation conditions were determined by ligating vector arms to genomic insert in various ratios. The molar ratios of λ arms to vector insert used were varied between approximately 1:3 and 1:0.5.

Ligation mixtures were prepared as follows:

λ vector BamH1 half-site arms	0.5μg
genomic insert	0.15-0.9μg
T4 DNA ligase buffer	1μl
T4 DNA ligase (10-15U/μl)	1μl
ddH_2O	Xμl to give a final

volume of 10µl

T4 DNA ligase buffer consisted of 300mM Tris-HCl, pH 7.8; 100mM $MgCl_2$; 100mM DTT; 10mM ATP. Ligations were incubated for 12-16 hours at 14°C

(iv) Packaging of ligated DNA, and titration

Ligated DNA was packaged by means of the Packagene® *in vitro* packaging system (Promega). A Packagene extract (50µl) was thawed on ice. The entire ligation reaction was added, gently mixed in and incubated at 22°C (3 hours). For storage at 4°C for up to 3 weeks, 445µl of phage buffer [200mM Tris-HCl, pH 7.4; 100mM NaCl; 10mM $MgSO_4$] was added with 25µl of chloroform.

Bacterial plating cells were grown till an A_{600} of 0.6 was reached. The cells were pelleted by centrifugation at 5000g (10 minutes) and resuspended in $^1/_{10}$th the volume of 10mM $MgSO_4$ and stored at 4°C prior to use. 100µl of appropriately diluted recombinant phage was added to 100µl of prepared bacterial cells. The phage was allowed to adsorb for 30 minutes at 37°C after which 3ml of molten (45°C) tryptone (TB) top agar was added. The mix was vortexed gently and immediately poured onto LB plates. The top agar was allowed to solidify and the plates were incubated overnight at 37°C.

(v) Plaque lifts and hybridization

Plates were incubated at 4°C (1 hour) to harden the agar. The membrane was gently placed on one edge of the plate and progressively laid down until all the surface was covered. Duplicate lifts were made for each plate, the first membrane being left for 30 seconds and the second for 1 minute. The orientation of the filter was noted by stabbing a 20-gauge needle through both the membrane and agar at asymetric points. The membrane was removed and allowed to air-dry for 10-20 minutes, plaque side up, before baking at 80°C (2 hours). The membrane was placed in 0.2M NaOH, 1.5M NaCl (2 minutes), followed by immersion in 0.4M Tris-HCl, pH 7.6, 2 x SSPE (2

minutes) and finally 2 minutes in 2 x SSPE. The membrane was allowed to dry in air and then used for hybridization experiments as described previously (Section 2.2.13).

(vi) Preparation of phage lysate

Bacterial plating cells were prepared as described previously. In a 15ml Falcon polypropylene tube, 100µl of cells were mixed with 1 x 10^5 plaque-forming units (pfu) and incubated at 37°C (30 minutes). Alternatively, if a putative positive clone was identified this was eluted from an agar plate. The phage plaque was picked with a pasteur pipette and expelled into a 15ml Falcon polypropylene tube containing 1ml of SM buffer [0.01% gelatin; 50mM Tris-HCl, pH 7.5; 100mM NaCl; 8mM $MgSO_4$] and incubated for one hour at room temperature.

A volume of 60-100µl of the eluted phage or 1 x 10^5 pfu was mixed with 100µl of plating bacteria at 37°C (30 minutes). A volume of 3ml of molten (45°C) TB top agarose [1% Bacto-tryptone (Difco); 0.5% NaCl; 0.6% agarose] was mixed in and poured onto an LB plate. The top agarose was allowed to solidify and the plates were incubated at 37°C until the plaques became confluent. The plate was overlaid with 3ml SM buffer. The top agarose was transferred to a centrifuge tube by means of a spatula and incubated at room temperature for 30 minutes. The tube was centrifuged at 10000g (10 minutes, 4°C). The supernatant was transferred to a fresh tube and chloroform was added to a final concentration of 0.3% (v/v). The phage lysate was stored at 4°C.

2.2.15 RNA Preparation, dot blotting and hybridization

Trizol reagent (Gibco-BRL) was used to extract RNA from *E. coli* cells according to the instructions provided with the solution.

1µg of RNA was dotted onto a Hybond-N membrane by means of a vacuum manifold (Schleicher and Schuell). RNA was fixed to the membrane by baking in a vacuum oven at 80°C for 2 hours. Hybridization was carried out as described in Section 2.2.13 with the exception that salmon sperm DNA (0.5ml of 1mg/ml sonicated, denatured solution) replaced Marvel in the same total volume.

Marvel was not included in work with RNA since an unacceptable level of

RNAse activity was anticipated.

2.2.16 Maxicell *in vivo* expression system

This system uses a uv-sensitive host strain to allow selective labelling of plasmid-encoded proteins. Cells are irradiated with uv light, causing damage to DNA, which is subsequently degraded. Due to their smaller target size (compared to the genome) and higher copy number, many plasmids will survive. Subsequent incubation with ^{35}S-methionine allows exclusive labelling of proteins encoded by the plasmids.

The procedure followed was as described by Stoker *et al.*, (1984).

2.2.17 Induction of cultures by use of IPTG

Following growth of cultures to the required optical density (A_{600} monitored with a CECIL CE303 spectrophotometer), isopropyl β-D-thiogalactoside (IPTG) (40mM stock) was added to a final concentration of 0.4mM and incubation continued. 10 Minutes post IPTG induction, rifampicin was added, to a final concentration of 10μg/ml. At specified time points, cell aliquots were removed and incubated with ^{35}S-methionine (Amersham, 1000Ci/mmole) at a final concentration of 30mCi/ml (10 min, 37°C). Cells were pelleted at 12000g (5 min). The supernatant was discarded and the pellet resuspended in sample buffer to give a solution such that 10μl corresponded to 0.1 A_{600} unit. All samples were analysed by SDS-PAGE and fluorography, (Section 2.2.21).

2.2.18 Preparation of periplasmic fraction from ^{35}S-labelled cells, by osmotic shock

100μl of cells labelled with ^{35}S-methionine were pelleted in the microfuge and resuspended in 125μl of, 20% sucrose; 30mM Tris; 1mM EDTA, pH 8.0. The suspension was incubated, with shaking, at room temperature (10 minutes). The cells were pelleted by centrifugation at 12000g (3 minutes).

The supernatant (SN1) was transferred to a fresh tube and held on ice. The pellet was resuspended in 312μl of ice-cold ddH_2O and was shaken on ice for 10

minutes. The suspension was centrifuged at 12000g (10 minutes). The supernatant (SN2) was removed to a fresh tube and kept on ice. Proteins from fractions SN1 and SN2 were recovered by TCA precipitation (half the volume of 30% TCA was mixed in and the tube held on ice for at least 30 minutes before centrifugation at 5000g (5 minutes, 4°C) to pellet precipitated proteins) and, together with the pellet were solubilized in sample buffer and analysed by SDS-PAGE and fluorography, (Section 2.2.21).

2.2.19 Cell fractionation of ^{35}S-labelled cells

The strain CSH26ΔF6 was grown in an orbital shaker (37°C) overnight in 10ml of nutrient broth. 1ml of this culture and 3.9ml of 10mM sodium phosphate buffer, pH 7.4 were added to 100µl of ^{35}S-methionine labelled cells. The suspension was sonicated for three 1 minute pulses with 1 minute cooling intervals between each pulse. The sonicated cells were centrifuged at 5000g (20 minutes, 4°C) to remove whole cells. The supernatant was transferred to fresh ultracentrifuge tubes and centrifuged in a TL-100 ultracentrifuge for 30 minutes at 38000 rpm (4°C) to pellet the membrane fraction.

The supernatant was transferred to a fresh tube. Cytoplasmic and periplasmic proteins in the supernatant were recovered by precipitation with TCA as before. The pellet was resuspended in 100µl of 0.5% (w/v) sarkosyl and incubated at room temperature for 20 minutes.

The sample was then centrifuged in a TL-100 ultracentrifuge at 38000 rpm (30 minutes, 4°C). The pellet, containing outer membrane proteins, was resuspended in 60µl ddH_2O. The inner membrane proteins in the supernatant were recovered by TCA precipitation as before. All samples were analysed by SDS-PAGE and fluorography, (Section 2.2.21).

2.2.20 *In vitro* transcription/translation

In vitro transcription-translation reactions were carried out by using S30 extracts prepared from strain MC4100 according to the method described by Nevin and Pratt (1991). MC4100, containing the plasmid pAR1219 (which encodes T7 RNA

polymerase), was used to prepare S30 extracts for T7-controlled transcription-translation reactions.

The following components were added to a microfuge tube;

Low molecular weight mixture	7.5µl
^{35}S-methionine	2.0µl
S30 extract	5.0µl
MgAc (0.1M)	3.5µl
DNA (2-3µg)	X µl
TE	Y µl to give a final volume of 30µl

Preparation of the low molecular weight mixture and the S30 extract was as described by Pratt (1984). The mix was incubated for 30 minutes at 37°C after which 30µl of ddH_2O and 60µl of 2 x sample buffer were added. The samples were boiled for 5 minutes before analysis by SDS-PAGE and fluorography, (Section 2.2.21).

2.2.21 SDS polyacrylamide gel electrophoresis (SDS-PAGE)

The procedure was based on that described by Laemmli (1970) and was carried out by using the Bio-rad PROTEAN® II XI slab cell kit or Bio-rad MINI-PROTEAN® II slab cell kit according to the manufacturer's instructions. Gels 0.75mm thick were prepared with a 7% polyacrylamide (Severn Biotech Ltd.) stacking gel cast over a 12.5% polyacrylamide resolving gel. Protein samples were prepared for SDS-PAGE by the addition of 1 volume of 2 x SDS sample buffer [4% (w/v) SDS; 20% (v/v) glycerol; 10% (v/v) β-mercaptoethanol; 125mM Tris-HCl, pH 6.8; bromophenol blue] and boiling for 5 minutes prior to loading. Electrophoresis was carried out at 200 volts for mini-gels, while large gels were run overnight at 10 volts per gel. After electrophoresis, gels were fixed for 10 minutes by immersion in fixing solution [100ml glacial acetic acid; 250ml isopropanol; 650ml ddH_2O].

2.2.22 Staining of SDS polyacrylamide gels

Gels were stained by soaking in fixing solution containing 0.05% (w/v) Coomassie brilliant blue (Sigma) for at least 4 hours. Background staining was reduced by sequential immersions in fresh destaining solution [10% acetic acid and 10% (v/v) isopropanol].

2.2.23 Fluorography of ^{35}S-methionine-labelled proteins separated by SDS-PAGE

The gel was drained of fixer and covered in dimethyl sulphoxide (DMSO) and incubated at room temperature (30 minutes). The DMSO was discarded and the process repeated with fresh DMSO. The gel was then covered in 22% (w/v) 2,5-diphenyloxazole (PPO) in DMSO and shaken at room temperature for 90 minutes.

The PPO in DMSO was poured off to be reused and the gel was washed in tap water for 1 hour at room temperature. The gel was dried onto Whatman No. 17 chromatography paper by means of a Biorad vacuum slab gel drier. The dried gel was placed next to a piece of X-ray film (Fuji RX-100) in a light-tight cassette and placed at -70°C for exposure. The film was developed after an appropriate exposure (AGFA Curix 60 developer).

2.2.24 Production of antibodies

Specific antisera can be raised by immunizing animals with short synthetic peptides (Lerner 1984). Sequences of 10 amino acids were chosen with no more than 4 adjacent hydrophobic residues, no more than 6 hydrophillic residues in total and a high number of charged amino acids. These properties facilitated easy linkage to a carrier protein such as keyhole limpet haemocyanin (KLH) through a carboxy-terminal cysteine residue that is not included in the 10 amino acid sequence. In addition the polypeptide is more likely to be on the surface of the target protein and thus accessible to antibody.

(i) Polypeptide synthesis

The proteins CrtY and CrtI were analysed by the peptide structure program within the GCG (Genetics Computer Group, Wisconsin, USA) package before amino acid sequences were chosen to fulfil the above requirements. The amino acid sequence chosen was WRLRQRYPQGGC from CrtY and QRDKPGGRAYGGC from CrtI. The polypeptides were synthesised at the Sussex Centre for Neuroscience, University of Sussex, Brighton.

(ii) Conjugation of polypeptide to keyhole limpet haemocyanin (KLH)

5mg of KLH was dissolved in 1.5ml of 0.05M sodium phosphate buffer, pH 7.0. 1mg of m-maleimidobenzoic acid N-hydroxysuccinimide ester (MBS) was added and the suspension was left at room temperature for 40 minutes. A Sephadex-200 gel filtration column (20ml bed volume) was used to remove unconjugated MBS. 5mg of polypeptide was added to the activated KLH and left to react overnight at room temperature. The conjugated antigen was stored at -20°C until needed.

(iii) Immunization of hens

Two brown laying hens were purchased from Whitehouse Farm, Lancaster. The conjugated antigen (1ml) was emulsified with an equal volume of Freunds complete adjuvant (Sigma) by sonication (4-6 x 15 second intervals with intermittent cooling). The suspension was injected into the pectoral muscle. Another injection of the antigen (prepared as described above) was given approximately 20 days later. The eggs were collected daily and stored at 4°C until use.

(iv) Purification of antibodies from hens egg yolk

The egg yolk was separated from the white and washed with ddH_2O. The volume was increased to 15ml with extraction buffer (EB) [10mM KH_2PO_4; 0.1M NaCl; pH 7.2] and mixed with 15ml 7% (w/v) PEG 6000 in EB. The precipitate was pelleted by centrifugation at 10000g (10 minutes, 4°C). The supernatant was filtered through one thickness of tissue paper and mixed with 1 volume of 24% (w/v) PEG

6000 in EB. This mixture was centrifuged as described previously, the supernatant discarded and the pellet resuspended in 10ml EB. When fully resuspended, 10ml of 24% (w/v) PEG 6000 in EB was added and the pellet precipitated by centrifugation as before. The pellet was resuspended in 10ml EB and stored at -20°C.

2.2.25 Western blotting

The procedure followed was based on that of Towbin *et al.*, (1979) and was carried out by means of the Bio-rad TRANS-BLOT® or Bio-rad MINI TRANS-BLOT®, onto Polyscreen PVDF transfer membrane (Du Pont) according to the manufacturer's instructions. Rabbit anti-chick IgG alkaline phosphatase conjugate was used (1 in 3000 dilution) as the second antibody, with 5-bromo-4-chloro-3-indolyl phosphate (BCIP) and nitro blue tetrazolium (NBT) colour development solutions. BCIP and NBT solutions were made fresh prior to use. To 20ml carbonate buffer [0.1M $NaHCO_3$; 1.0mM $MgCl_2$; pH to 9.8 with 5M NaOH], 200µl of 30mg/ml NBT [in 70% (v/v) dimethylformamide] and 200µl of 15mg/ml BCIP [in 100% dimethylformamide] were added and poured onto the membrane. Incubation was carried out for up to 30 minutes for colour development. To stop colour development when the required sensitivity was reached, the membrane was rinsed in a number of changes of ddH_2O.

2.2.26 Characterization of Pigments

(i) Pigment isolation

Whole cells were extracted with 1 volume of hot (55°C) acetone/methanol (7:2). The extract was filtered through Whatman No. 1 filter paper into a buchner flask and the cell debris re-extracted until virtually colourless. The extract was taken up into ether and washed with water to remove any traces of acetone/methanol. The ether phase was evaporated to dryness. The carotenoids were then dissolved in dichloromethane (DCM) or ether and filtered through a pasteur pipette plugged with cotton wool before evaporation under nitrogen. Carotenoids and extracts were stored under nitrogen at

-20°C.

(ii) Reversed-phase HPLC of carotenoid pigments

A Waters 660 solvent delivery system was used with a Waters 990 on-line multichannel photodiode array detector. Data were stored and analysed on a dedicated NEC Powermate personal computer. The stationary phase employed was a Spherisorb ODS2 reversed phase HPLC column (5µm particle size; 25cm x 4.6mm diameter). The mobile phase consisted of a gradient of 0-100% ethyl acetate in acetonitrile - water (9:1 v/v) at 1ml/min with the solvents continuously sparged with helium.

Carotenoids were identified by comparison with known standards where possible, and tentative identification from the retention time and absorption spectrum of each peak.

(iii) Thin-layer chromatography of carotenoids

Carotenoids were separated on silica thin-layer chromatography (tlc) plates (20cm x 20cm). Plates (thickness 0.5mm) were prepared from a slurry of silica gel G (30g) and 0.3% KOH in water (60ml), by means of a 'Quickfit' 8CR thin-layer plate spreader (J. A. Jobling Ltd., Stone, Staffs.), and dried for 1 hour at 110°C. Samples were dissolved in a small volume of ether or DCM and applied to plates from finely drawn pasteur pipettes. Generally, ether:petrol in a ratio of 1:1 was sufficient to resolve bacterial carotenoids. The separated bands were scraped off, eluted with ether and evaporated under nitrogen.

(iv) Mass spectrometry

Carotenoids isolated from tlc plates were identified by their molecular ions and fragmentation patterns. Electron impact ionization mass spectra were recorded by means of a VG Quattro quadruple mass spectrometer operated in the positive ion mode. The source temperature was 240°C, emission current 200µA, electron energy 70eV. The mass range scanned was *m/z* 40-750 Daltons in scan time 4 seconds.

CHAPTER 3

INCREASING THE EXPRESSION OF CAROTENOID BIOSYNTHESIS ENZYMES IN *E. COLI*

3.1 General introduction

Biotechnology involves the use of micro-organisms for the production of commercial products on an industrial scale. Recombinant DNA technology is a powerful tool in this process. Micro-organisms are engineered to have either novel or enhanced genetic attributes.

This chapter describes the genetic manipulation of the *crt* genes of *E. herbicola* with a view to increase carotenoid production.

Various expression studies on the individual genes have been performed. A plasmid encoding the carotenoid biosynthesis genes of *E. herbicola* was originally isolated from a cosmid library. All the genes were present on one isolate, designated pPL376. This plasmid conferred an orange phenotype on *E. coli* cells. Carotenoid production was shown to be inhibited by glucose, and cAMP was thought to be involved in regulation (Perry *et al.*, 1986). The genes for the enzymes CrtX, CrtY and CrtZ have been cloned individually downstream of a T7 promoter and their activities studied *in vitro* (Hundle *et al.*, 1992, 1993). Manipulation (by deletion and subsequent addition of base pairs) of *crtE* at the 5' and 3' ends resulted in increased enzyme activity when the genes were expressed in *E. coli*, (Ausich, 1994). Ausich also successfully manipulated *crtY* such that it could be expressed in both *E. coli* and *S. cerevisiae*. Phytoene desaturase from *E. herbicola* has been introduced into a strain of *R. sphaeroides* lacking CrtI. In *R. sphaeroides*, CrtI is responsible for the production of neurosporene. Along with lycopene (previously not found in *R. sphaeroides*) two new carotenoids were produced, spirilloxanthin and 2-ketospirilloxanthin (Appendix I) (Ausich, 1994). The introduction of *E. herbicola* genes into *R. sphaeroides* has since been taken a step further. The *crtI*, *Y*, *B* and *Z* genes were transformed into various

mutant strains of *R. sphaeroides*, resulting in carotenoid compositions that differed from normal (Hunter *et al.*, 1994).

3.2 The T7 expression system

T7 RNA polymerase and its cognate promoter provide a widely used system to increase the expression of genes (Studier and Moffatt, 1986). The T7 *in vivo* expression system comprises cloning the gene of interest downstream of a T7 promoter and placing this plasmid into a host *E. coli* strain which carries the gene for T7 RNA polymerase under the control of the *lac*uv5 promoter (Dubendorf and Studier, 1991).

The T7 RNA polymerase is structurally and functionally distinct from *E. coli* RNA polymerase. It recognizes a highly conserved 23 base pair promoter region, no additional cofactors are required for transcription and it has a greater elongation rate and processivity than *E. coli* RNA polymerase. T7 RNA polymerase is resistant to rifampicin whereas *E. coli* RNA polymerase is not, thus allowing for selective gene expression from the T7 promoter in the presence of rifampicin (Tabor and Richardson, 1985).

Cells are grown and then induced with IPTG (T7 RNA polymerase is synthesized), the cells are treated with rifampicin and proteins synthesized may be labelled with ^{35}S-methionine. Since the gene of interest is the only one transcribed in the presence of rifampicin its product is the major radioactive protein present. Therefore by using rifampicin it can be ensured that the only genes that are newly transcribed will be those expressed from the T7 promoter. The latter approach allows the monitoring of expression of given proteins in the absence of available monospecific antisera.

The aims of the work described in this chapter were to clone the *crt* genes of *E. herbicola* so that they are under the control of the T7 promoter. The results of these cloning procedures were compared at the protein, RNA and carotenoid levels to determine whether an increase in carotenoid production was achieved.

3.3 Genetic manipulation of *crt* genes

(i) Placing the *crt* genes under the control of T7 RNA polymerase

Within the six-gene carotenoid biosynthesis cluster of *E. herbicola*, *crtZ* is transcribed in the opposite direction. By using simple cloning strategies the 10Kb *Hin*dIII-*Eco*RI fragment (Fig. 3.1) of the plasmid pAPU211 could be cloned in both directions downstream of a T7 promoter by using the two vectors pT7-5 and pT7-6 (Appendix II). These vectors were chosen as they both have the strong T7 promoter ϕ10.

The plasmids pAPU211, pT7-5 and pT7-6 were all digested with *Eco*RI and *Hin*dIII in a buffer compatible for both enzymes (2.2.8). The vectors pT7-5 and pT7-6 were then treated with alkaline phosphatase to remove 5' phosphates (2.2.9). A small portion of the digest was analysed by agarose gel electrophoresis (2.2.6) to confirm that digestion had gone to completion (data not shown). The remainders of the digests were then separated by agarose gel electrophoresis. Bands containing the digested pT7-6, pT7-5 and the 10Kb *crt* fragment were excised and the DNA was recovered (2.2.7). A small portion of the recovered DNA was analysed on an agarose gel and the estimated recovery was ~70%.

10ng of the 2.5Kb digested vector was mixed with 50ng and 150ng of the 10Kb *crt* fragment to give molar ratios of vector to fragment of 1:1 and 1:3 respectively. The vectors and fragment were ligated overnight at 14°C along with appropriate controls (2.2.10). Competent DH5α cells were transformed with the overnight ligations (2.2.1). Transformants were selected on nutrient agar (NA) ampicillin plates (25μg/ml) and, after overnight incubation at 37°C, a number of colonies were visible.

On closer inspection, some of these colonies were orange when compared with the control transformants, suggesting that the 10Kb fragment was correctly inserted into each vector. These colonies were used to inoculate 10ml of nutrient broth (NB) and plasmid DNA was isolated (2.2.2). Plasmids were analysed by agarose gel electrophoresis after digestion with *Eco*RI and *Hin*dIII [Fig. 3.2 (b)]. The resulting

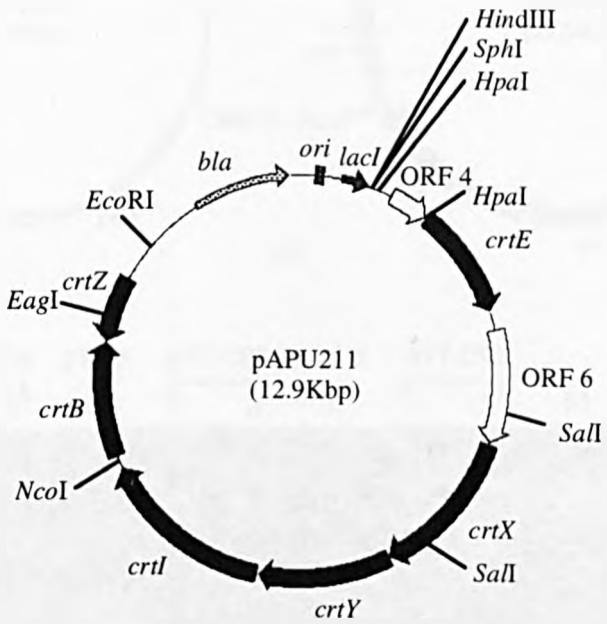
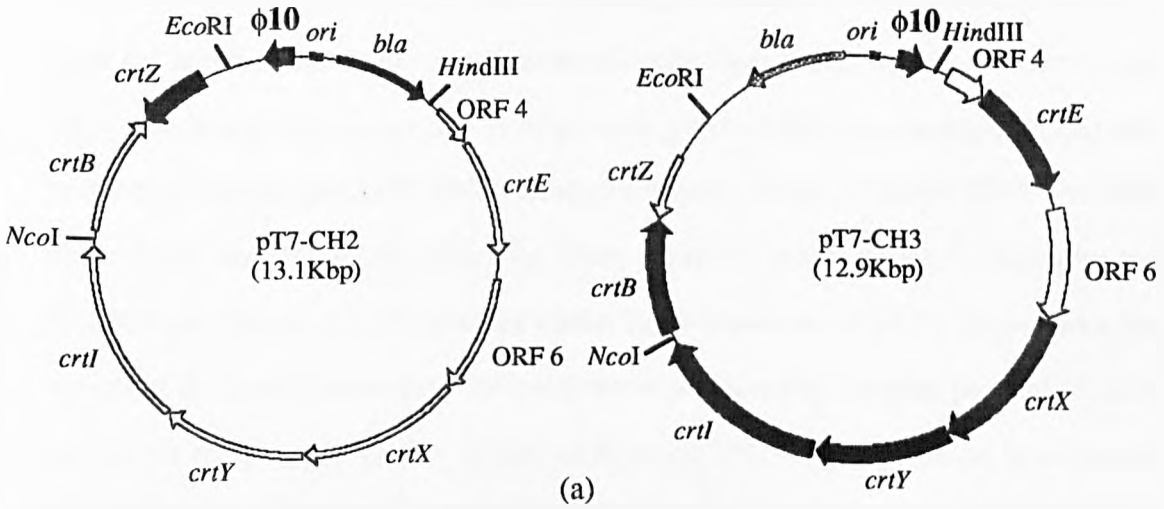
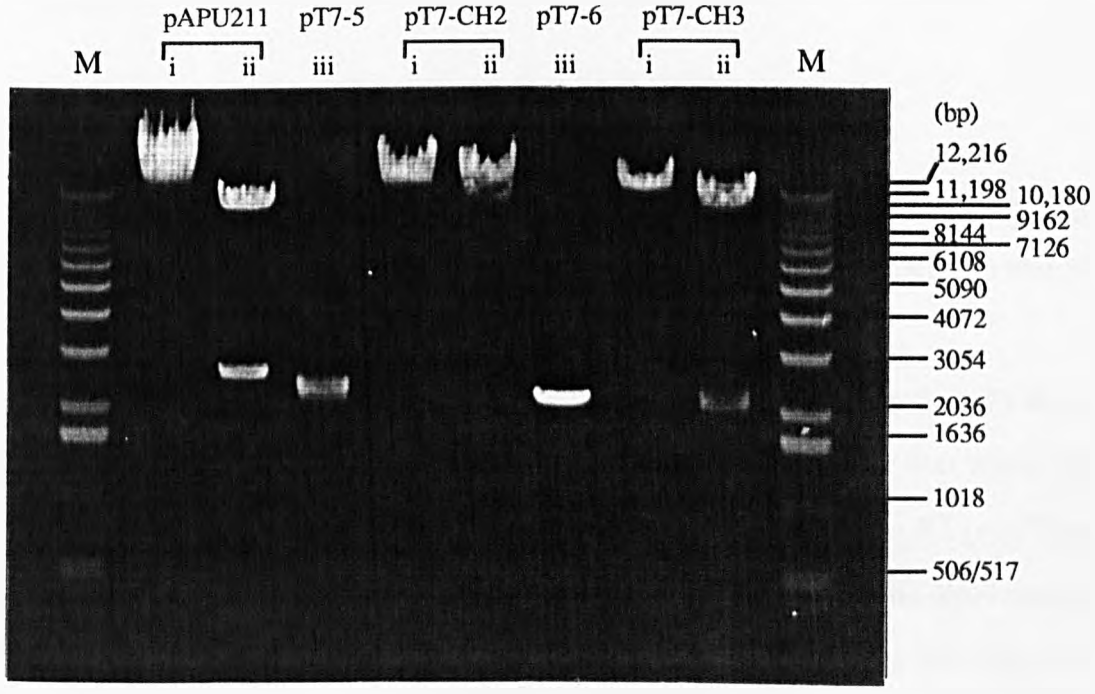


Figure 3.1 Restriction enzyme map of pAPU211.
The 10.185Kbp *Sph*I-*Eco*RI fragment from the plasmid pPL376 (Perry *et al.*, 1986) was cloned into the vector pUC19 (Appendix II) (Hundle *et al.*, 1994) at these sites to produce the recombinant plasmid pAPU211. pAPU211Z was generated by the deletion in pAPU211 of the *Eag*I-*Eco*RI fragment, blunting the overhangs and ligating them (Hundle *et al.*, 1994).



(a)



(b)

Figure 3.2 Restriction enzyme maps and agarose gel analysis of the constructs pT7-CH2 and pT7-CH3.
(a) Restriction enzyme maps of the plasmids pT7-CH2 and pT7-CH3. The*Hin*dIII-*Eco*RI fragment of the plasmid pAPU211 was cloned into the vectors pT7-5 and pT7-6 (Appendix II) to produce the recombinant plasmids pT7-CH2 and pT7-CH3 respectively.
(b) Agarose gel (0.8%) analysis of restriction enzyme digests of the above constructs. Digests performed are; (i)*Nco*I, (ii)*Hin*dIII-*Eco*RI, (iii)*Hin*dIII. M - 1 kilobase ladder (Gibco-BRL).

digest patterns, and the fact that only one orientation of the fragment was possible from the strategy, confirmed that the required constructs pT7-CH2 and pT7-CH3 [Fig. 3.2 (a)] had been obtained. In this cloning strategy, endogenous *E. coli* promoter sites were not removed and so the transformant colonies were orange in colour as carotenoid biosynthesis was still occurring. A large-scale plasmid DNA preparation (2.2.3) was performed and the yield of DNA was approximately 75µg. Plasmid DNA was then transformed into BL21(DE3)pRG1 for future expression studies (2.2.1). The gene for T7 RNA polymerase (DE3) is present within the chromosome of BL21 and is under the control of the *lac*uv5 promoter. The promoter is repressed by the gene product of *lacI*^{q} carried on the plasmid pRG1. Upon addition of IPTG this repression is removed allowing expression of the T7 RNA polymerase and subsequent transcription from the T7 promoter.

(ii) Amplification of *crtZ* by PCR

As *crtZ* is transcribed in the opposite direction to the rest of the *crt* genes (Fig. 3.1) it was decided to clone this gene individually and also to turn it round so that all six genes would be transcribed in the same direction.

The gene was amplified by means of the polymerase chain reaction (PCR) as this conveniently allowed the introduction of specific restriction enzyme sites within the primers to facilitate cloning. The primers chosen are shown in Figure 3.3 (a). They were designed so that *crtZ* could be inserted into pAPU211 at the *Hin*dIII-*Hpa*I sites in front of *crtE* and the same fragment could be inserted into pUC19 as a *Hin*dIII-blunt ended fragment to yield *crtZ* on its own.

The PCR amplification of *crt*Z was carried out as described in 2.2.11 with 100ng of the template pAPU211 and an annealing temperature of 50°C. An aliquot was removed and analysed by agarose gel electrophoresis (data not shown). The expected 530bp product was generated and is depicted in Figure 3.3 (b). The remainder was treated with proteinase K (to remove any *Taq* polymerase) and then treated with T4 polynucleotide kinase and Klenow DNA polymerase to add 5' phosphates and to fill in the ends of the fragment (2.2.12). The fragment was then digested with *Hin*dIII

***crtZ* 5' primer**

(CA)

5' - TTGATTAAGCTTGTCTGAGTCGCAGGAGC - 3'

*Hin*dIII

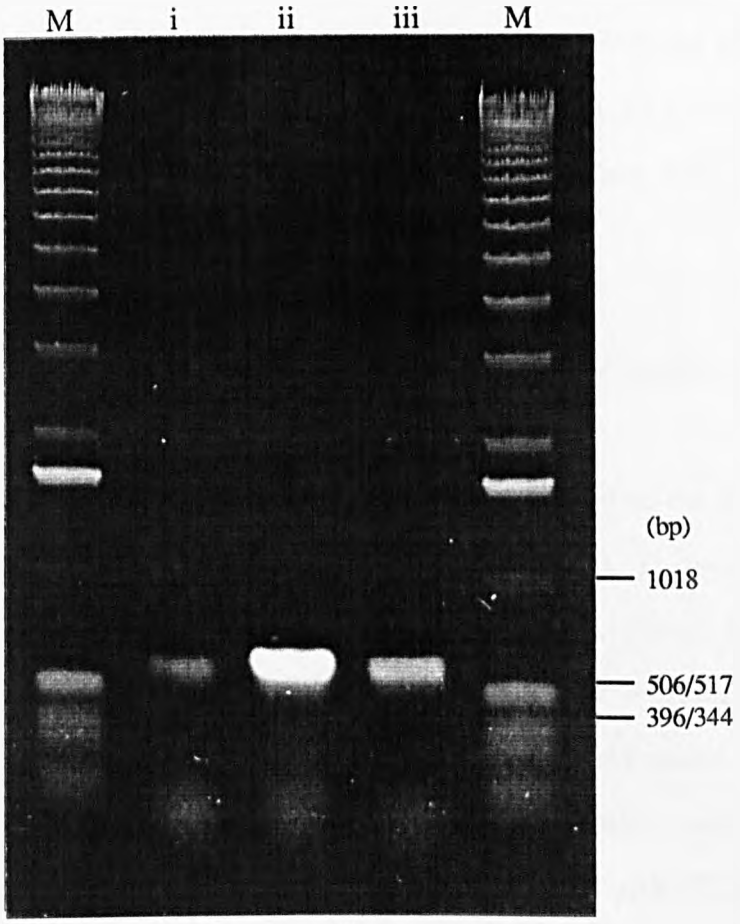
***crtZ* 3' primer**

(A) (GC)

5' - GTGAAAAAGTTAACATGCTGATGGCGGCA - 3'

*Hpa*I

(a)



(b)

Figure 3.3 PCR primers designed to amplify *crtZ* gene and analysis of PCR products. (a) PCR primers designed to amplify *crtZ* from the plasmid pAPU211. 5' primer with base alteration shown in bold to incorporate a *Hin*dIII site underlined and 3' primer with base alterations shown in bold to incorporate a *Hpa*I site underlined.
(b) Agarose gel (0.8%) analysis of PCR products obtained by using pAPU211 as the template. PCR was carried out as described in 2.2.11 with PCR buffer containing a Mg^{2+} concentration of (i) 2.5mM, (ii) 1.5mM and (iii) 0.5mM and an annealing temperature of 50°C. M - 1 kilobase ladder (Gibco-BRL).

(2.2.8). The vector pUC19 was digested with *Sma*I, gel purified (2.2.7) and then digested with *Hin*dIII. The PCR fragment was then ligated overnight with the digested vector plus the appropriate controls (2.2.10) and then transformed into competent DH5α cells (2.2.1).

Plasmid DNA was isolated from a number of transformants (2.2.2), and analysed by agarose gel electrophoresis (2.2.6). Restriction enzyme digests of *Hin*dIII and *Hin*dIII-*Hpa*I were performed to check if the inserted *crtZ* gene could be excised out of the vector pUC19. Figure 3.4 (b) shows that this was the case confirming that the required construct pCH4 [Fig. 3.4 (a)] had been obtained. The construct pT7-CH5 was obtained by digesting pCH4 with *Hin*dIII-*Eco*RI to excise *crtZ*, the fragment was gel purified and then ligated with *Hin*dIII-*Eco*RI digested vector, pT7-6. The required recombinant was confirmed by restriction enzyme digestion and agarose gel electrophoresis [Fig. 3.4 (b)].

(iii) Transcription of all *crt* genes in the same direction under the control of T7

The plasmid pAPU211Z has been generated by deleting the *Eag*I-*Eco*RI fragment which encompassed the gene *crtZ* (Hundle *et al.*, 1994). This vector plasmid was chosen to clone *crtZ* in front of the other five *crt* genes. If *crtZ* were cloned into a plasmid already containing *crtZ* this could lead to instability, rearrangements and deletions. By using the chosen plasmid the potential problem was avoided.

The plasmids pAPU211Z and pCH4 were each digested with *Hin*dIII and *Hpa*I in a buffer compatible for both enzymes (2.2.8). The plasmid pAPU211Z was then treated with alkaline phosphatase to remove 5' phosphates (2.2.9). Small portions of each digest were analysed by agarose gel electrophoresis (data not shown). Once digestion was complete the remainders of the digests were separated by agarose gel electrophoresis (2.2.6). Bands containing the large (approximately 10Kb) pAPU211Z fragment and the small (approximately 500bp) *crtZ* fragment were excised from the gel and purified (2.2.7). The *crtZ* fragment was ligated to the pAPU211Z fragment in molar ratios of 1:1 and 3:1, respectively, overnight at 14°C, along with appropriate

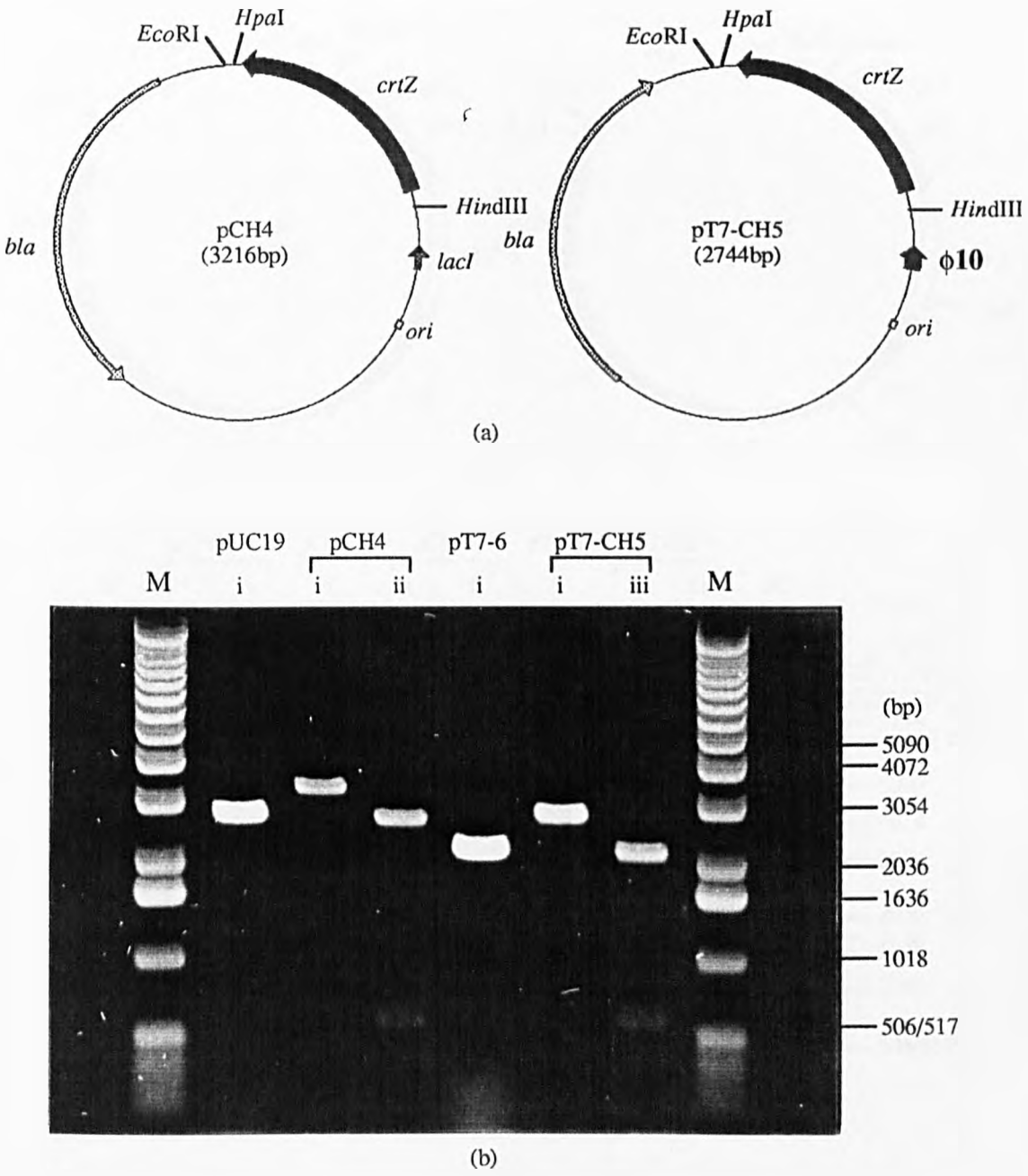
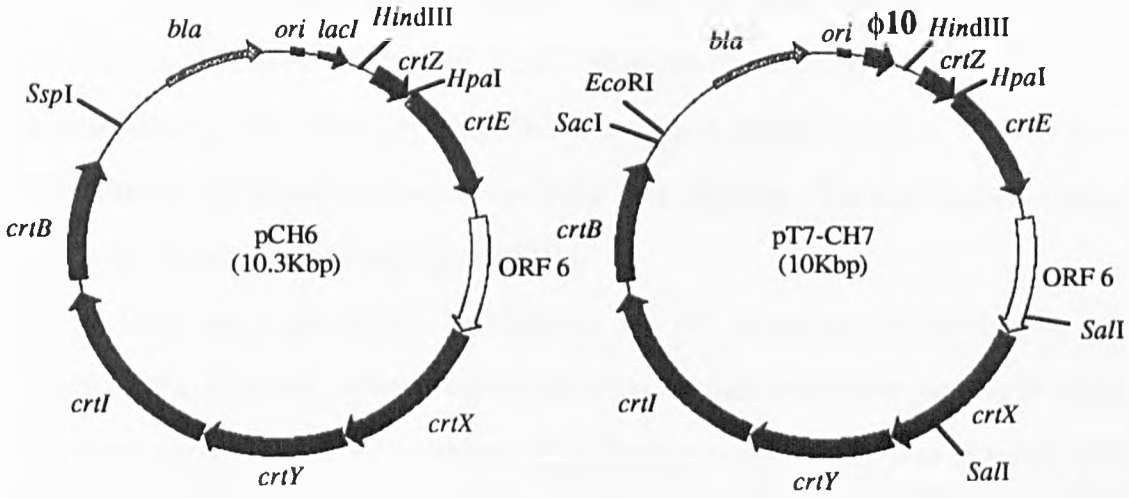
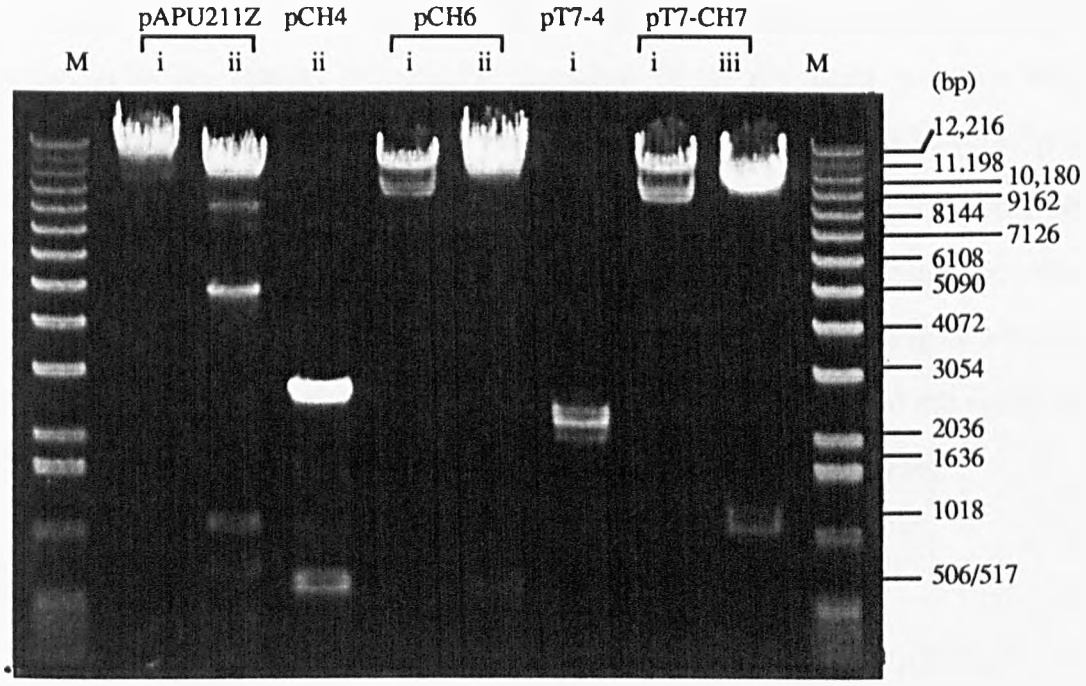


Figure 3.4 Restriction enzyme maps and agarose gel analysis of the constructs pCH4 and pT7-CH5.
(a) PCR amplified *crtZ* was cloned into the vector pUC19 (Appendix II) at the sites *Hin*dIII and *Sma*I as a *Hin*dIII blunt fragment to produce the recombinant plasmid pCH4. pCH4 was subsequently digested with *Hin*dIII and *Eco*RI to allow the cloning of *crtZ* at these sites into the vector pT7-6 (Appendix II) to produce the recombinant pT7-CH5.
(b) Agarose gel (0.8%) analysis of restriction enzyme digests of the above constructs. Digests performed are; (i) *Hin*dIII, (ii) *Hin*dIII-*Hpa*I, (iii) *Hin*dIII-*Eco*RI. M - 1 kilobase ladder (Gibco-BRL).



(a)



(b)

Figure 3.5 Restriction enzyme maps and agarose gel analysis of the constructs pCH6 and pT7-CH7.

(a) Restriction enzyme maps of the recombinant plasmids pCH6 and pT7-CH7. The *Hin*dIII-*Hpa*I fragment of pCH4 was cloned into these sites of the plasmid pAPU211Z to give the recombinant pCH6. The *Hin*dIII-*Ssp*I fragment of pCH6 was cloned into the *Hin*dIII-*Sma*I sites of the vector pT7-4 (Appendix II) to produce the recombinant pT7-CH7.

(b) Agarose gel (0.8%) analysis of restriction enzyme digests of the above constructs. Digests performed are; (i) *Hin*dIII, (ii) *Hin*dIII-*Hpa*I, (iii) *Sal*I. M - 1 kilobase ladder (Gibco-BRL).

controls (2.2.10). Competent DH5α cells were transformed with the ligations (2.2.1). Transformants were selected on NA ampicillin plates.

A single transformant was used to inoculate 10ml of NB and plasmid DNA was isolated as described in Section 2.2.2. Plasmids were analysed by agarose gel electrophoresis after digestion with *Hin*dIII and *Hpa*I [Figure 3.5 (b)]. The results of these digests confirmed that the construct had been obtained. The colonies were not as distinctly orange as cells carrying pAPU211.

Once the clone pCH6, in which all the *crt* genes are transcribed in one direction, was obtained, a further cloning procedure would enable the genes to be under the direct control of the T7 promoter. The cloning vector pT7-4 was digested with *Hin*dIII and *Sma*I and dephosphorylated. The *crt* cluster was removed from the plasmid pCH6 by digestion with *Hin*dIII and *Ssp*I and, the 8.5Kb fragment was purified from the agarose gel (2.2.7). Ligations of the *crt* insert to vector were performed in the ratios of 1:1 and 3:1 as previously. Competent DH5α cells were transformed with the ligation mixes and recombinants were selected on NA ampicillin plates as before. Plasmid DNA was isolated from a number of individual colonies (2.2.2). Agarose gels were used to analyse restriction enzyme digests [Figure 3.5 (b)]. From the sizes of fragments generated in the digests it was concluded the construct pT7-CH7 had been obtained. The transformed colonies were not pigmented.

3.4 *In vivo* induction of the carotenoid biosynthesis genes

In vivo induction of the *crt* genes in pT7-CH2, pT7-CH3 and pT7-CH7 was performed by using IPTG at a final concentration of 0.4mM, rifampicin (10μg/ml) was added 10 minutes post IPTG induction (2.2.17). No visible induction of the Crt enzymes was revealed on stained profiles or by using radioactive pulse labelling with ^{35}S-methionine, with the constructs pT7-CH2 and pT7-CH3 [Figure 3.6 (a) and (b)]. The radioactive profile of pT7-CH7 [Figure 3.7 (b)] showed the presence of proteins at 54KDa, 45KDa, 43KDa, 36KDa and 33KDa, that were absent from the vector pT7-4 (used as a negative control) These could correspond to the Crt enzymes CrtI, CrtX, CrtY, CrtB and CrtE respectively. No protein corresponding to CrtZ (20KDa) was

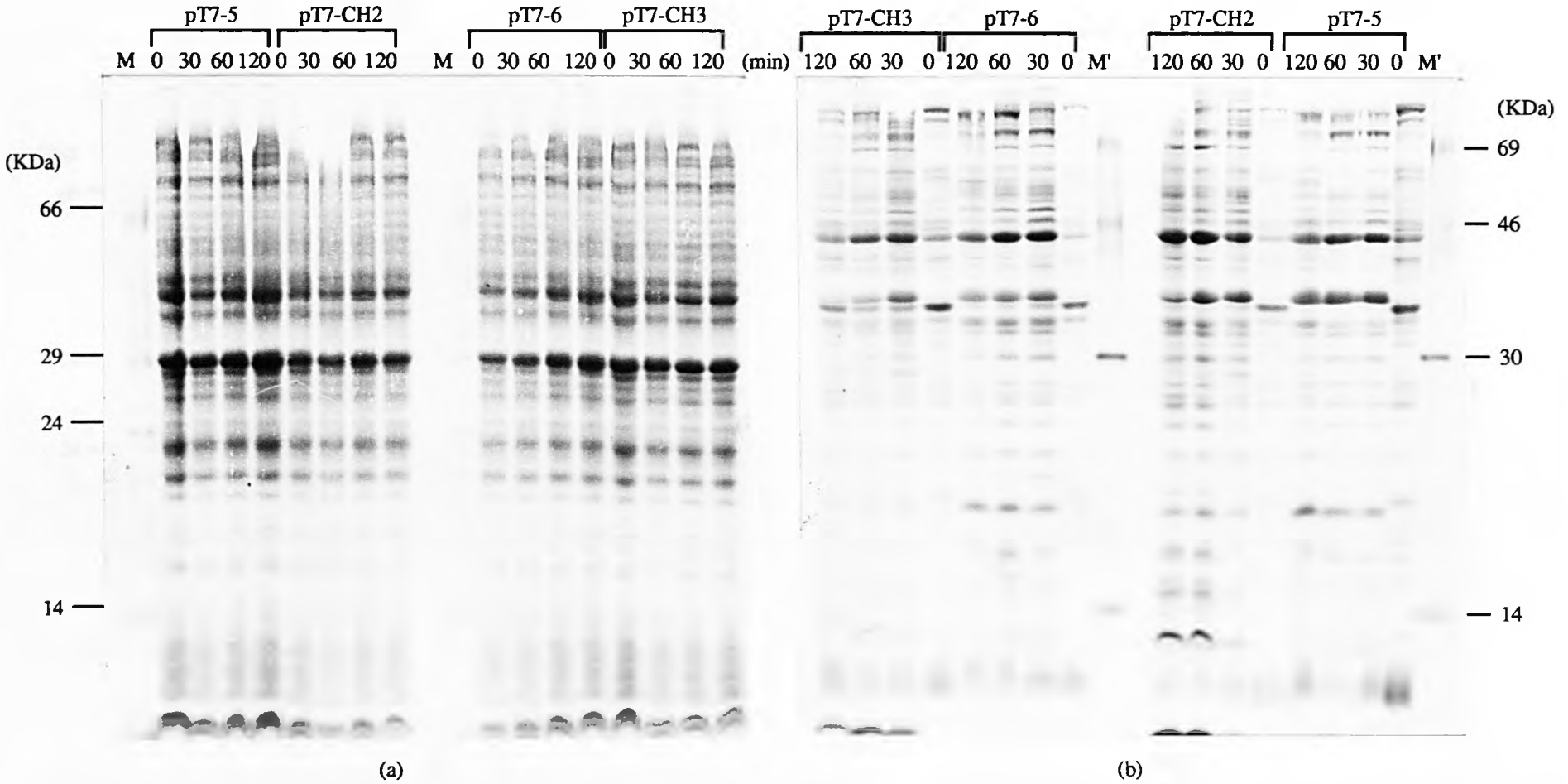


Figure 3.6 *In vivo* inductions of BL21(DE3)pRG1 containing the plasmids pT7-5, pT7-6 and their respective recombinants pT7-CH2 and pT7-CH3. Cultures were grown to late log phase in M9 minimal medium and IPTG was added to a final concentration of 0.4mM (2.2.17). Samples were removed at the times indicated and incubated with ^{35}S-methionine and rifampicin at a final concentration of 0.5 mg/ml (10 min, 37°C). Cell equivalents were analysed by (a) SDS-PAGE and Coomassie blue staining or (b) SDS-PAGE and fluorography as described in 2.2.21-23. M - SDS7 markers (Sigma), M' - ^{14}C markers (Sigma).

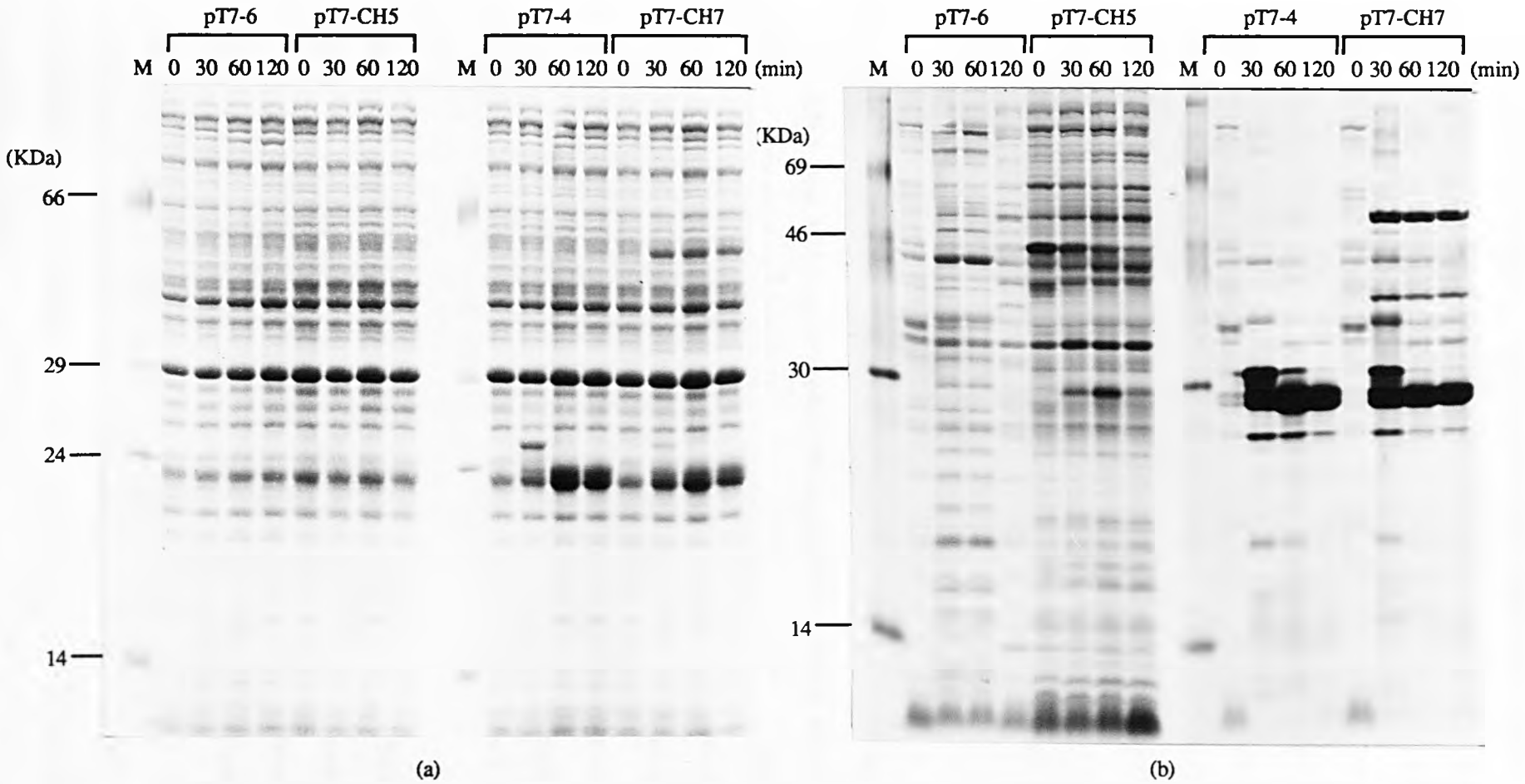


Figure 3.7 *In vivo* inductions of BL21(DE3)pRG1 containing the plasmids pT7-4, pT7-6 and their respective recombinants pT7-CH7 and pT7-CH5. Cultures were grown to late log phase in M9 minimal medium and IPTG was added to a final concentration of 0.4mM (2.2.17). Samples were removed at the times indicated and incubated with ^{35}S-methionine and rifampicin at a final concentration of 0.5 mg/ml (10 min, 37°C). Cell equivalents were analysed by (a) SDS-PAGE and Coomassie blue staining or (b) SDS-PAGE and fluorography as described in 2.2.21-23. M - SDS7 markers (Sigma), M' - ^{14}C markers (Sigma).

visible either in the products from the pT7-CH7 or the pT7-CH5 construct [Figure 3.7 (a) and (b)]. A prominent band apparent in the stained and radioactive profiles at 30KDa corresponds to the β-lactamase protein of the plasmids pT7-4 and pT7-CH7. In the vectors pT7-5 and pT7-6 (Appendix II) and the constructs pT7-CH2 and pT7-CH3 (Figure 3.2) the β-lactamase gene is transcribed in the opposite direction with respect to the T7 promoter and so would not be observed upon induction with IPTG. A protein approximately 54KDa in size was also visible in the stained profile of pT7-CH7 but not of pT7-4 and is most likely CrtI [Figure 3.7 (a)]. The other proteins were not apparent in the stained profile as they are probably masked by other *E. coli* proteins of equivalent sizes. An induced protein, approximately 33KDa in size was apparent in the radioactive profile of pT7-CH5, however it is unclear what this could correspond to as it is too large to be CrtZ.

3.5 RNA synthesis of the *crt* transcript after *in vivo* induction

RNA was extracted and analysed to observe whether transcription of the constructs pT7-CH2 and pT7-CH3 was proceeding after IPTG induction. Total RNA was dot blotted onto a membrane followed by hybridization reactions with probes for *crtI*, *crtB* and *crtE*. Total RNA was extracted from 1ml of IPTG induced cells (IPTG induction was performed as previously, 3.4) taken at each time point indicated in Figure 3.8 and isolated by means of Trizol reagent (2.2.15). As a negative control for the hybridization reaction, RNA was also extracted from the plasmid vectors pT7-4, pT7-5 and pT7-6 at the same time points post IPTG induction. An aliquot was analysed by agarose gel electrophoresis and was found to be contaminated with DNA. The DNA was removed by digestion with DNAseI (0.1U, 30 minutes, 37°C). The RNA content was determined spectrophotometrically and 1μg was dotted onto a nylon (Hybond) membrane and fixed by uv irradiation. The RNA was hybridized with *crtI*, *crtB* and *crtE* probes obtained by PCR and described in Chapter 5.2. Figure 3.8 only depicts results obtained for *crtI*. No hybridization with the probe was observed with the plasmid vectors. Hybridization experiments with both *crtB* and *crtE* gave similar results (data not shown). These results indicate that transcription from the T7 promoter

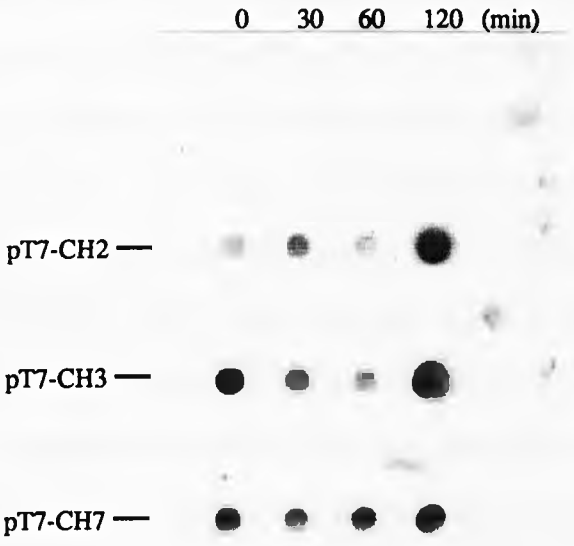


Figure 3.8 Dot blot time courses of the expression of *crt* genes.
BL21(DE3)pRG1 containing the plasmids pT7-CH2, pT7-CH3 and pT7-CH7 was grown and induced with IPTG as previously (Figs. 3.6 and 3.7). Samples were removed at the times indicated, total RNA was extracted and 1μg was used for each dot (2.2.15). The RNA was probed for *crtI* (5.4.2 (ii)) as shown and also *crtB* and *crtE*.

was proceeding as the hybridization signal increases at 120 minutes. Hybridization at time 0 for the constructs pT7-CH2 and pT7-CH3 is indicative of transcription proceeding from the *E. coli* promoters. Hybridization of pT7-CH7 observed at time 0 is possibly due to 'leaky' repression of T7 RNA polymerase.

3.6 Transcription/translation reactions of the *crt* cluster in *vitro*

Plasmid constructs were also analysed *in vitro* to assess whether translation was proceeding. Plasmid DNA was used to prime the *E. coli* RNA polymerase and T7 RNA polymerase dependent *in vitro* transcription translation system (2.2.20) both in the presence and absence of rifampicin. After 30 minutes incubation (37°C) samples were analysed by SDS-PAGE and fluorography (2.2.21 and 2.2.23). As can be seen in Figure 3.9, lanes (ii), (iv) and (v) correspond to the plasmids pAPU211, pAPU211Z and pCH6 respectively. Bands are apparent corresponding to molecular weights of approximately 36 and 54KDa proteins which are absent from the analysis of vector alone [Figure 3.9 lane (i)]. These could correspond to the CrtB and CrtI protein products respectively. There may also be a band at 33KDa (*crtE* gene product) which is difficult to see as it could be masked by other bands. The other bands that would be of interest, the 45, 44KDa and 20KDa proteins corresponding to the CrtX and CrtY and CrtZ proteins, are not visible.

In vitro transcription translation experiments were also performed on the T7 cloning vectors pT7-4, pT7-5 and pT7-6 (as negative controls) and the recombinants pT7-CH2, pT7-CH3 and pT7-CH7 (2.2.20). Reactions were primed and analysed as previously and the results are also given in Figure 3.9. The bands in lanes (viii) pT7-CH2, (x) pT7-CH3 and (xiii) pT7-CH7 are the same molecular weight bands as those observed for pAPU211, pAPU211Z and pCH6 and again could correspond to CrtB and CrtI protein products. It is not clear why the other Crt proteins were not visible.

3.7 Quantitative analysis of carotenoid production

Total carotenoid content of the transformed cells was analysed. This was performed to answer three questions: (i) is there a decrease in carotenoid content of *E.*

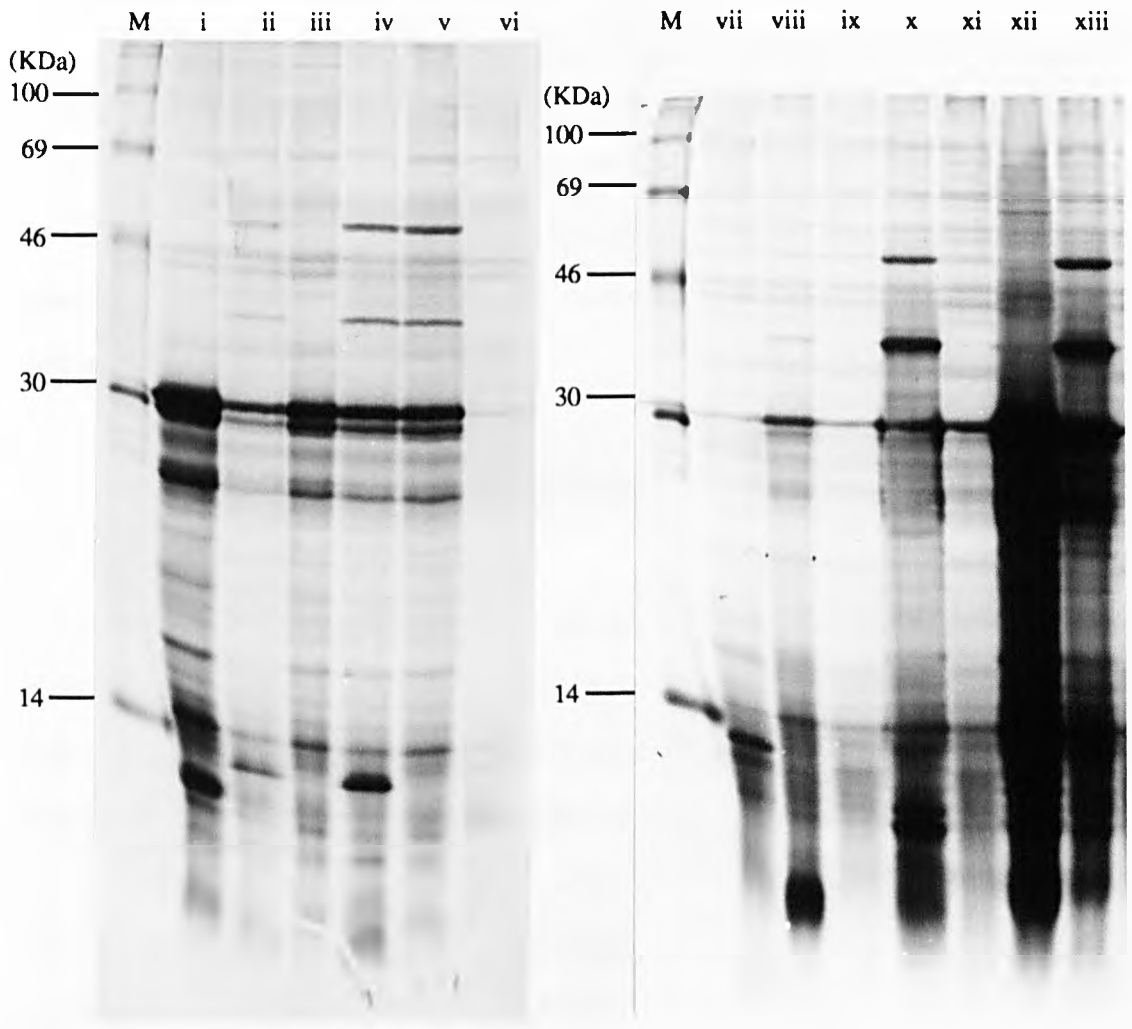


Figure 3.9 *In vitro* expression of carotenoid biosynthesis enzymes.
In vitro transcription-translation reactions were carried out by using the following plasmid templates; (i) pUC19, (ii) pAPU211, (iii) pCH4, (iv) pAPU211Z, (v) pCH6, (vi) no plasmid, (vii) pT7-5, (viii) pT7-CH2, (ix) pT7-6, (x) pT7-CH3, (xi) pT7-CH5, (xii) pT7-4, (xiii) pT7-CH7 as described in 2.2.20. Reactions vii-xiii were carried out in the presence of rifampicin (0.5mg/ml). All reactions were analysed by SDS-PAGE and fluorography (2.2.23). M - ^{14}C markers (Sigma).

coli cells transformed with the plasmid pCH6?, (ii) is there any increase in carotenoid production at all in *E. coli* cells transformed with the plasmids pT7-CH2 and pT7-CH3 when compared to pAPU211?, (iii) does BL21(DE3)pRG1 pT7-CH7 show any carotenoid production upon induction with IPTG?

In vivo induction with IPTG was carried out exactly as previously (3.4) and carotenoids were extracted from 1ml aliquots at the same time points. Samples at each time point were obtained in triplicate. Carotenoids were extracted from cell equivalents at each time point. The carotenoid extract was analysed by a scanning spectrophotometer and the carotenoid content was determined by applying the following equation;

$$A^{1\%1cm}=2500, \text{ and therefore for } 1\mu g/ml,\ A=0.25$$

Figure 3.10 shows a bar chart of the results obtained. The highest carotenoid content was obtained with *E. coli* transformed with the plasmid pAPU211. Placing five or one *crt* gene under the control of T7 did not produce greater levels of carotenoid than pAPU211. Placing all six genes under the control of the T7 promoter resulted in no carotenoid biosynthesis at all even after induction with IPTG. The six genes transcribed in the same direction, on the plasmid pCH6 did produce carotenoid but at lower levels than pAPU211. With all the transformed strains of *E. coli* higher carotenoid levels were detected than in *E. herbicola*.

3.8 Carotenoid biosynthesis in stationary phase

RpoS is a stationary phase specific sigma factor required for the expression of many genes as cells enter and continue in stationary phase. The orange colour of the transformed *E. coli* cells was not apparent after overnight incubation at 37°C but appeared after further incubation i.e. in stationary phase. It was thought that the *crt* cluster may be reliant on *rpoS* for expression. The plasmids pAPU211, pT7-CH2 and pT7-CH3 were used to transform *E. coli* strains carrying wild type *rpoS* (ZK126) and an *rpoS* deficient strain (ZK1000). The wild type transformants were orange in colour

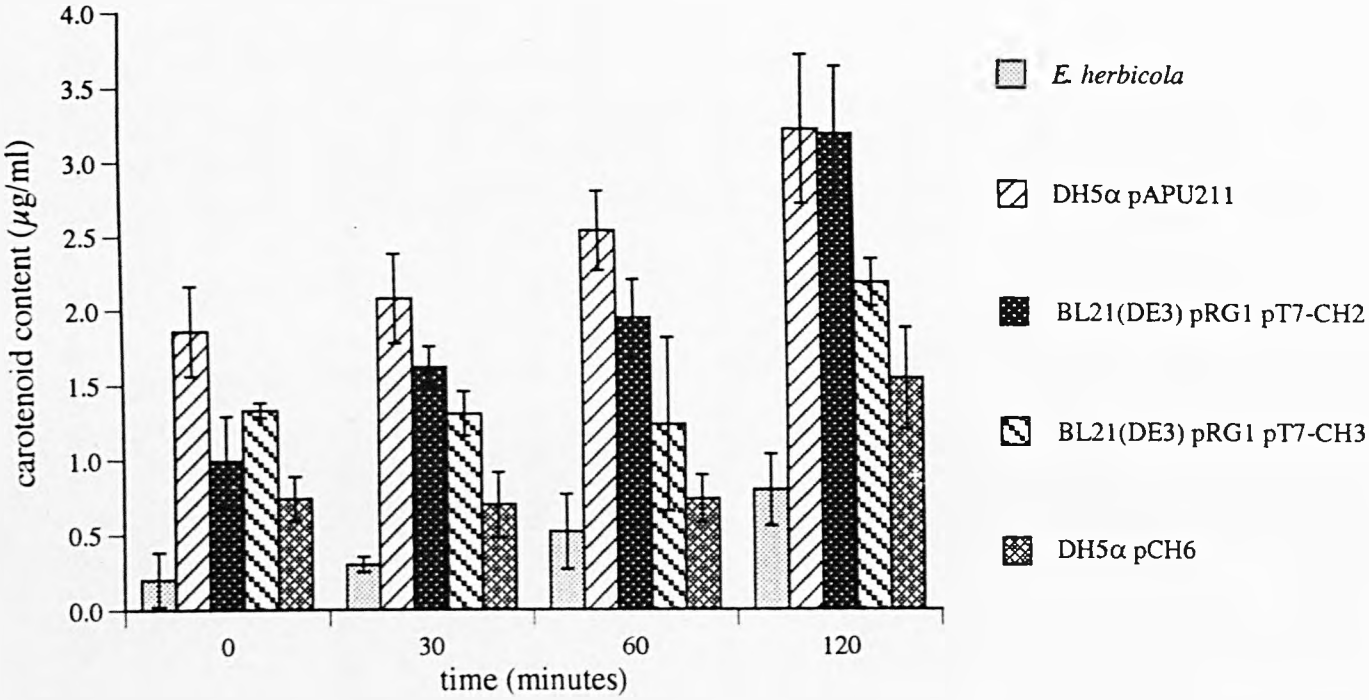


Figure 3.10 Quantitative carotenoid analysis of *E. herbicola* and *E. coli* transformed with the plasmids, pAPU211, pT7-CH2, pT7-CH3 and pCH6. Carotenoids were extracted as described in Section 2.2.26 (i) from cell equivalent samples in triplicate. Samples for pT7-CH2 and pT7-CH3 were taken at the indicated time points post IPTG induction (3.4), the other samples were taken at the same time points. Cell growth was monitored by optical density readings at A_{600nm}, at time zero all A_{600} readings were equivalent at 0.9. Standard deviations are indicated by error bars.

whereas the *rpoS* deficient transformants were white. This suggests that the *crt* cluster is reliant on *rpoS* for expression.

3.9 Discussion

The T7 RNA polymerase expression system is widely used for increasing gene expression. This system was employed with the aim of increasing *crt* gene expression and consequently increasing carotenoid production.

Unusually after the genes were placed under the control of the T7 RNA polymerase no proteins from the constructs pT7-CH2 and pT7-CH3 were visible upon induction of the polymerase with IPTG. To determine whether the problem was with transcription, RNA was analysed at specific time points after induction. The hybridization of probes *crtE*, *crtB* and *crtI* with the dot blotted RNA at all time points indicated that transcription was proceeding, suggesting that translation of the *crt* transcript was the problem.

Endogenous *E. coli* promoters have not been removed from the constructs pT7-CH2 and pT7-CH3. The *crt* genes are constitutively expressed and are not directly under the control of T7 until IPTG induction (although expression due to low levels of T7 RNA polymerase expression cannot be ruled out), at which point the presence of rifampicin will inhibit any *E. coli* RNA polymerase activity. Therefore carotenoids are being synthesized before induction. Tnis almost certainly means that RNA is being transcribed by the endogenous *E. coli* RNA polymerase and subsequently translated from the *crt* transcript. The distance between the promoter sequence and initiation codon is approximately 35bp which is considerably smaller than the distance observed with the T7 constructs (approximately 1.5Kb). It was thought that the 1.5Kb distance between the T7 promoter and the AUG initiation codon in the plasmids pT7-CH2 and pT7-CH3 could lead to RNA secondary structure formation. The observation oi translation occurring from the endogenous *E. coli* promoters supports the theory that the 1.5Kb distance is too long for efficient, if any, translation of the transcript to occur due to RNA secondary structure formation. This secondary structure may prevent ribosome binding and thus translation.

Turning the *crtZ* gene round and placing it in front of *crtE* in the plasmid pAPU211Z enabled transcription of all genes in the same direction. By placing the *crt* cluster under the control of the T7 promoter in the vector pT7-4 the endogenous

promoters were removed. Expression of the *crt* genes was therefore solely reliant on T7 RNA polymerase. In this construct, expression of, at least a number of, the Crt enzymes occurred but carotenoid biosynthesis was absent. It may be that the proteins are produced in such high quantities that misfolding occurs, inhibiting complex formation, membrane attachment (if required) and function. The distance between the T7 promoter sequence and the AUG initiation codon is approximately forty bases in pT7-CH7 thus significant secondary structure should be absent in this construct. Secondary structure effects on the efficiency of ribosome function have been observed before (Looman *et al.*, 1986), where the efficiency of ribosome binding (and subsequent protein expression) were correlated with the lack of secondary structure. Alterations 5' to the ribosome binding site significantly reduced the rate of mRNA translation, in a vector series specifically constructed to analyse this hypothesis (Stanssens *et al.*, 1985). RNA secondary and/or tertiary structure was suggested to be the reason for this observation.

With the construct pCH6 in which *crtZ* is transcribed in the same direction as the other five *crt* genes, noticeably smaller amounts of carotenoids were synthesized, when compared to pAPU211. It is therefore possible that *crtZ* is transcribed in the opposite direction as a controlling feature of carotenoid biosynthesis. The CrtZ protein was also never observed by SDS-PAGE, but this could be due to its being rapidly degraded or being very hydrophobic, as indicated by the hydropathy plots shown in Figure 4.1. Hundle *et al.* (1993) did observe CrtZ on SDS-PAGE after *in vitro* expression. This is not directly comparable with the *in vivo* experiments performed in this Chapter, as in Hundles work, the enzyme was expressed individually and not with the other Crt enzymes which may play a crucial role in the enzymes expression and activity of CrtZ.

T7 RNA polymerase directed transcription/translation reactions were performed *in vitro* to discover if any proteins could be translated from the *crt* transcripts in the constructs pT7-CH2 and pT7-CH3. The results showed the presence of two proteins corresponding to CrtI and CrtB in all the constructs made and also in the plasmids pAPU211 and pAPU211Z. The presence of these two proteins contradicts the idea that

RNA secondary structure, may be preventing translation. Because all the constructs gave this result, secondary structure does not seem to be a factor in their expression *in vitro* and the reason for this is unclear. Possibly some intrinsic property of the *in vitro* system or the rapid degradation of the proteins *in vivo* could be responsible for not observing protein expression. However proteins are less rapidly degraded *in vitro*. Alternatively a protein present *in vivo* but not *in vitro* may bind to RNA secondary structure and occlude ribosome binding sites. However this does not explain only ever observing two proteins with these reactions.

Carotenoid content was measured in triplicate for all the constructs, pAPU211 and *E. herbicola*. The highest carotenoid production was observed with the plasmid pAPU211 when expressed in *E. coli*. This means that the T7 cloning procedures did not increase carotenoid production in *E. coli* when compared to the plasmid pAPU211. Carotenoid production increased over time with the constructs pT7-CH2 and pT7-CH3, however this could be due to the presence of existing enzymes. Alternatively synthesis may be due to newly synthesized enzymes (T7 directed) as earlier enzymes have degraded. Either postulate is possible and it is unclear from these experiments which is correct, though one would tend to favour the former explanation as the Crt enzymes were not observed from these constructs. However the levels of protein needed for carotenoid biosynthesis are unknown and may be minimal. If this is the case although the proteins were not observed from these constructs, levels may be sufficient for biosynthesis of carotenoids. All the constructs showed increased carotenoid production when compared to *E. herbicola*, which suggests that simply increasing gene copy number leads to greatly increased carotenoid production, even in heterologous hosts.

Overproducing and underproducing strains of *Thermus thermophilus* HB27 have recently been isolated (Hoshino *et al.*, 1994b) and one overproducing strain was transformed with a multicopy plasmid carrying the gene *crtB* of *T. thermophilus* HB27. This transformed strain produced approximately twenty times more carotenoid than the parental strain (Hoshino *et al.*, 1994a). This suggests a further experiment to atempt to increase carotenoid production could be to transform *E. herbicola* with the plasmids pAPU211, pCH6, pT7-CH2, pT7-CH3 and pT7-CH7. Obviously T7 RNA

polymerase would have to be cotransformed under the control of an inducible promoter alongside the T7 constructs as this gene will not be present on the *E. herbicola* chromosome. It is possible that similar results to those observed for *T. thermophilus* would be obtained. The addition of 0.1% mevalonic acid to the culture media of the yeast *Phaffia rhodozyma* stimulated astaxanthin synthesis with total carotenoid biosynthesis increases of 400% (Calo *et al.*, 1995). Experiments such as this are simple to execute and could lead to the required end result of increased carotenoid production, negating the need for genetic manipulation.

Carotenoid biosynthesis appeared to be dependent on the stationary phase sigma factor, RpoS, as no carotenoids were apparent in the RpoS deficient strain transformed with the plasmid pAPU211. The RpoS sigma factor shows certain similarities to RpoD (the major sigma factor of *E. coli*) that suggest similar mechanisms of recognizing and binding to promoter sequences. The consensus promoter sequences recognized by RpoD are TATAAT and TTGACA at positions -10 and -35 respectively, upstream from the start site of transcription. A recent report has described a conservation of sequence at the -35 position of RpoS dependent genes (Wise *et al.*, 1996) where cytosine nucleotides replace the thymidine nucleotides at the 5' end. The start site of transcription has not been determined with the *E. herbicola* cluster, although it has been determined in *E. herbicola* Eho13 (To *et al.*, 1994) and lies 147bp upstream from *crtE*, so it is difficult to say whether an RpoS promoter sequence is present or whether expression of *crt* genes is dependent upon a protein whose expression is RpoS-dependent. Inspection of the *E. herbicola* sequence does not reveal any striking RpoS promoter sequences, although these are not easy to identify since they are rather degenerate.

In conclusion the *crt* genes were successfully cloned such that their expression was under the control of the T7 promoter. However this did not facilitate increased carotenoid production when compared to the *crt* cluster under the control of the *lac* promoter in the plasmid pAPU211. Carotenoid production was increased in all *E. coli* strains carrying the *crt* cluster on plasmid constructs compared to the *E. herbicola*

chromosomal copy. Increasing the expression of protein is a feasible approach, but this will not necessarily mean an increase in carotenoid levels, as observed with the strategies employed here. A novel approach needs to be employed if carotenoids are to be mass produced by bacteria.

CHAPTER 4

CELLULAR LOCATION OF THE ENZYMES OF CAROTENOID BIOSYNTHESIS

4.1 General introduction

The various enzymic steps of carotenoid biosynthesis in *E. herbicola* (Figure 1.9) have been assigned to specific loci within the gene cluster and studied in detail. It has been generally accepted that the enzymes of carotenoid biosynthesis are localized in membranes. The carotenoid biosynthesis enzymes of *E. herbicola* are thought to be organized into a multi-enzyme complex that is associated with, bound to or even situated within the membrane. When the *crt* genes of *E. herbicola* are expressed in *E. coli* the same membrane associations are presumed to occur (Fraser and Sandmann, 1992). However nothing is known about the enzyme-membrane and enzyme-enzyme associations that would be essential for such a complex to function.

CrtB or phytoene synthase is a 34KDa bifunctional enzyme which performs two sequential reactions synthesizing phytoene from GGDP. It is similar in sequence to pTOM5 encoding phytoene synthase that is differentially expressed during tomato fruit ripening (Armstrong *et al.*, 1990).

Phytoene dehydrogenase is the largest of the enzymes at 54KDa. It has a high amino acid sequence identity (41.7%) with *R. capsulatus* CrtI and also shows conservative substitutions. This one enzyme performs four sequential dehydrogenations that convert phytoene into lycopene (Armstrong *et al.*, 1990). The phytoene dehydrogenase enzymes from *R. capsulatus*, *E. herbicola* and *N. crassa* all possess a hydrophobic N-terminal domain containing a putative ADP-binding βαβ fold characteristic of enzymes that are known to bind FAD or NAD(P) cofactors (Armstrong *et al.*, 1993). CrtI has been purified from inclusion bodies and its activity restored upon the removal of urea. FAD was involved in its activation and NAD(P) was inhibitory. The FAD was not covalently bound to the enzyme (Fraser *et al.*, 1992).

The *crtE* gene encodes the enzyme geranylgeranyl diphosphate synthase which is 33KDa in size. The enzyme has 55% and 30% amino acid identity to those in *E. uredovora* and *R. capsulatus*, respectively. The gene has been isolated from the *crt* gene cluster by PCR amplification and cloned into an *E. coli* expression vector (Math *et al.*, 1992). The function of *crtE* was ascertained by using radiolabelled substrates and monitoring reaction products by HPLC and TLC. Weidmann *et al* (1993) have overexpressed and purified the enzyme from *E. uredovora* and discovered that replacement of the first 13 amino acids by 16 heterologous amino acids did not significantly affect activity.

Zeaxanthin glycosylase (CrtX), an enzyme of 45KDa, has been expressed *in vitro* under the control of an inducible T7 promoter (Hundle *et al.*, 1992) and has 45% similarity to that of *E. uredovora.* The enzyme requires UDP-activated substrates and has a distinct region of homology with six other UDP-glucosyl and glucoronosyl transferases. This homologous region is thought to correspond to the UDP-binding motif. Cell fractions were assayed and activity was present in the washed membrane and supernatant fractions, but activity was five-fold higher per mg of protein in the membrane than in the supernatant. The amino acid sequence has been analysed by a Kyte and Doolittle hydropathy plot (Kyte and Doolittle, 1982). Two regions were discovered that have primarily hydrophobic residues but are too small to form transmembrane helices. It was proposed that these could serve as anchors to associate the protein with the membrane. As activity is seen in both the membrane and supernatant fractions it was decided that Crt X is loosely associated with the membrane and is not an integral membrane protein.

Lycopene cyclase (43KDa) and β-carotene hydroxylase (22KDa) have also been placed under the control of the T7 promoter and their activities expressed *in vitro* (Hundle *et al.*, 1993). The enzymes are 58% and 73% similar, respectively, to those of *E. uredovora.* Lycopene cyclase was found to be more hydrophilic based on amino acid composition than the β-carotene hydroxylase and exhibited enzymic activity in a crude cell-free lysate although to a lesser extent than did β-carotene hydroxylase.

After cell fractionation β-carotene hydroxylase activity was found only in the supernatant and the membrane fraction was inactive although a hydropathy plot showed several regions of hydrophobicity consistent with a membrane-bound enzyme. β-Carotene hydroxylase catalyses two hydroxylations to synthesise zeaxanthin from β-carotene via the intermediate β-cryptoxanthin. The enzyme is thought to have a mixed function oxygenase mechanism as it can proceed in the absence of externally added cofactors, but electron donors such as NAD and NADPH stimulate its activity. This could explain why no enzyme activity is seen in the membrane fraction, as complex redox reactions could be disrupted on cell fractionation. The Eho 10 *crt* cluster expressed in *E. coli* and *R. sphaeroides* may facilitate the understanding of enzyme assembly and organization in carotenoid biosynthesis (Hunter *et al.*, 1994).

The aim of the work described in this chapter was to discover the location of the enzymes. This was attempted by use of the T7 RNA expression system followed by cell fractionation. The production of antibodies against CrtY and CrtI was also employed in an attempt to determine the location of these two enzymes.

4.2 Predicting possible membrane spanning regions

Good estimations of the potential of a protein to associate with a membrane can be made by the use of hydropathy plots. The six carotenoid biosynthesis enzymes of *E. herbicola* were analysed by means of the University of Wisconsin Genetics Computer Group (GCG) program Pepplot. This program showed possible hydrophobic areas of each protein which could be indicative of membrane association. Hydrophilicity of the proteins was also indicated. Figure 4.1 illustrates the results of these analyses. A large hydrophobic stretch of amino acids was observed with the protein CrtZ. CrtE contained large areas of hydrophilic amino acids. The other proteins CrtX, CrtY, CrtB and CrtI showed a similar number of hydrophobic and hydrophilic residues. The difficulty encountered in the interpretation of hydropathy plots is that areas of hydrophobicity within a protein do not necessarily mean membrane association. The hydrophobic region of CrtZ could serve a number of functions. It could be involved in membrane association or membrane anchoring but it could also represent a region necessary for substrate binding i.e. the active site of the enzyme.

4.3 Maxi-cell *in vivo* expression

Initially, in the absence of antisera to individual *crt* gene products, the gene expression system of 'maxi-cells' was used as a means of identifying the plasmid-encoded proteins and their cellular location.

The *recA* gene product of *E. coli* is required for the repair of uv-damaged DNA. *E. coli* strains carrying a *recA* mutation will not be able to repair chromosomal DNA damaged by uv irradiation. If cells are incubated in the dark after uv exposure, alternative photoactivated repair pathways (*phr*) are also inhibited. Such uv-damaged DNA is subsequently degraded by intracellular nucleases. This results in minimal protein synthesis due to the lack of transcribable DNA.

If such strains carry a multicopy plasmid, many plasmids will survive uv-irradiation due to their smaller target size and high copy number. These plasmids can serve as templates for transcription and so plasmid-encoded proteins can be identified by preferentially labelling them with ^{35}S-methionine.

Figure 4.1 Hydropathy plots of the carotenoid biosynthesis gene products of *E. herbicola*. Analysis was performed by means of the University of Wisconsin Genetics Computer Group package (Devereux *et al.*, 1984).

Figure 4.1 (continued)

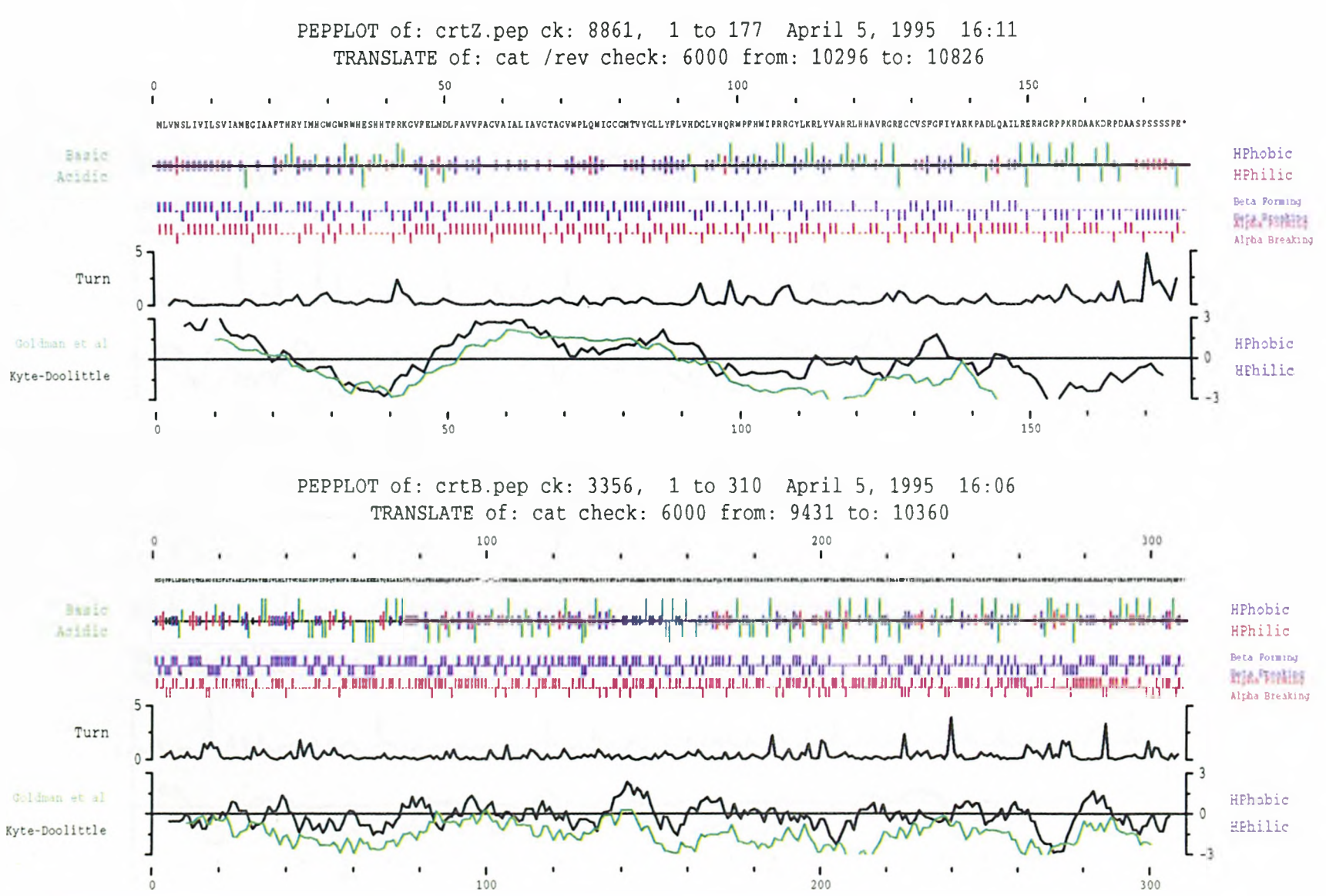
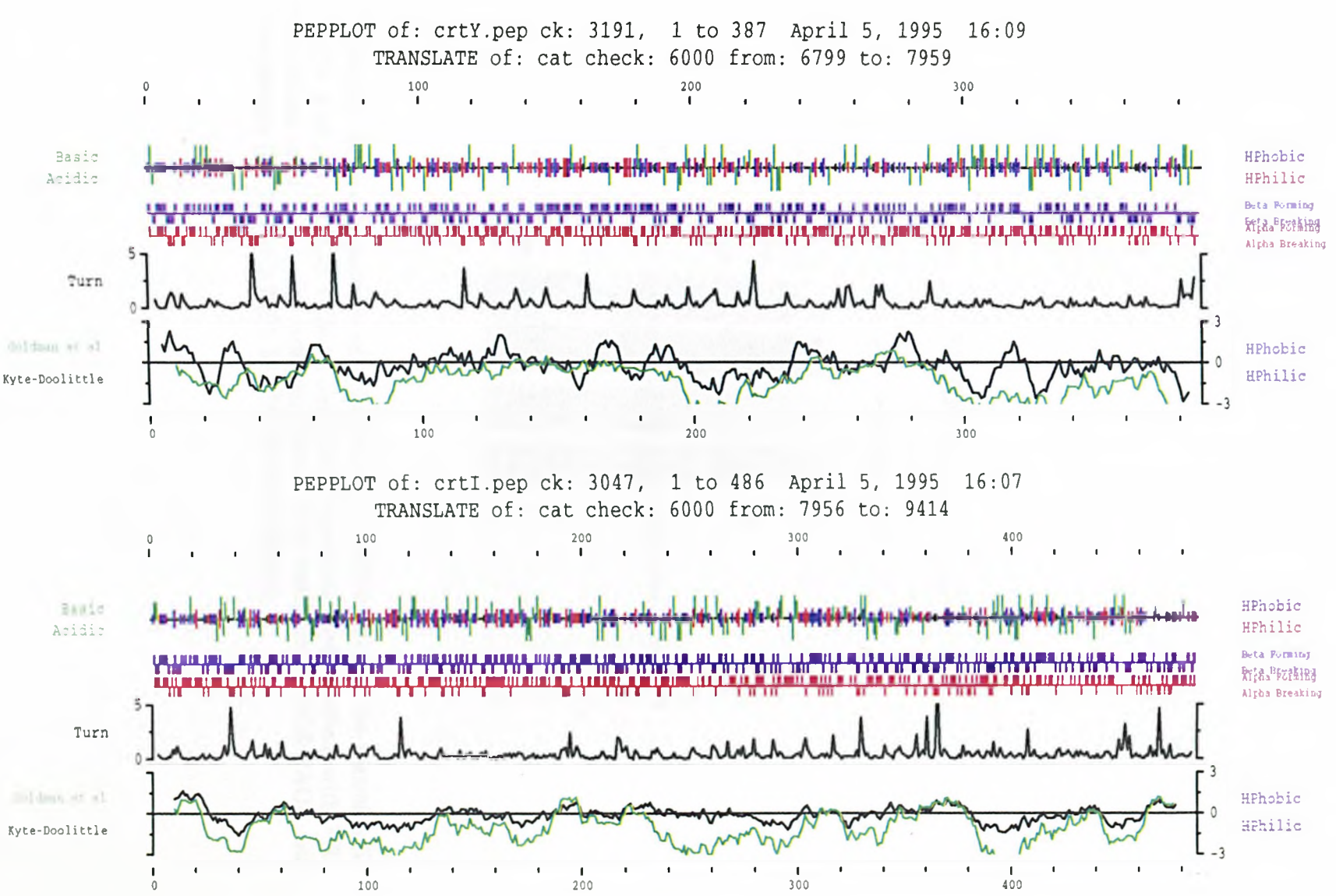


Figure 4.1 (continued)



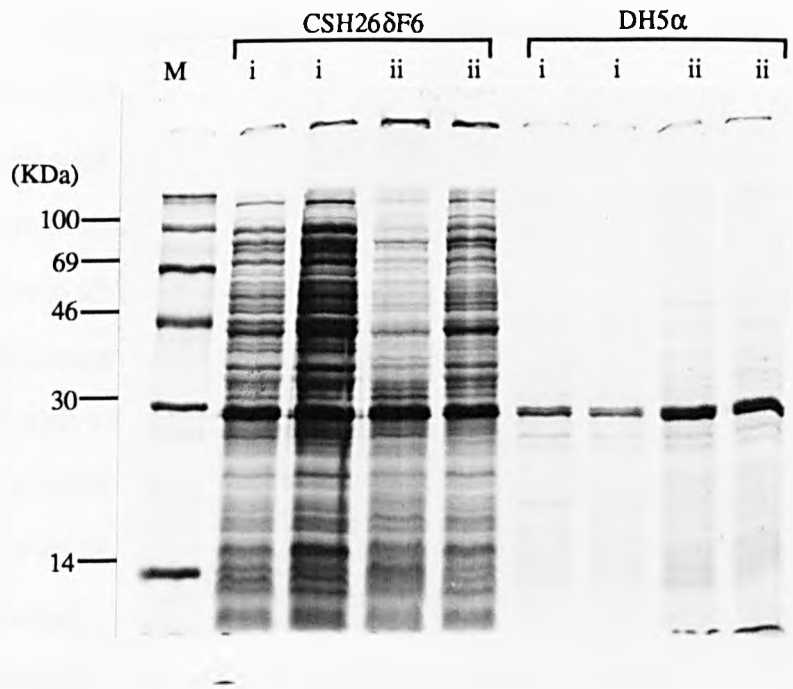


Figure 4.2 'Maxi-cell' proteins of CSH26ΔF6 and DH5α carrying the plasmids (i) pUC19 and (ii) pAPU211. Maxi-cells were produced and labelled with ^{35}S-methionine as in 2.2.16 and cell equivalents were analysed by SDS-PAGE and fluorography (2.2.21 and 2.2.23). M - ^{14}C markers (Sigma).

Two strains were used, namely CSH26ΔF6 and DH5α, both of which are *recA*. The strains were transformed with the plasmids pUC19 and pAPU211 (2.2.1).

Cultures of transformed CSH26ΔF6 and DH5α were grown till an A_{600} of 0.4 was reached, irradiated for 2 minutes at 254nm and labelled with ^{35}S-methionine as described in 2.2.16. The labelled proteins were analysed by SDS-PAGE and fluorography (2.2.21 and 2.2.23) and are depicted in Figure 4.2. In the case of the CSH26ΔF6 [lane (ii)] no distinctly labelled proteins were visible in the cells transformed with pAPU211 when compared with pUC19 (lane (i), used as a negative control). Neither plasmid gave discrete bands distinguishable from *E. coli* proteins. DH5α transformed with pAPU211 did show proteins that were not present with pUC19 and were distinct from the background of *E. coli* proteins [Figure 4.2 lane (ii)]. These were of the correct sizes to correspond to the Crt proteins. The signal was weak, however even after seven days of exposure, so these systems were not really suitable to use in localization experiments.

Longer exposure times of uv irradiation of the CSHΔF6 cells did not result in a decrease of labelled proteins i.e. a reduction in background. Decreasing the length of exposure with DH5α led to an increase in numbers of labelled proteins i.e. an increase in background. Increasing the time of exposure reduced the background but bands were still faint. An optimum exposure was found but the signal produced was too weak to enable localization.

4.4 Cell fractionation

The expression of the carotenoid biosynthesis enzymes was weak in maxi-cells and consequently the proteins could not be localized by using this system. The recombinant clone pT7-CH7 (Figure 3.5) showed high level expression of the enzymes so cell fractionation of *E. coli* containing this plasmid could enable localization experiments to proceed.

Cell fractionation experiments were performed on an aliquot of ^{35}S-methionine labelled cells of BL21(DE3)pRG1 pT7-CH7 at 60 and 120 minutes post IPTG

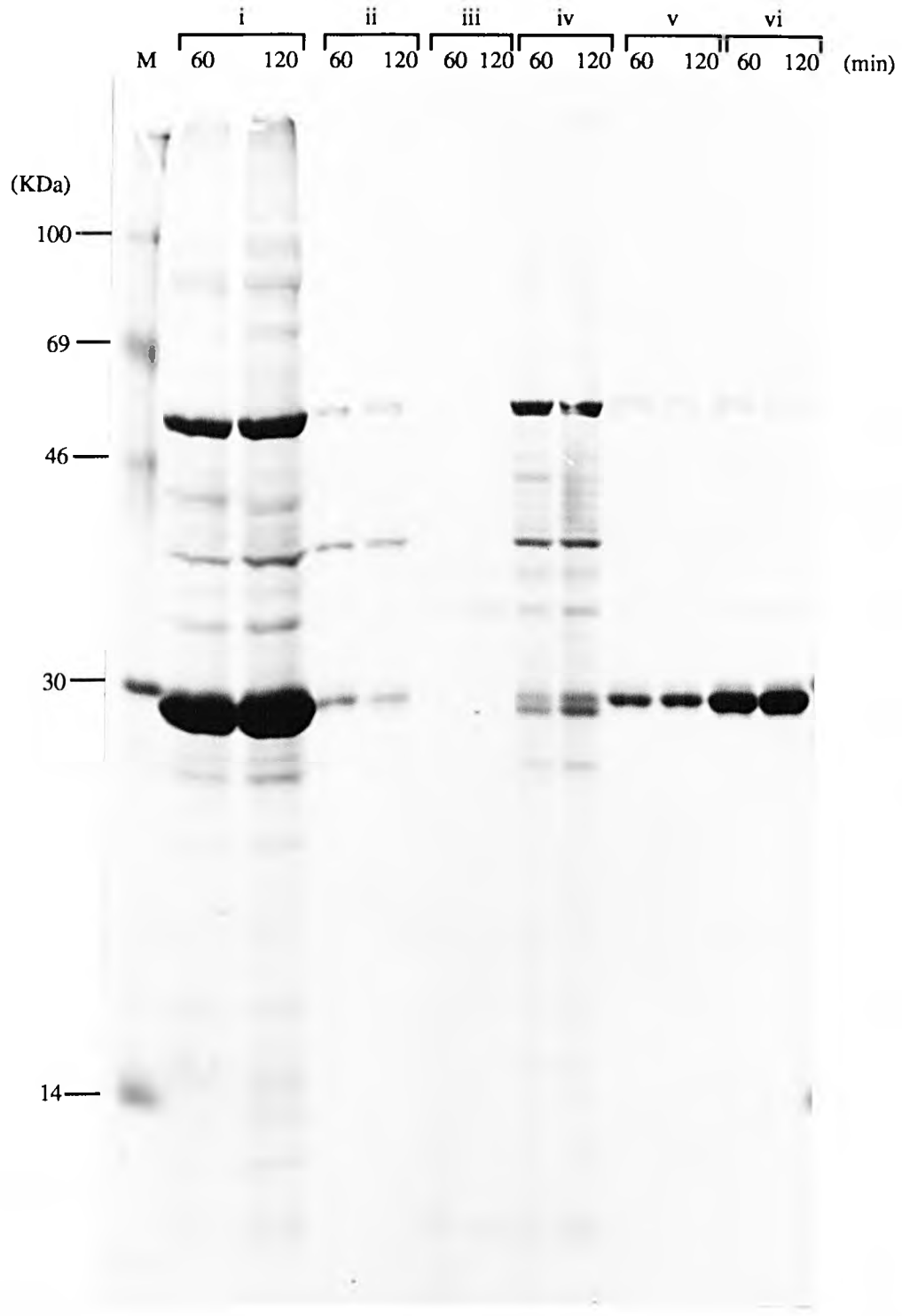


Figure 4.3 Cell fractions of BL21(DE3)pRG1 pT7-CH7.
In vivo inductions of BL21(DE3)pRG1 pT7-CH7 were performed as in 2.2.17. Cell equivalents obtained at 60 and 120 minutes post induction were separated into the following fractions (i) cytoplasm and periplasm, (ii) inner membrane, (iii) outer membrane, (iv) total cells minus periplasm, (v) and (vi) periplasm (2.2.18 and 2.2.19), and analysed by SDS-PAGE and fluorography (2.2.21 and 2.2.23). M - ^{14}C markers (Sigma).

induction (Figure 3.7) as described in 2.2.18 and 2.2.19. The proteins present in the cytoplasm, periplasm, inner and outer membranes were all then analysed by SDS-PAGE and fluorography (2.2.21 and 2.2.23). The results of the analysis are given in Figure 4.3. The protein band at 30KDa corresponds to β-lactamase, the ampicillin resistance protein encoded in the vector pT7-4. This protein is located in the periplasm as indicated in lanes (v) and (vi). The protein at 54KDa corresponding to CrtI is also apparent in the periplasmic fractions [lanes (v) and (vi)] as well as in the inner membrane [lane (ii)] and cytoplasmic fractions [lane (i)]. A protein of approximately 36KDa, likely to correspond to CrtB, is present in the outer membrane [lane (iii)] and cytoplasmic fractions [lane (i)]. Protein corresponding to 33KDa (CrtE) is observed in the outer membrane [lane (iii)], cytoplasm [lane (i)] and the periplasmic fractions [lanes (v) and (vi)]. There is a 45KDa protein (CrtX) present only in the cytoplasmic fraction [lane (i)]. A protein at 41KDa (CrtY) is present in the inner membrane [lane (ii)] and cytoplasm [lane (i)]. All the Crt enzymes with the exception of CrtX appear to have some association with the membrane, but are mainly located to the cytoplasm. It is possible that the membrane association is genuine and not a result of cross contamination of cell fractions which always occurs in low levels. It could be postulated that increasing protein expression may not give accurate results of precise cell location. The use of antibodies negates the need for genetic manipulation to enable visualization and so location of normally expressed proteins could be achieved.

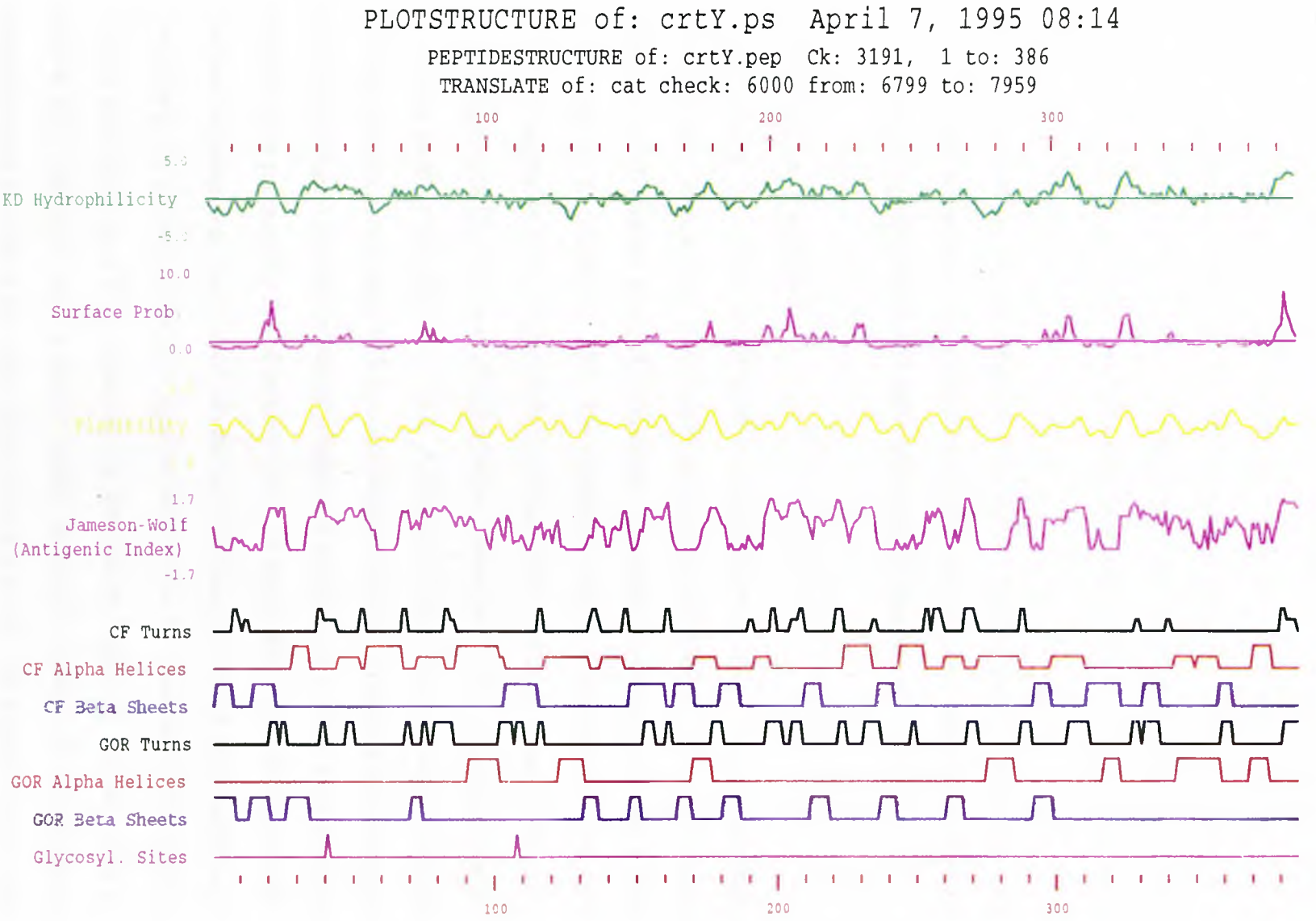
4.5 The generation of polyclonal antibodies against CrtY and CrtI

(i) Designing polypeptides

The University of Wisconsin GCG package was used in the design of suitable polypeptides to be used as antigens for the production of antibodies. Figure 4.4 depicts the results of these analyses. The polypeptide needs to fulfil a number of requirements. It needs to be in a fairly hydrophilic area of the protein and also needs to be on or near the surface i.e. not embedded within domains of the protein. The amino acids need also to be antigenic ensuring that an immune response is elicited. The 10-mers chosen for

Figure 4.4 Peptide structure plots of the carotenoid biosynthesis gene products, CrtY and CrtI of *E. herbicola*. Analysis was performed by means of the University of Wisconsin Genetics Computer Group package (Devereux *et al.*, 1984).

Figure 4.4 (contiued)



the proteins phytoene dehydrogenase (CrtI) and lycopene cyclase (CrtY) fit these prerequisites. An extra three amino acids were added to the end to allow linkage to keyhole limpet haemacyanin (KLH) and to give a distance between the polypeptide and KLH so that the polypeptide is exposed as much as possible. The polypeptides were synthesized as described [2.2.24 (i)]. Before injection the polypeptides were linked to KLH, purified by gel filtration chromatography and mixed with Freunds complete adjuvant [2.2.24 (ii)].

Two chickens were immunized with each antigen and eggs were collected daily from the date of injection. Appendix III shows the dates of collection and how the eggs were grouped together before antibodies were purified from the egg yolk.

(ii) Titration of antisera

Figure 4.5 shows the titration of antisera raised against the two proteins. Proteins were separated by SDS-PAGE of the BL21(DE3)pRG1 pT7-CH7 sample obtained at 120 minutes post IPTG induction. The proteins were blotted onto PVDF membranes as described in 2.2.25. The membrane was cut into strips and each strip was then incubated with grouped antisera obtained from the eggs at sequential time-points (Appendix III). No band of hybridization was seen for CrtY indicating that no antibodies were present [Figure 4.5 (a)]. A band of 54KDa in size is clearly observed with the antisera raised against CrtI at approximately twenty-five days post-injection and five days post booster injection [Figure 4.5 (b) lane (v)]. This indicates that antisera were successfully raised against CrtI. To verify this, and to ensure that no immune response was seen with CrtY, the antisera were incubated with membranes blotted with the plasmid vector (as a negative control) and the recombinant pT7-CH7 at various time points post IPTG induction (as depicted in Figure 3.7 and described in section 3.4). The CrtI band is clearly visible in the 30, 60 and 120 minutes post induction lanes and not in the vector lanes [Figure 4.6 (b)]. No discrete bands were visible for CrtY [Figure 4.6 (a)], thus supporting the theory that antisera were not obtained against this protein.

High background was observed for the CrtI antisera. This was decreased by

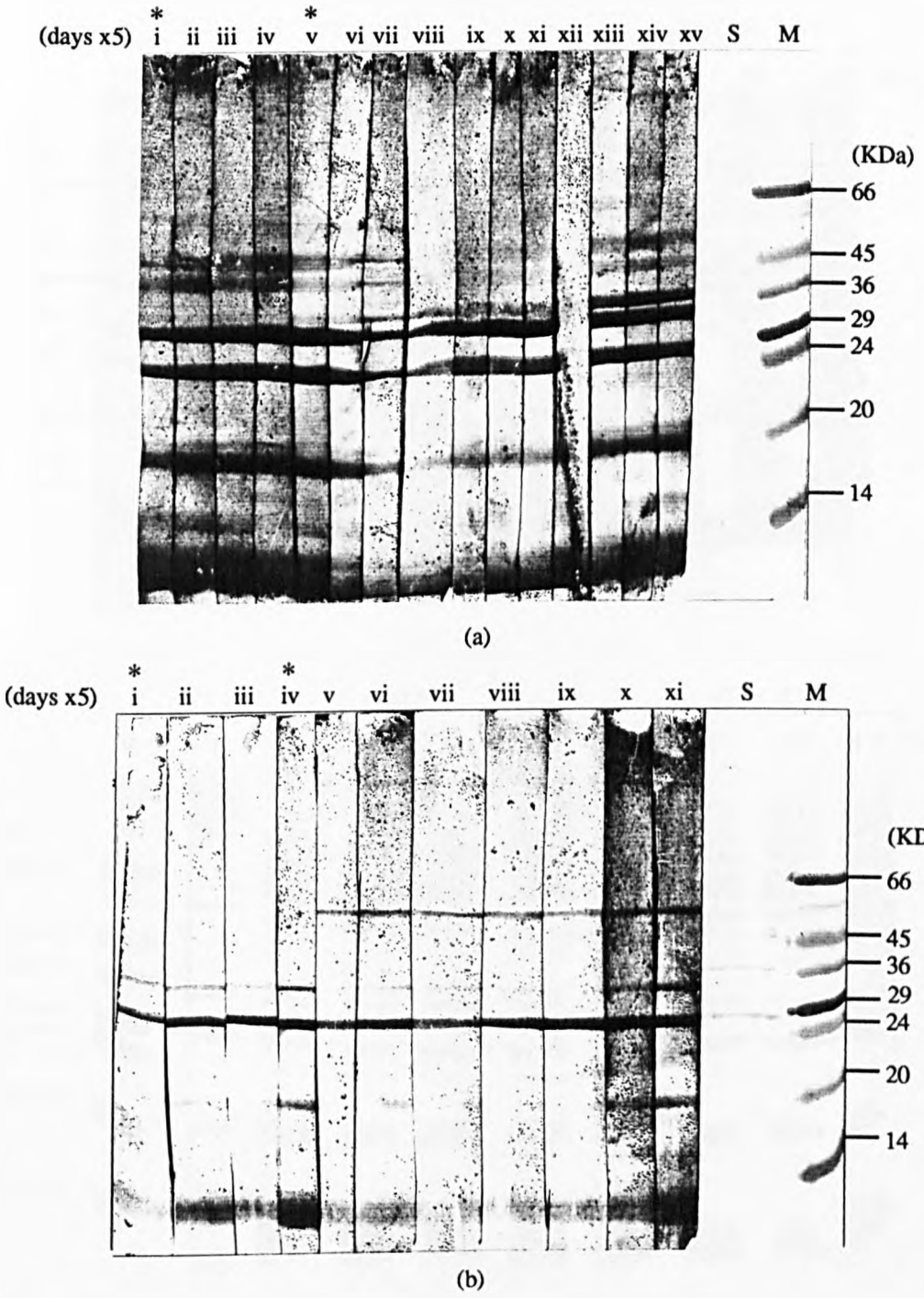


Figure 4.5 Titration of the antibodies raised against the *crtI* and *crtY* gene products. The cell sample of BL21(DE3)pRG1 pT7-CH7 obtained at 120 minutes post induction (Fig. 3.7) was separated by SDS-PAGE and blotted onto PVDF membranes (2.2.21 and 2.2.25). Membrane strips were then used to analyse antibodies raised against (a) CrtY and (b) CrtI. Antibodies were purified from individual eggs and then grouped together in sequential batches of five (Appendix III). * - innoculation of hen. S - Coomassie blue stain of BL21(DE3)pRG1 pT7-CH7 at 120 minutes post induction. M - SDS7 markers (Sigma).

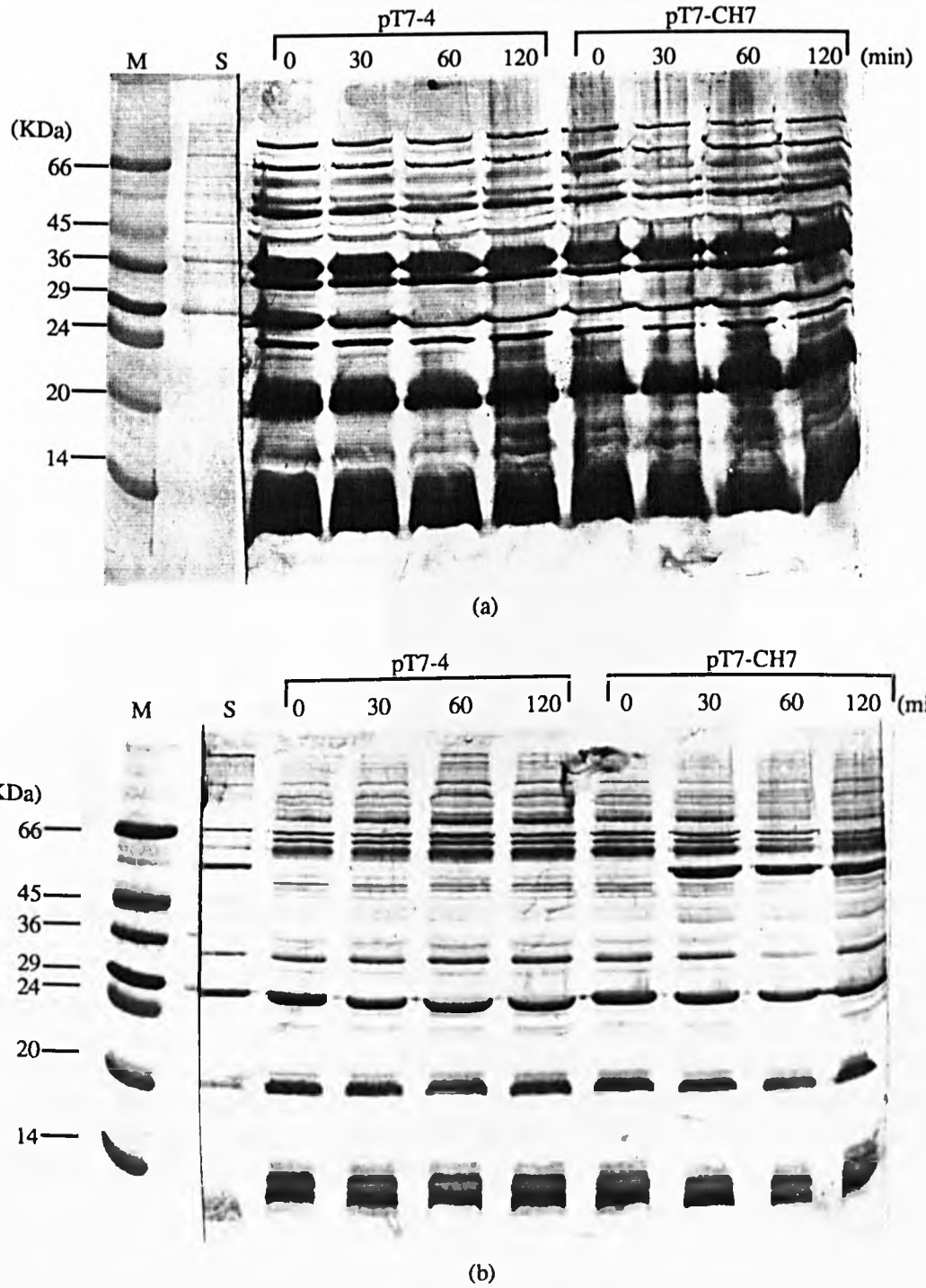


Figure 4.6 *In vivo* inductions of BL21(DE3)pRG1 containing the plasmids pT7-4 and pT7-CH7 analysed by using antibodies raised against (a) CrtY and (b) CrtI. Samples are as described in Figure 3.7. Cell equivalents were separated by SDS-PAGE and blotted onto PVDF membranes (2.2.21 and 2.2.25). S - Coomassie blue stain of BL21(DE3)pRG1 pT7-CH7 at 120 minutes post induction. M - SDS7 markers (Sigma).

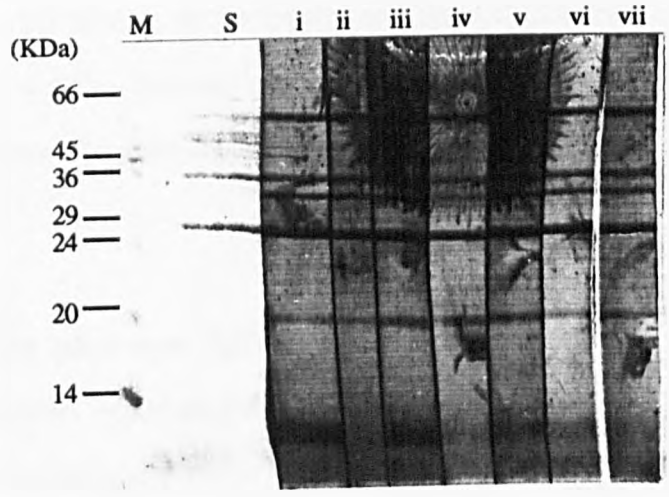


Figure 4.7 Preincubation of the antibody raised against CrtI, with sonicated *E. coli* cells.
The cell sample of BL21(DE3)pRG1 pT7-CH7 obtained at 120 minutes post induction (Fig. 3.7) was separated by SDS-PAGE and blotted onto PVDF membranes (2.2.21 and 2.2.25). Antisera were preincubated with the following amounts (µl) of 1.5ml of sonicated BL21(DE3)pRG1 cells (i) 0, (ii) 10, (iii) 20, (iv) 40, (v) 80, (vi) 160 and (vii) 240 for 1 hour before incubation with the membrane. S - Coomassie blue stain of BL21(DE3)pRG1 pT7-CH7 at 120 minutes post induction. M - SDS7 markers (Sigma).

pre-incubating the sera with sonicated *E. coli* cells as shown in Figure 4.7.

(iii) Western blotting of cell-fractionated recombinant pT7-CH7

The cellular fractions obtained from BL21(DE3)pRG1 pT7-CH7 at 120 minutes post IPTG induction (Figure 4.3) were separated on an SDS-PAGE gel and blotted onto a PVDF membrane as previously. The membrane was incubated with the CrtI antisera raised and the results are shown in Figure 4.8. The 54KDa band appears in the same lanes as those in Figure 4.3 confirming that the 54KDa protein (located mainly to the cytoplasm but also to the periplasm and inner membrane) is CrtI. The plasmid pAPU211 was also fractionated into the different cellular compartments and analysed with the anti-sera and gave the same results (data not shown) again confirming this band as CrtI.

(iv) Localizing phytoene dehydrogenase in *E. herbicola*

Localization experiments of CrtI in its natural host *E. herbicola* were also performed. Total cellular proteins of *E. herbicola* were separated by SDS-PAGE and blotted onto PVDF membranes as previously. A protein band of 54KDa corresponding to CrtI was not observed with the antisera [Figure 4.9 lane (ii)]. Other bands appeared to be dominant but these did not correspond to the size of CrtI when expressed in *E. coli*. It was proposed that the CrtI of *E. herbicola* may only be present in very small quantities and so the antibody may not cross-react at such low levels. To test this theory increasing amounts of total cellular protein were analysed by Western blotting [Figure 4.10 (a)] but no different results were observed. It was also possible that CrtI may be expressed at certain stages of the cell cycle i.e. early, mid or late log phase or even stationary phase. Cell equivalents taken at these times of growth were also analysed by Western blotting [Figure 4.10 (b)] but again no proteins were visible at the expected size of 54KDa.

The predominant 45KDa band of *E. herbicola* with which the antisera cross-reacted was excised from the membrane, and sequenced from the N-terminus, by automated Edman degradation, on an Applied BIOSYSTEMS model 471A Sequenator.

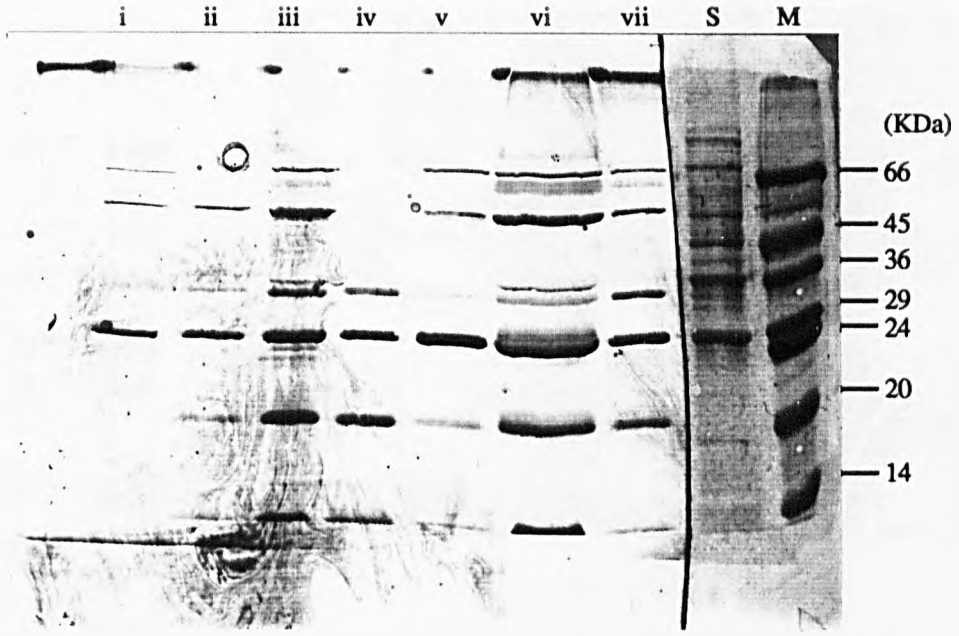


Figure 4.8 Cell fractions of BL21(DE3)pRG1 pT7-CH7.
Samples were obtained as described in Figure 4.3. Fractions are (i) and (ii) periplasm, (iii) total cell minus periplasm, (iv) outer membrane, (v) inner membrane (vi) cytoplasm and periplasm (vii) total cell. S - Coomassie blue stain of BL21(DE3)pRG1 pT7-CH7 at 120 minutes post induction. M - SDS7 markers (Sigma).

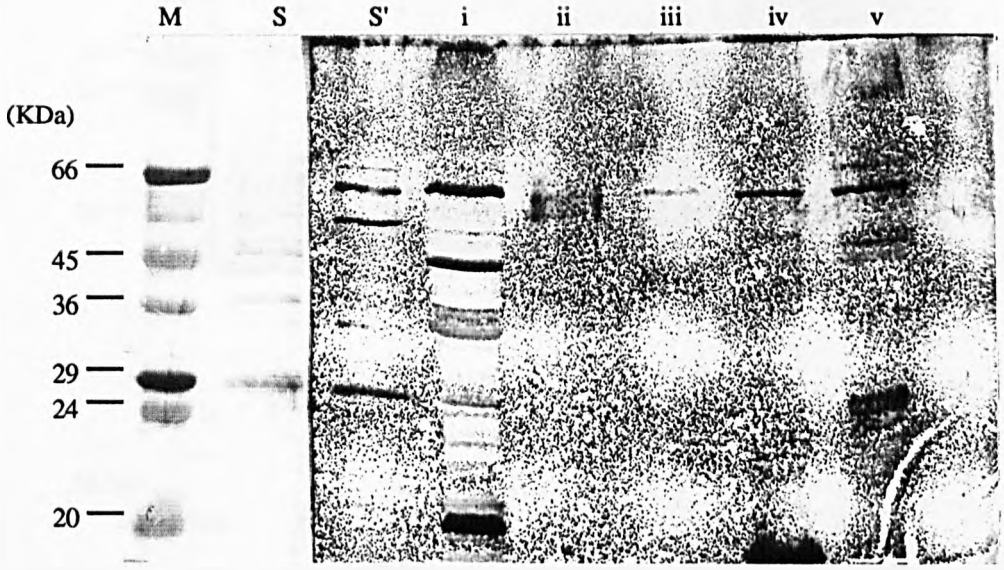
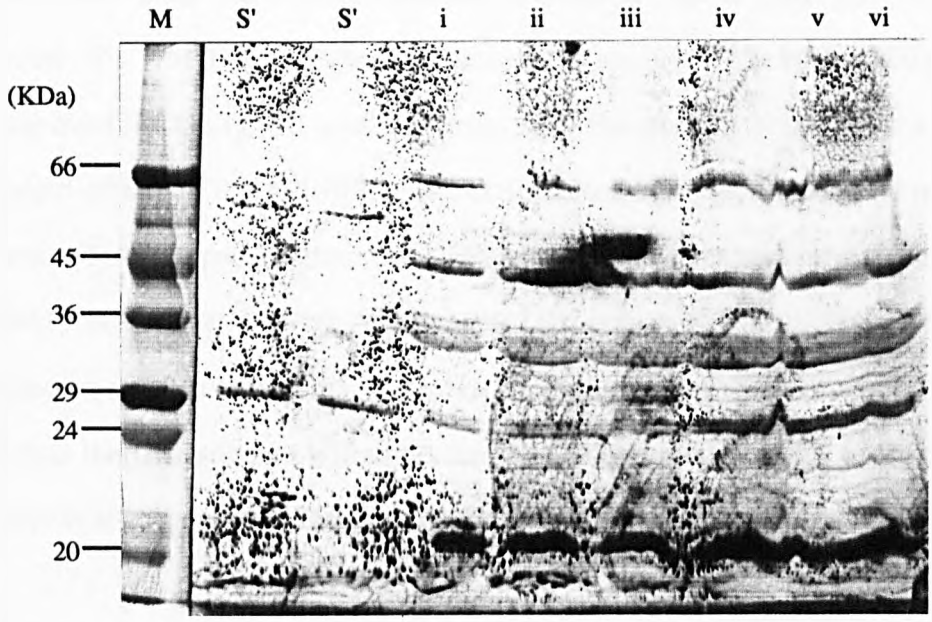
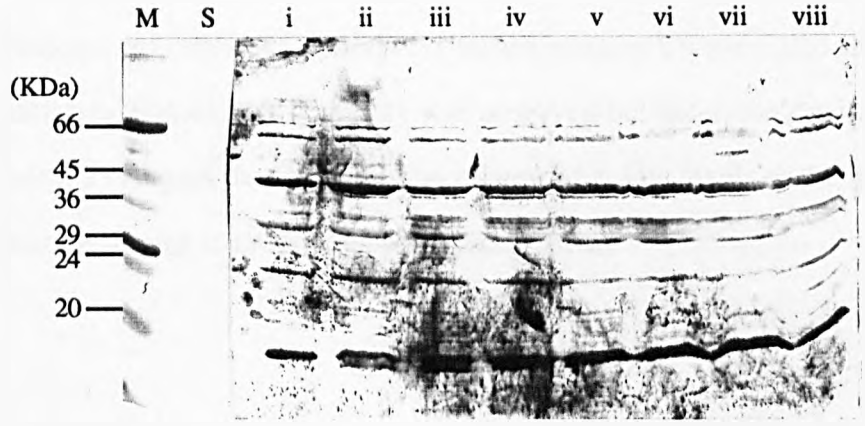


Figure 4.9 Investigation of cross reactions of bacterial CrtI proteins with antibodies raised against *E. herbicola* CrtI.
Cell equivalent proteins of (i) *E. herbicola*, (ii) *B. linens*, (iii) *M. luteus*, (iv) *R. ruber* and (v) *S. multivorum* were separated by SDS-PAGE and blotted onto PVDF membranes and incubated with antisera raised against CrtI (2.2.21 and 2.2.25). S - Coomassie blue stain of BL21(DE3)pRG1 pT7-CH7 at 120 minutes post induction. S' - as S except Western blotted and incubated with antibody. M - SDS7 markers (Sigma).



(a)



(b)

Figure 4.10 Optimisation of Western blotting conditions with antibodies raised against CrtI.
Cell proteins were separated by SDS-PAGE and blotted onto PVDF membranes. Increasing amounts of total cell proteins were loaded as follows in A_{600} optical density (O.D.) units (a) (i) 0.1, (ii) 0.2, (iii) 0.4, (iv) 0.8 and (v) 1. (b) Equivalent total cell proteins were loaded at the following growth stages (measured in A_{600} O.D. units) (i) 0.21, (ii) 0.38, (iii) 0.7, (iv) 0.8, (v) 0.9, (vi) 1.0, (vii) 1.05, (viii) 1.3. S - Coomassie blue stain of BL21(DE3)pRG1 pT7-CH7 at 120 minutes post induction. S' - as S except Western blotted and incubated with antibody. M - SDS7 markers (Sigma).

The sample was crude and so difficulties arose in obtaining sequence due to high background. The tentative sequence, M?P/EAVILSLK was obtained. No matches to this sequence were made from database searching. An *E. coli* 2D gel map was analysed for 45KDa proteins (Swiss-2D page, http://expasy.hcuge.ch/cgi-bin/map2/def?ECOL1). It was proposed that the protein could be the *Erwinia* equivalent of one of the following, isocitrate dehydrogenase, S-adensoyl methionine synthetase, NADP-specific glutamate dehydrogenase. The latter enzyme certainly is a possibility as the antisera were raised against the region of CrtI characteristic of βαβ folds known to bind cofactors. It may be that this protein is expressed in such high levels that the CrtI antisera will cross react. It is possible that CrtI expression in *E. herbicola* is at such low levels that detection with this crude antiserum is not possible.

(v) Screening other bacteria for phytoene dehydrogenase

The total cellular protein equivalents of *Brevibacterium linens*, *Micrococcus luteus*, *Rhodococcus ruber* and *Sphingobacterium multivorum* were also analysed with the CrtI antisera. Some cross-reactivity was observed but not at the predicted size for CrtI (Figure 4.9). Again the CrtI could be expressed at low levels or the antisera may not be sensitive enough to cross-react with other species.

4.6 Discussion

Semi *in vivo* gene expression systems have been used previously to identify and localize plasmid encoded gene products (Stoker *et al.*, 1984; Meagher *et al.*, 1977). In the absence of specific antisera or in cases where overproduction of gene products is either impossible to achieve, due to toxicity or translational control mechanisms, or undesirable, such that genuine localization, perhaps to limited membrane binding sites, is not apparent, gene expression systems can be very useful.

Two different approaches have been used to try to localize the *crt* gene products. Maxi cells carrying the *crt* genes are labelled then fractionated or advantage is taken of the T7 RNA polymerase dependent expression in the presence of rifampicin, subsequent labelling with ^{35}S-methionine, revealing only newly transcribed/translated products. In neither case is overproduction needed to obtain results as products present in low levels are highly radiochemically labelled and cell fractionation can identify their cellular location.

Although the maxi-cell system is thought to be a good expression system in which to study protein expression from plasmids, this did not prove to be the case with the plasmid pAPU211. The levels of protein needed for carotenoid biosynthesis are not known. It may be that, even though carotenoid synthesis is high, only small amounts of protein are needed to achieve this. Alternatively, proteins could be relatively unstable or toxic to the cell at high levels. If either postulate is true, then difficulties will arise in determining cellular location as low levels of protein are very difficult to observe.

As Crt protein expression was weak in maxi cells, and antisera to the Crt enzymes were not available, a radioactive labelling approach was employed in localization experiments. This involves reaching a balance between making enough protein to enable visualization but not so much that available membrane sites are swamped. Expression of the proteins was increased such that they were visible on SDS-PAGE gels with the plasmid construct pT7-CH7 as shown in Figure 3.7 (b). Cell fractionation showed some membrane association of the proteins CrtI, CrtB, CrtY, and CrtE, however the majority of the protein appeared to be located in the cytoplasm. If

the proteins are truly located to the membrane, overexpression could lead to the complete saturation of available sites within the membrane, resulting in the excess proteins being deposited in the cytoplasm, however the distribution observed for the Crt enzymes would also be consistent with cytoplasmic location of the complex with some cross-contamination of fractions. The *E. coli* cells carrying the Crt cluster were also fractionated by passing through a French Press (data not shown) and the same results were observed.

Antibodies used for protein localization negate the need for altering protein expression. Antibodies raised against CrtI of *Synechococcus* PCC7942 have been successfully used in *in situ* localization experiments with tobacco and spinach (Linden *et al.*, 1993). The enzyme was located in the thylakoid membranes of the chloroplasts. The same techniques have been employed for localization, of CrtI to the photosynthetic membrane, in *R. capsulatus* SB1003, *Synechocystis* PCC 6714 and *Anabaena variabilis* ATCC 29413 (Serrano *et al.*, 1990). Antibodies raised against GGDP synthase of *Capsicum* fruits have located the enzyme to distinct regions within the stroma, the plastid membranes and plastoglobuli of chloroplasts (Cheniclet *et al.*, 1992). In all these experiments antibodies were raised in rabbits against purified proteins.

The use of chickens as an alternative source of polyclonal antibodies has been described by Polsen *et al.* (1980). Chicken egg yolks have proved a rich and inexpensive source of polyclonal antibodies raised against mammalian proteins (Gassmann *et al.*, 1990). The procedure was chosen because it was a simple, cheap and non-invasive method, that held numerous advantages over more conventional polyclonal antibody production. An unlimited supply of antibody could be obtained as eggs are laid daily, the amount of specific antibody from one egg is the same as that from one bleed of rabbit and the process is a non-terminal procedure were just a few of the advantages considered when this protocol was decided upon.

Approximately one month after inoculation a response was seen with antisera raised against CrtI but not with antisera raised against CrtY. The epitope chosen for CrtY may not have not been sufficiently antigenic or exposed enough to elicit a

response. The epitope could have different characteristics when part of the whole protein or on SDS-PAGE gels, compared to a chemically synthesised polypeptide. However these parameters did not apply to CrtI, for which a distinct response was seen. This is supported by observing a cross-reaction of protein of appropriate size with the antisera after IPTG induction of the plasmid pT7-CH7. Cell fractionation experiments were performed on this construct and pAPU211. The results were the same for both and supported those obtained in the cell fractionation experiments (Section 4.4). This also implies that cell fractionation of pT7-CH7 to localize the proteins was acceptable as the same results were observed as for the cell fractionation of the plasmid pAPU211 where much lower levels of protein were apparent. This implies that location of the Crt enzymes in *E. coli* within the cytoplasm was genuine and not due to the saturation of available membrane sites, although the question of some membrane association or cross-contamination remains unanswered. Overexpression of CrtI from *R. sphaeroides* in *E. coli* resulted in the majority of the enzyme activity being associated with the membrane fraction despite the abundance of CrtI in the cytoplasm (Lang *et al.*, 1994). Perhaps this observation could be applicable to the overexpression of CrtI from *E. herbicola*. This could account for some membrane association. A recent paper by Gorham *et al.*, (1996) suggests that carotenoid biosynthesis occurs in the cytoplasm in *M. xanthus* and that the regulating anti-sigma factor CarR is membrane associated. Clearly the location, expression and regulation of Crt enzymes is different in different organisms, so no general rule of location can be applied.

Attempts were made to discover any enzyme complex formation with CrtI. In initial gene assignment of the *crt* genes of *E. herbicola*, each *crt* gene was interrupted by a specific mutation (Hundle *et al.*, 1994). The resultant plasmids containing mutations in one gene of the entire *crt* cluster were received as a gift from Dr Hundle. Cell fractionation experiments were performed on these constructs to discover if the presence or absence of another protein of the pathway altered the location of CrtI, as a crude attempt to discover any complex formation. CrtI was always located in the same place as previously, regardless of whether another protein was present or absent (data

not shown). This does not rule out a transient complex formation occurring upon carotenoid biosynthesis.

Attempts to observe CrtI with the antisera in *E. herbicola*, and the other bacteria were negative. The lack of any positive result with the four bacteria could be because the epitope used for antisera production was not as well conserved in these bacteria, and so no cross reactivity would occur. This is thought unlikely, as the epitope chosen encompassed some of the βαβ fold, characteristic of the bacterial CrtI enzymes. Antisera raised against *R. capsulatus* CrtI and *Synechococcus* sp. CrtI and CrtB have shown cross-reactivity and have been used to locate the corresponding enzymes in plants (Linden *et al.*, 1993; Schmidt *et al.*, 1989; Fraser *et al.*, 1994), however these antibodies were raised against the whole protein, not an epitope, which could be fundamental for cross-reactivity between species. It may be, as is thought with *E. herbicola* that the protein is expressed at very low levels in these bacteria.

Cross reactivity was observed with CrtI antisera. Chicken immunoglobulins have previously been raised against mammalian proteins. Non-specific binding is reduced as chicken immunoglobulins do not cross-react with mammalian immunoglobulins or Fc receptors. Freunds complete adjuvant consists of a mineral oil base containing dead mycobacteria, and is used in immunization to ensure a general stimulation of the immune system. The cross-reactivity seen with CrtI antisera is possibly due to antisera produced against the mycobacteria reacting with equivalent *E. coli* proteins. If this technique is used in the future mineral oil alone could be used instead as the adjuvant. To eliminate cross-reactivity, the antibody could be purified by affinity columns, but resources were not available for this procedure. Commercial kits are now availble for purifying chicken immunoglobulins (Promega). The signal to noise ratio was increased by pre-incubating the antisera with sonicated *E. coli* extracts before incubation with the membrane.

Antibody production was successful for CrtI. Location of the Crt enzymes is very difficult in a heterologous host as one has to assume that membrane association sites, if needed, are present. One has to assume that conditions are the same because

the carotenoid pathway is actively functioning. Western blotting was not sensitive enough to locate the protein in its natural host, probably due to extremely low levels of protein synthesis. The question of true location in *E. herbicola* remains unanswered and will until a more sensitive technique is applied, or a better antibody produced, to allow *in situ* hybridizations. Only when different experimental methods such as *in situ* hybridization, lead to the same conclusion, can the location of a protein be established (Randall *et al.*, 1987).

CHAPTER 5

CHARACTERIZATION OF UNUSUAL CAROTENOIDS AND THE IDENTIFICATION OF THEIR BIOSYNTHETIC GENES

5.1 General introduction

Carotenoid biosynthesis is important from both commercial and scientific viewpoints. As discussed earlier, many bacterial carotenoid pigments have been isolated and analysed but only in a few bacteria have the carotenoid biosynthetic pathways been genetically characterized. Thus there is great scope for the exploitation of carotenogenesis in these organisms.

The bacteria studied in this work and the carotenoids they are believed to produce are shown in Figures 5.1 and 5.2 respectively.

Rhodococcus ruber is reported to produce canthaxanthin as its main pigment. Canthaxanthin is used to colour various foodstuffs and is supplied along with astaxanthin in the feed of salmon and trout to produce the pink colour of the flesh. If the genetics of carotenoid biosynthesis could be manipulated such that canthaxanthin [or astaxanthin (Appendix I)] could be mass produced it would prove a cost effective alternative to production by chemical synthesis or the use of large algal lakes.

Zeaxanthin is synthesized by *E. herbicola*; many aspects of the production have been patented, however *Sphingobacterium multivorum* (formely *Flavobacterium*), however, also synthesises zeaxanthin and is open to commercial exploitation.

Brevibacterium linens synthesizes the phenolic carotenoid 3,3'-dihydroxyisorenieratene. The pathways leading to its synthesis have not been fully analysed, and the formation of the aromatic ring is not understood. Since *B. linens* is already used in the food industry it is more amenable for use in a commercial production process.

Micrococcus luteus (formerly *Sarcina lutea*) synthesizes a C_{50} carotenoid, sarcinaxanthin, and also glucosides of this compound. The reactions involved in the

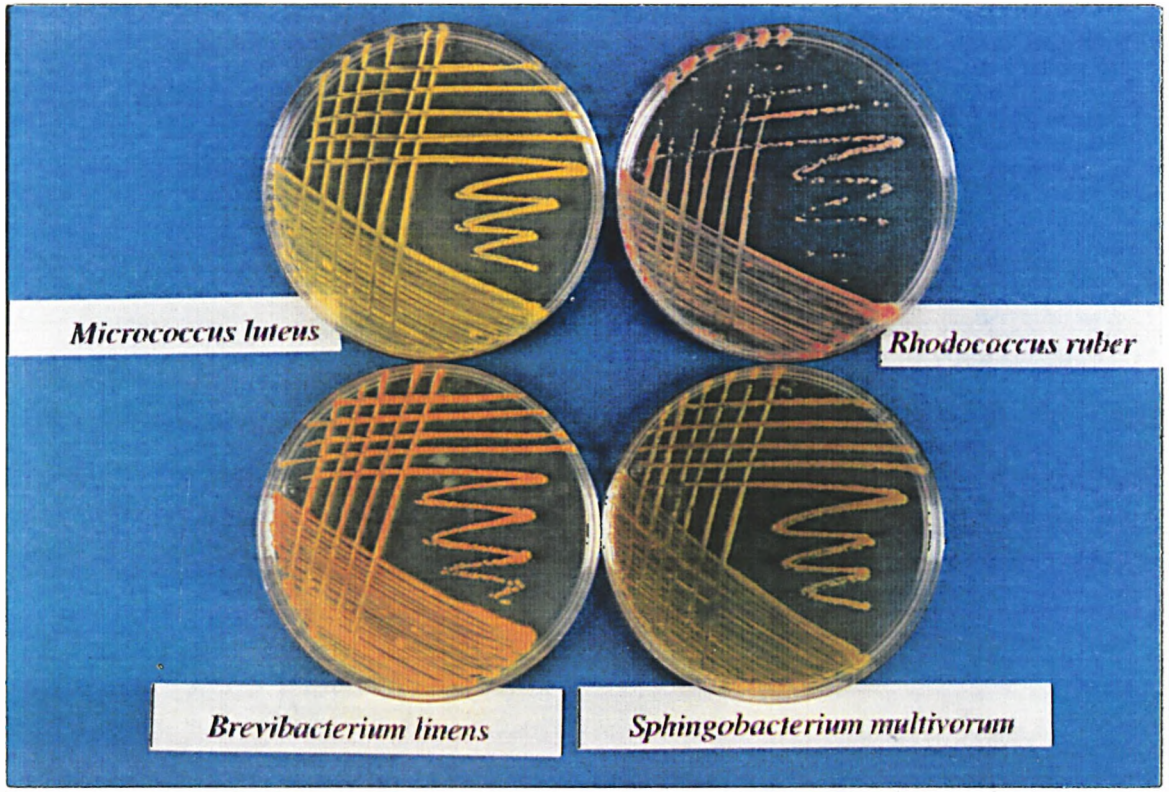
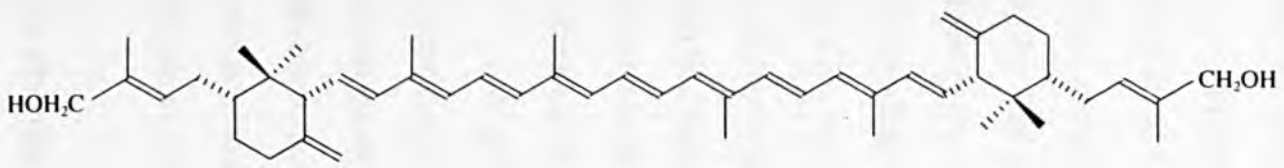
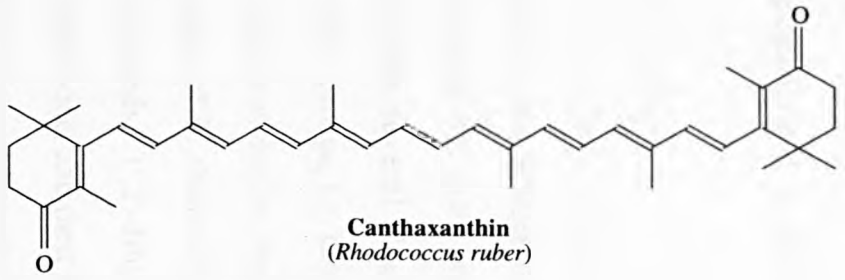


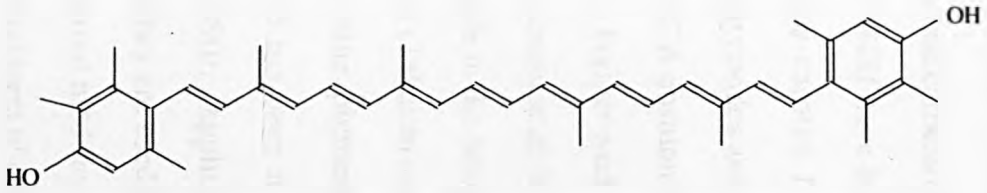
Figure 5.1 Colonies of the carotenoid-synthesising bacteria *B. linens*, *M. luteus*, *R. ruber* and *S. multivorum* growing on nutrient agar plates (2.1.2 (ii)).



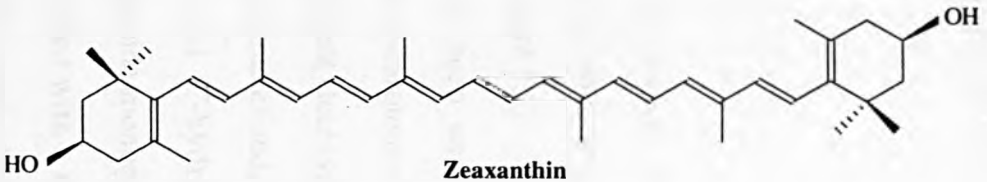
Sarcinaxanthin
(*Micrococcus luteus*)



Canthaxanthin
(*Rhodococcus ruber*)



3, 3' - Dihydroxyisorenieratene
(*Brevibacterium linens*)



Zeaxanthin
(*Sphingobacterium multivorum*)

Figure 5.2 Main carotenoids synthesized by the bacteria *B. linens*, *M. luteus*, *R. ruber* and *S. multivorum*.

formation of sarcinaxanthin have not been studied in any detail and are of interest because no enzymes or genes for C_{50} carotenoid biosynthesis have been isolated.

5.2 Bacterial strains used in this work

(i) *Rhodococcus ruber*

R. ruber is a Gram positive, non-photosynthetic bacterium that forms ringed/filamentous colonies with an irregular margin when grown on agar plates. The colonies are red in colour due to the presence of carotenoids (Figure 5.1).

Pigments from several *Rhodococcus* species have been analysed and categorized into three groups, i.e. those containing β-carotene, γ-carotene or neither. These analyses were inconclusive about the total carotenoid content (Ichhiyama *et al.*, 1989). The pigments of *R. rhodochrous* RNMSI have been characterised. The two monocyclic carotenoids, 1',2'-dihydro-β,ψ-caroten-1'-ol and 1'-hydroxy-1',2'-dihydro-β,ψ-caroten-4-one, and their monoglycosides and glycoside monoesters were present. The glucoside moities were at the C-6 position esterified with a mixture of nine major fatty acids (Takaichi *et al.*, 1990). Further work on this strain showed that a decrease in growth temperature resulted in an increase in the total carotenoid content, an increase in the content of unsaturated fatty acids in the carotenoid glucoside esters and a decrease in the content of saturated fatty acids (Takaichi and Isshidsu, 1993).

A gene responsible for pink and blue pigmentation, but not carotenoid biosynthesis, in *Rhodococcus* ATCC 21145 has been studied to enable its use as a chromogenic cloning vector (Hart *et al.*, 1990). Applications of various species of *Rhodococcus* are patented because in Japan they are used for the production of acrylic acid/acrylamide and they are also used as steroid modifiers (Finnerty, 1992). None of the patents involves their use as large scale producers of carotenoids.

(ii) *Brevibacterium linens*

B. linens is a non-motile, Gram-positive bacterium of the family Corynebacterium. Carotenogenesis-stimulating chemicals have been added to

Brevibacterium sp. strain KY-4313 growth medium to optimize canthaxanthin production (Nelis and De Leenheer, 1989). No actual chemical stimulation was discovered, but periodical renewal of the media enhanced canthaxanthin production. Alcohols and retinols also stimulated carotenogenesis but this was not conducive to improved growth. The orange colour of *B. linens* is due to the presence of the carotenoid 3,3'-dihydroxyisorenieratene (Figure 5.1).

(iii) *Sphingobacterium multivorum*

S. multivorum (formerly *Flavobacterium*) is a Gram-negative bacterium isolated from seawater and fresh water and food. The main carotenoid biosynthesized, zeaxanthin, has the same chirality (3*R*, 3'*R*) as that produced in plants.

Zeaxanthin is the major pigment of *Flavobacterium aquatile* and both zeaxanthin and β-carotene are produced in other *Flavobacterium* sp. Experiments on *Flavobacterium* R1519 showed that nicotine blocks zeaxanthin formation by inhibiting lycopene cyclization. At low nicotine concentrations the monocyclic rubixanthin was formed and converted into β-cryptoxanthin and zeaxanthin showing that alternative pathways exist (McDermott *et al.*, 1974). The experiments also showed that cyclization precedes hydroxylation. Brown *et. al.*, analysed cell-free preparations of *Flavobacterium* R1560 and showed that phytoene formation was enhanced by co-factors (Brown *et al.*, 1975).

(iv) *Micrococcus luteus*

The genera *Micrococcus* are Gram-positive bacteria forming opaque colonies that are either bright yellow or pink, when grown on agar. The type species is *M. luteus* which has formerly been known as *Sarcina lutea.*

As early as 1936 the pigments of *M. luteus* were known to be carotenoids (Nakamura, 1936), the main components being considered to be xanthophyll-like esters. Pigments have been extracted from *M. luteus*, *S. flava* and *S. radiodurans* and were identified as carotenols, ketocarotenoids and carotenoids complexed with proteins and glucose (Thirkell, 1969; Thirkell and Hunter, 1969; Thirkell and Strang, 1967).

The main pigment of *M. luteus* is the C_{50} carotenoid sarcinaxanthin and glucoside derivatives. The main pigment of *M. roseus* was shown to stabilize synthetic membranes and was identified as bisdehydro-β-carotene-2-carboxylic acid (Jagannadham *et al.*, 1991).

The carotenoids of *M. luteus* protect the cell against lethal photosensitization by toluidine blue where the membrane proteins are a target. Protection is not offered against 8-methoxypsoralen where cell DNA is damaged (Mathews, 1963). Studies have been performed on the effect of polyene chain length in offering protection against photosensitization, and showed that 8 conjugated double bonds (cdb) offered no protection (Mathews-Roth *et al.*, 1974) when compared to the wild-type pigment of *M. luteus* and β-carotene. Mutants have been generated that differ in the number of cdb that their carotnoids contain. Carotenoids with 9 cdb were effective in quenching singlet oxygen, (as much as β-carotene), those with fewer cdb being less effective (Mathews-Roth and Krinsky, 1970a). When compared to wild type *M. luteus* it was proposed that the important factor in protection is the total amount of carotenoid per cell and not the amount of various individual pigments (Mathews-Roth and Krinsky, 1970b). In a study on singlet oxygen lethality to different bacteria, *M. luteus* strains containing carotenoids were more resistant than their carotenoid free counterparts and other Gram-positive bacteria lacking carotenoids (Dahl *et al.*, 1989).

Many enzymes involved in the synthesis of prenyldiphosphates in *M. luteus* have been isolated and characterized. Prenyldiphosphates are precursors of carotenoids and are the main components of the side chain of quinones. The enzymes solanesyl diphosphate synthase and GGDP synthase have been isolated from *Micrococcus lysodeikticus* (Kandutsch, 1969; Sagami *et al.*, 1977). In 1980 three distinct prenyl transferases were isolated that are responsible for the generation of compounds varying in carbon chain length from C_{20} to C_{60} (Baba and Allen, 1980). The regulation of isoprenoid synthesis in *M. luteus* whole cells was rate limited by the enzymes IDP synthase and FDP synthase (Takatsuji *et al.*, 1983). More recently the level of IDP-Mg has been shown to be decisive in determining the chain lengths produced by solanesyl diphosphate (Ohnuma *et al.*, 1992). The enzyme has been purified and shown to

produce *all-E*-solanesyl and octaprenyl diphosphates and was thought to have a cofactor involvement (Ohnuma *et al.*, 1991).

This chapter describes the isolation and identification of the carotenoid pigments produced by the above bacteria. Preliminary investigations into the carotenoid biosynthesis genes have also been performed by the technique of Southern *et al* (1975) with probes derived from the carotenoid biosynthesis genes of *E. herbicola*.

5.3 Analysis of pigments extracted from the bacteria

(i) Extraction of carotenoids

Bacterial growth medium was inoculated with one colony of the appropriate bacterial stocks and growth was continued in light conditions (2000lux) until the cells reached stationary phase. Carotenoids were extracted from each of the four bacterial cultures as described in Section 2.2.26 (i). The pigments were then analysed by reversed-phase HPLC [2.2.26 (ii)].

(ii) Reversed-phase HPLC

The solvent system chosen to separate the carotenoids is described in Section 2.2.26 (ii). Chromatograms illustrating the separation of carotenoids from each bacterium are given in Figure 5.3 (a-d). Where possible, identification of peaks is given in the Figure legend.

The HPLC analysis of the main carotenoids of *M. luteus* identified three peaks with identical uv-vis spectra and close retention times. The identical uv-vis spectra indicate that the carotenoids have the same chromophore. The spectra showed no *cis*-peaks, indicating that the isomers are all likely to be (all *E*)-isomers [Figure 5.3 (b)].

HPLC analysis of the carotenoids of *B. linens* showed five peaks that could be putatively identified on the basis of their uv-vis spectra [Figure 5.3 (a). The major component of *B. linens* was dihydroxyisorenieratene. β-Carotene was also present, together with a *cis*-isomer.

The HPLC analysis of the carotenoids of *R. ruber* revealed a very complex mix of carotenoids, some not commonly found in bacteria. The composition was far more complex than simple ketocarotenoid production. No positive identification of canthaxanthin could be made, although compounds with similar spectra were present, the retention times differed for that of canthaxanthin. A major component however, could be identified as dihydroxyisorenieratene which has previously only been identified in *Brevibacterium* sp. [Figure 5.3 (c)].

Figure 5.3 Separation of carotenoids from the bacteria *B.linens*, *M. luteus*, *R. ruber* and *S. multivorum* by reversed-phase HPLC. The solvent system used is described in 2.2.26 (ii). Absorption spectra for each peak indicated were determined by a photo-diode array detector [2.2.26 (ii)].

(a) *B.linens*

Peak identifications;

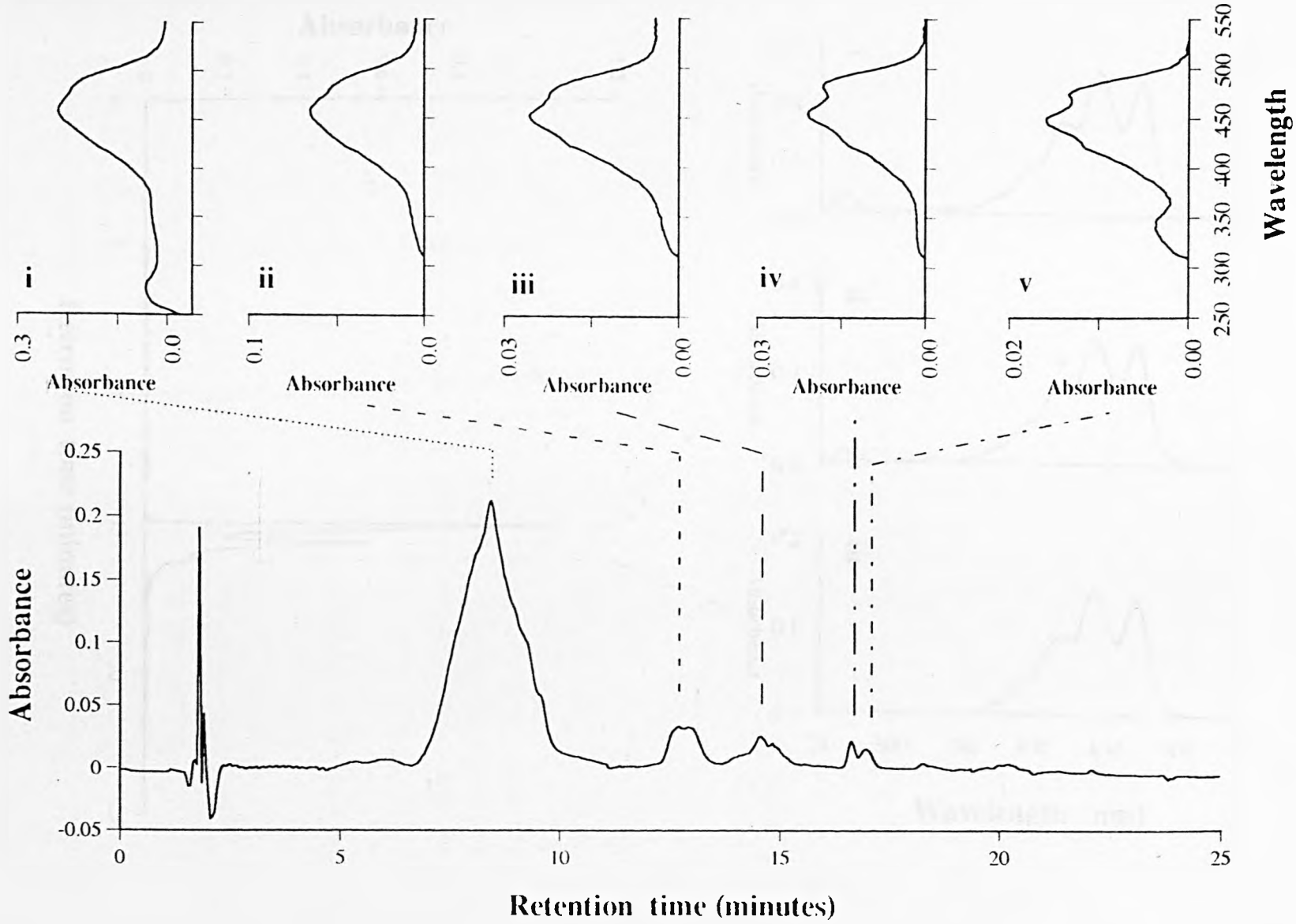
(i) dihydroxyisorenieratene, λ_{max} nm; 425, 455, 488

(ii) monohydroxyisorenieratene, λ_{max} nm; 425, 455, 480

(iii) monohydroxy-β-isorenieratene, λ_{max} nm; 425, 450, 475

(iv) β-carotene, λ_{max} nm; 425, 450, 479

(v) cis-β-carotene, λ_{max} nm; 425, 450, 475


i
ii
iii
iv
v
0.3
0.0
0.1
0.0
0.03
0.00
0.03
0.00
0.02
0.00
Absorbance
Absorbance
Absorbance
Absorbance
Absorbance
250 300 350 400 450 500 550
Wavelength
0.25
0.2
0.15
0.1
0.05
0
-0.05
Absorbance
0
5
10
15
20
25
Retention time (minutes)

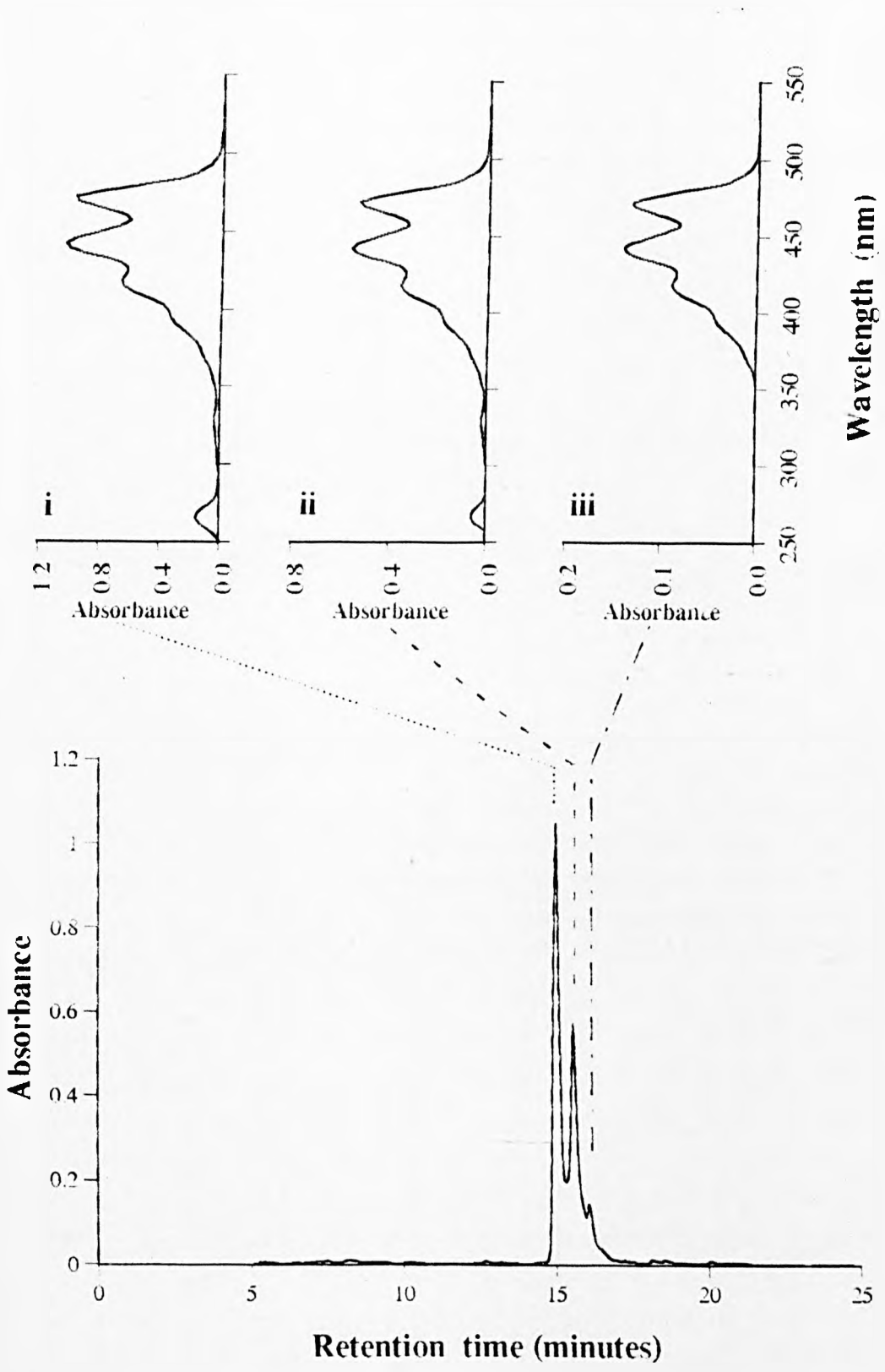
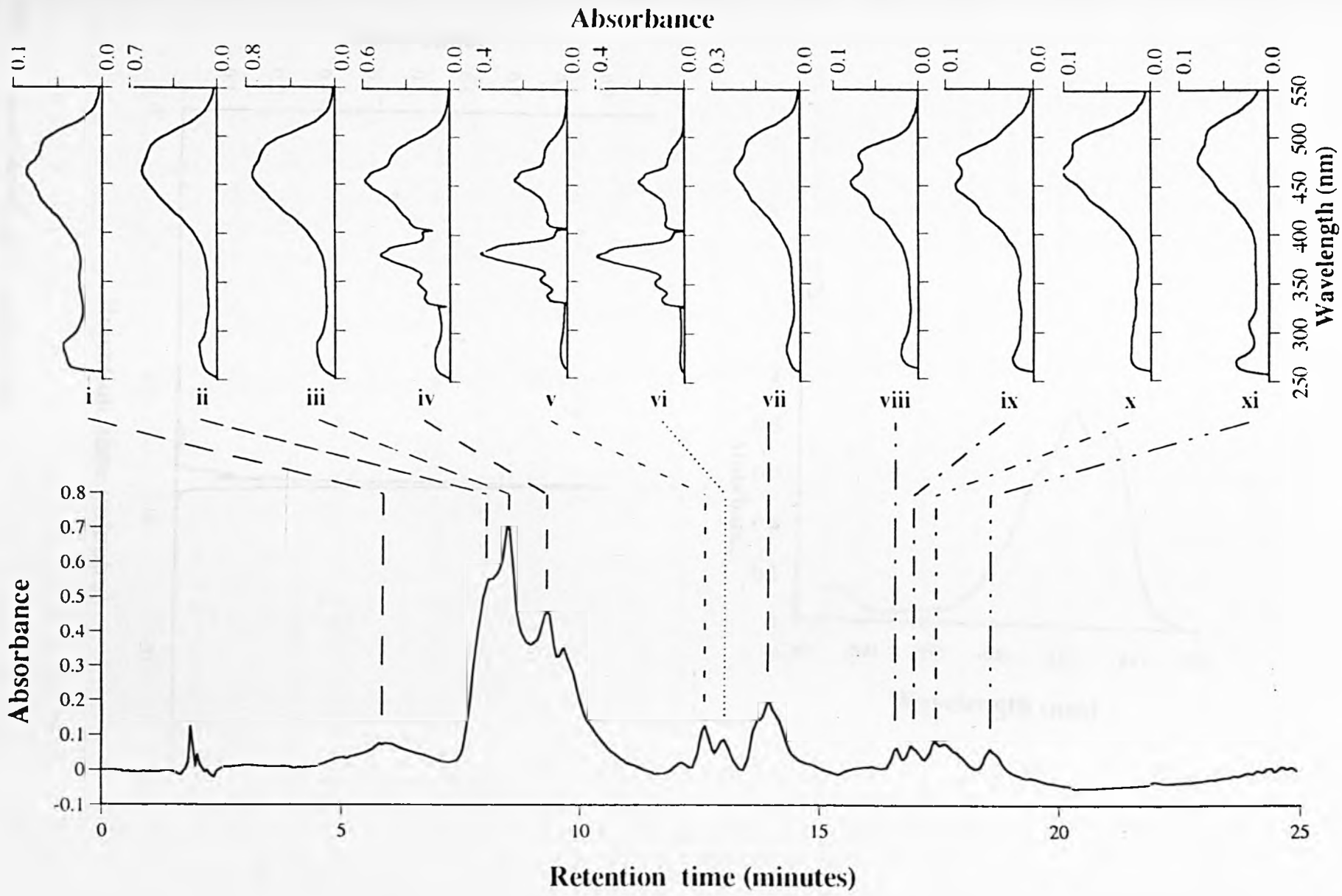


Figure 5.3 (continued), (b) *M. luteus*

Peak identifications, (i), (ii) and (iii) isomers of sarcinaxanthin λ_{max} nm; 420, 440, 471

Figure 5.3 (continued)
(c)R. ruber
Peak identifications;
(i) uncharacterized λ_{max} nm; 435, 460, 495
(ii) dihydroxyisorenieratene, λ_{max} nm; 425, 455, 485
(iii) monohydroxyisorenieratene, λ_{max} nm; 425, 455, 485
(iv)uncharacterized λ_{max} nm; 420, 455, 480
(v) uncharacterized λ_{max} nm; 425, 455, 480
(vi) uncharacterized λ_{max} nm; 425, 455, 480
(vii) uncharacterized λ_{max} nm; 435, 465, 495
(viii) uncharacterized λ_{max} nm; 420, 450, 480
(ix) uncharacterized λ_{max} nm; 420, 450, 475
(x) uncharacterized λ_{max} nm; 435, 470, 490
(xi) uncharacterized λ_{max} nm; 444, 474, 505


Absorbance
0.1
0.0
0.7
0.0
0.8
0.0
0.6
0.0
0.4
0.0
0.4
0.0
0.3
0.0
0.1
0.0
0.1
0.0
0.1
0.0
0.1
0.0
250 300 350 400 450 500 550
Wavelength (nm)
i
ii
iii
iv
v
vi
vii
viii
ix
x
xi
Absorbance
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
0
5
10
15
20
25
Retention time (minutes)

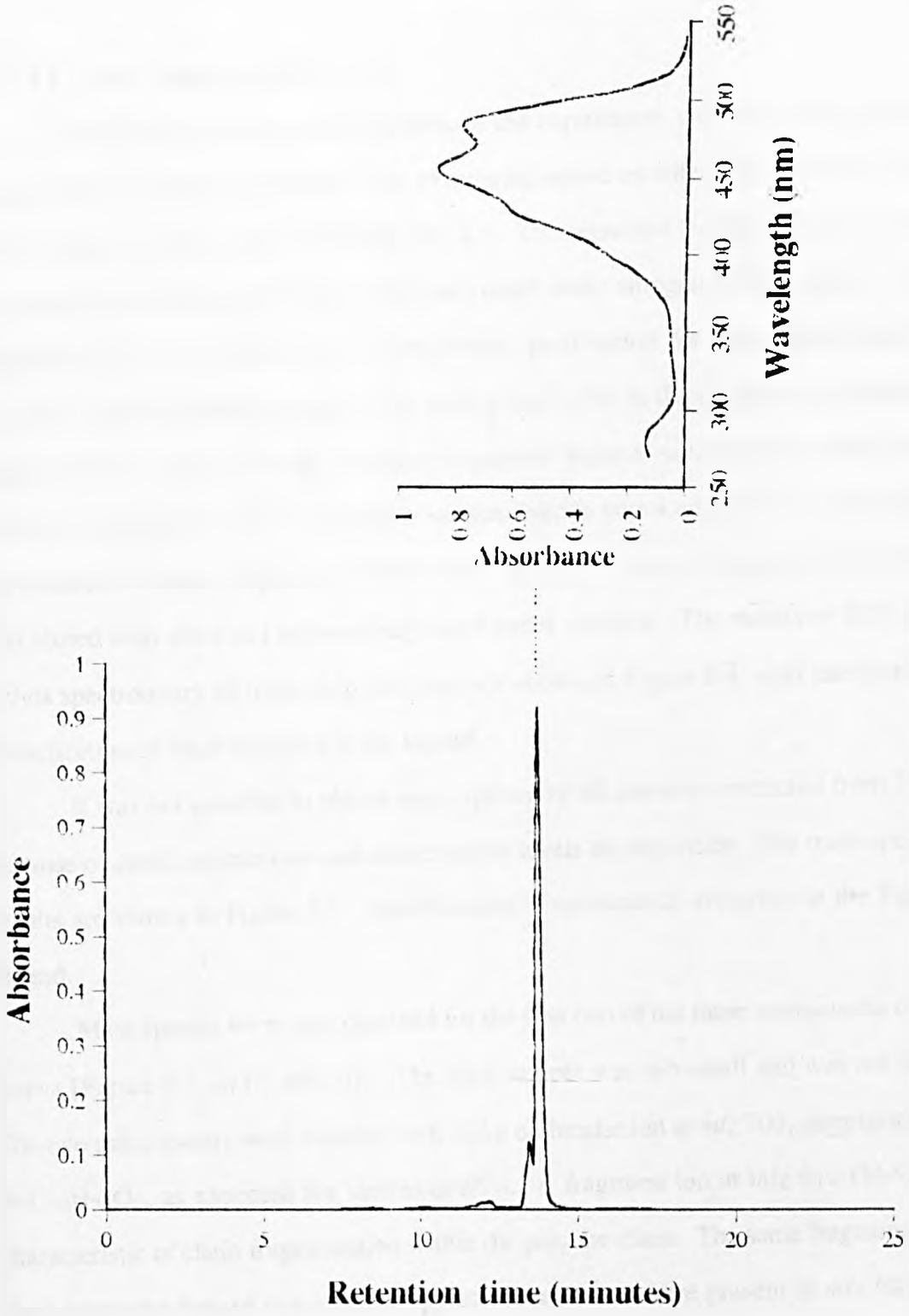


Figure 5.3 (continued), (d) ***S. multivorum***

Peak identification; zeaxanthin, λ_{max} nm; 425, 450, 480

The major pigment of *S. multivorum* was identified from the HPLC analysis as zeaxanthin. No other carotenoids were observed [Figure 5.3 (d)].

(iii) TLC and mass spectrometry

Further information on the identity of the carotenoids was obtained by mass spectrometry. Pigments extracted were chromatographed on silica TLC plates [2.2.26 (iii)], with a solvent system of ether:petrol, 1:1. The separated fractions were scraped off, dissolved in dichloromethane (DCM) and dried under nitrogen. The fractions were analysed by uv/vis spectrometry before further purification for mass spectrometry. Alumina was deactivated to grade III by adding H_2O (6%) to the alumina under petrol. Small columns made from the alumina in pasteur pipettes were used to purify the fractions obtained by TLC. Fractions were dissolved in petrol:ether 3:1 and applied to the alumina column. Impurities were washed through with petrol and the carotenoid was eluted with ether and subsequently dried under nitrogen. The results of TLC and uv/vis spectrometry of individual fractions are shown in Figure 5.4, with preliminary identification of fractions given in the legend.

It was not possible to obtain mass spectra of all fractions extracted from TLC because of small sample size and unacceptable levels of impurities. The mass spectra results are shown in Figure 5.5. Identification of carotenoids are given in the Figure legend.

Mass spectra were only obtained for the first two of the three components of *M. luteus* [Figure 5.5 (b) (i) and (ii)]. The third sample was too small and was not pure. The two mass spectra were similar; both had a molecular ion at *m/z* 704, corresponding to $C_{50}H_{72}O_2$, as expected for sarcinaxanthin. A fragment ion at *m/z* 612 (M-92) is characteristic of chain fragmentation within the polyene chain. The same fragmentation ions due to the loss of one and two molecules of water, were present at *m/z* 686 and 668 but with different relative intensities for the two compounds. Without detailed nuclear magnetic resonance (nmr) study the compounds cannot be fully characterized.

The mass spectra of the *B. linens* carotenoids show that the main pigment has a molecular ion at *m/z* 560, corresponding to $C_{40}H_{48}O_2$ [Figure 5.5 (b)]. A major

Figure 5.4 Separation of bacterial carotenoids by TLC and analysis by uv/vis absorption spectra. Carotenoids were extracted as in 2.2.26 (i) and analysed by TLC [2.2.26 (iii)] with a solvent system of ether:petrol, 1:1 and by uv-vis absorption spectrophotometry.

(a) ***B. linens***

Fraction identifications;

(i) cis-β-carotene, λ_{max} nm; 425, 450, 475

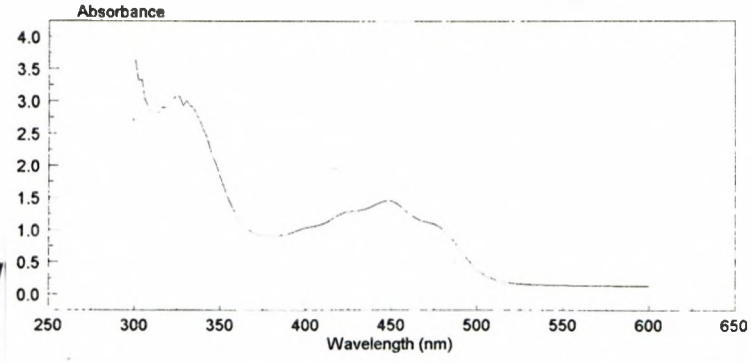
(ii) β-carotene, λ_{max} nm; 425, 450, 475

(iii) dihydroxyisorenieratene, λ_{max} nm; 425, 455, 485

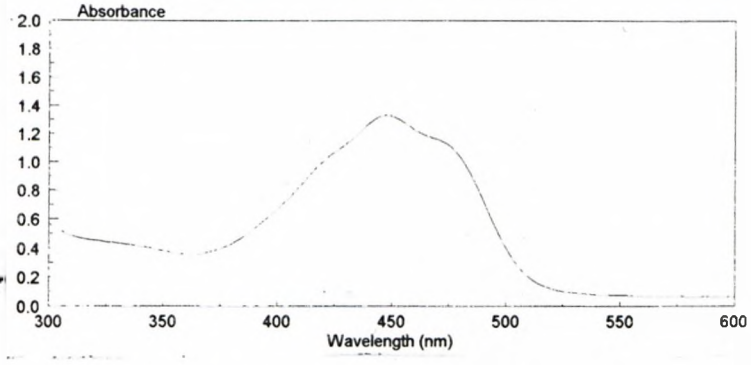
(iv) uncharacterized

(v) uncharacterized

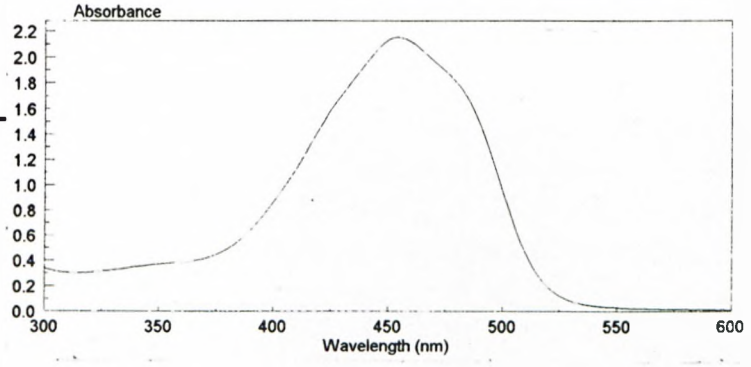
i


Absorbance
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
250
300
350
400
450
500
550
600
650
Wavelength (nm)

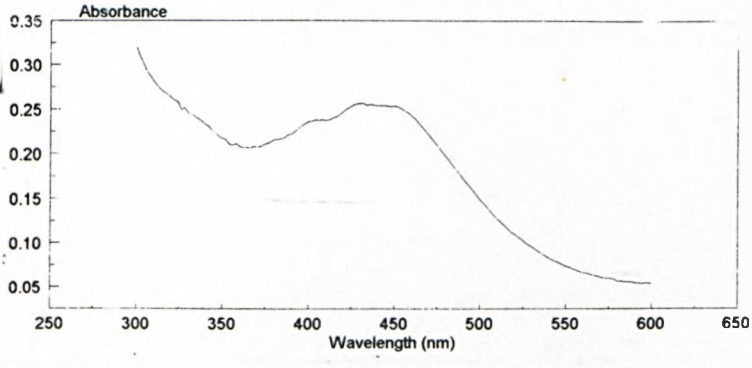
ii


Absorbance
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
300
350
400
450
500
550
600
Wavelength (nm)

iii


Absorbance
2.2
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
300
350
400
450
500
550
600
Wavelength (nm)

iv


Absorbance
0.35
0.30
0.25
0.20
0.15
0.10
0.05
250
300
350
400
450
500
550
600
650
Wavelength (nm)

v

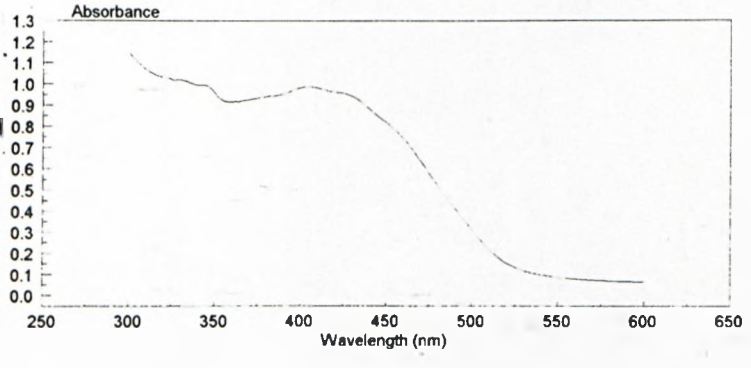
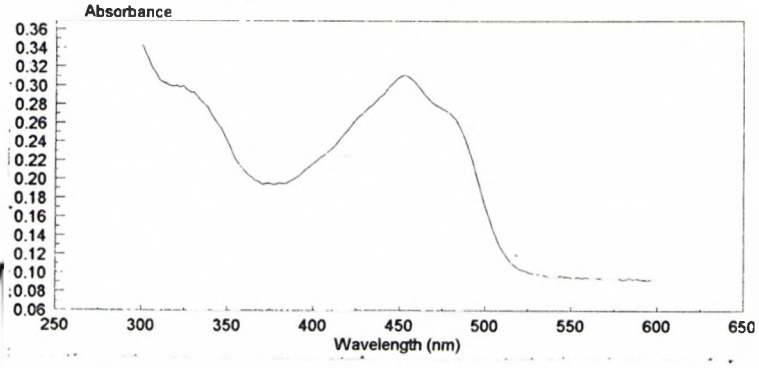
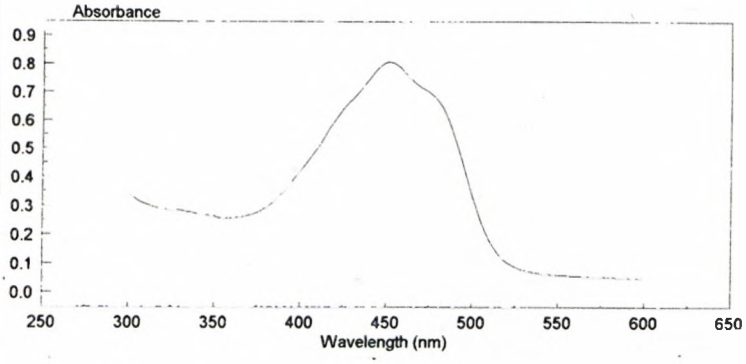

Absorbance
1.3
1.2
1.1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
250
300
350
400
450
500
550
600
650
Wavelength (nm)

Figure 5.4 (continued)
(b) *R. ruber*
Fraction identifications;
(i) cis-β-carotene, λ_{max} nm; 425, 450, 480
(ii) β-carotene, λ_{max} nm; 425, 450, 480
(iii) dyhydroxyisorenieratene, λ_{max} nm; 430, 460, 485
(iv) uncharacterized, λ_{max} nm; 425, 460, 495
(v) glycosides of monocyclic carotenoids, λ_{max} nm; 435, 460, 495

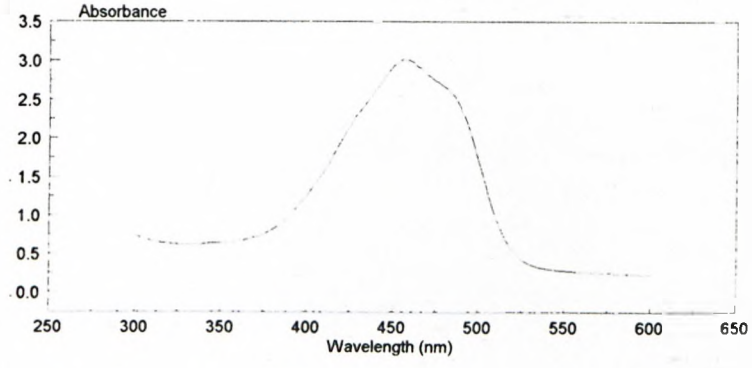
i


Absorbance
0.36
0.34
0.32
0.30
0.28
0.26
0.24
0.22
0.20
0.18
0.16
0.14
0.12
0.10
0.08
0.06
250
300
350
400
450
500
550
600
650
Wavelength (nm)

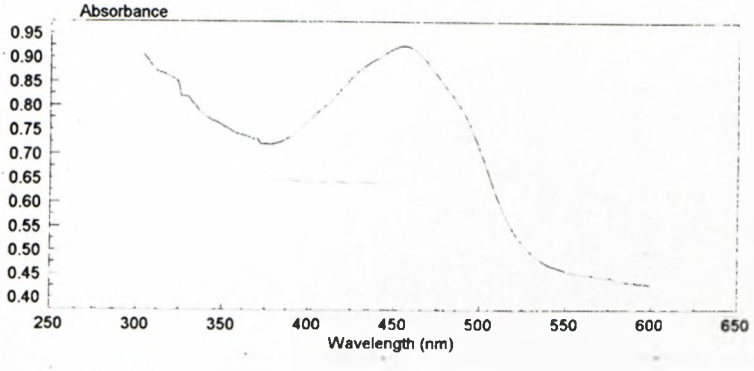
ii


Absorbance
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
250
300
350
400
450
500
550
600
650
Wavelength (nm)

iii


Absorbance
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
250
300
350
400
450
500
550
600
650
Wavelength (nm)

iv


Absorbance
0.95
0.90
0.85
0.80
0.75
0.70
0.65
0.60
0.55
0.50
0.45
0.40
250
300
350
400
450
500
550
600
650
Wavelength (nm)

v

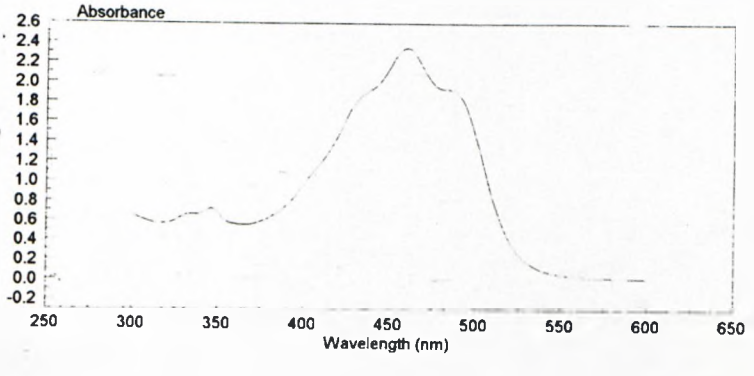
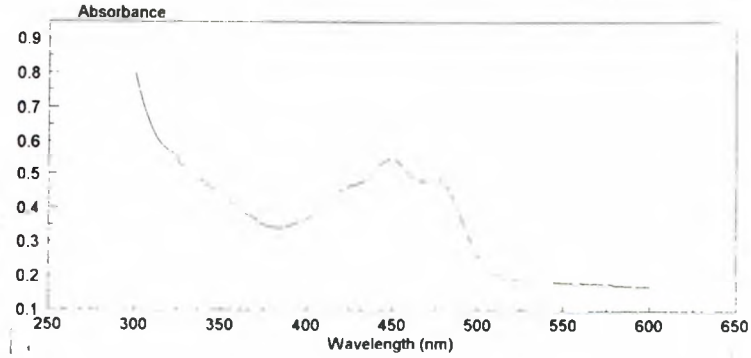
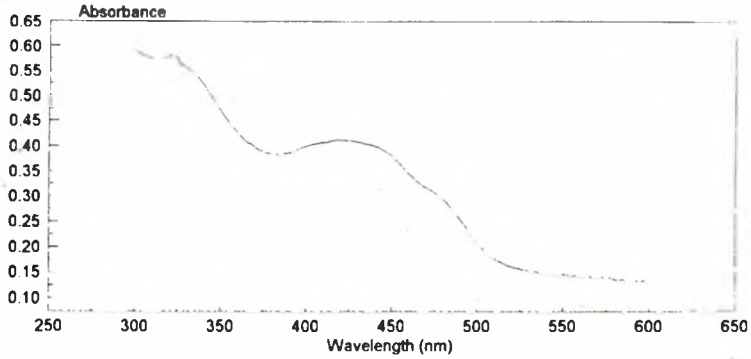

Absorbance
2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
250
300
350
400
450
500
550
600
650
Wavelength (nm)

Figure 5.4 (continued)
(c) *S. multivorum*
Fraction identifications;
(i) β-carotene, λ_{max} nm; 425, 450, 478
(ii) uncharacterized
(iii) zeaxanthin, λ_{max} nm; 425, 450, 475
(iv) uncharacterized

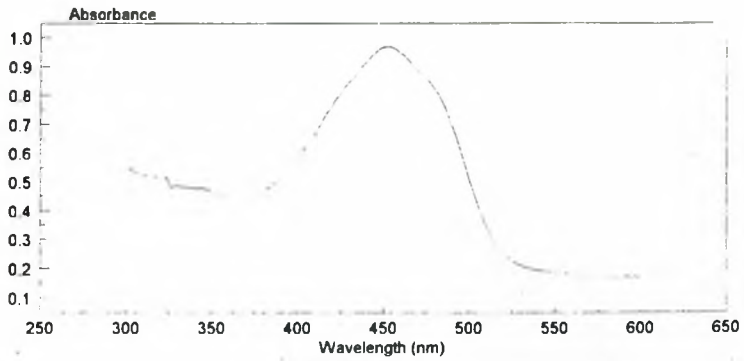
i


Absorbance
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
250
300
350
400
450
500
550
600
650
Wavelength (nm)

ii


Absorbance
0.65
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
250
300
350
400
450
500
550
600
650
Wavelength (nm)

iii


Absorbance
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
250
300
350
400
450
500
550
600
650
Wavelength (nm)

iv

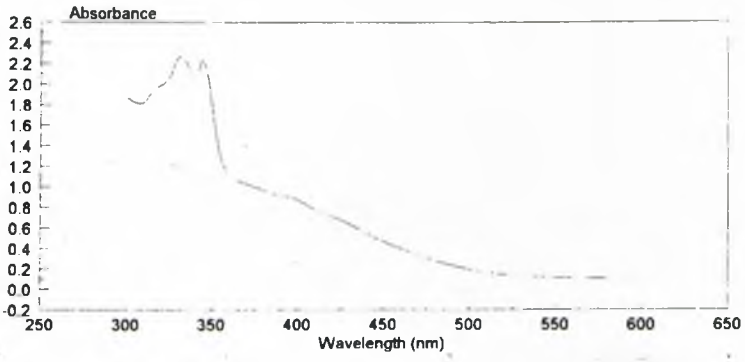
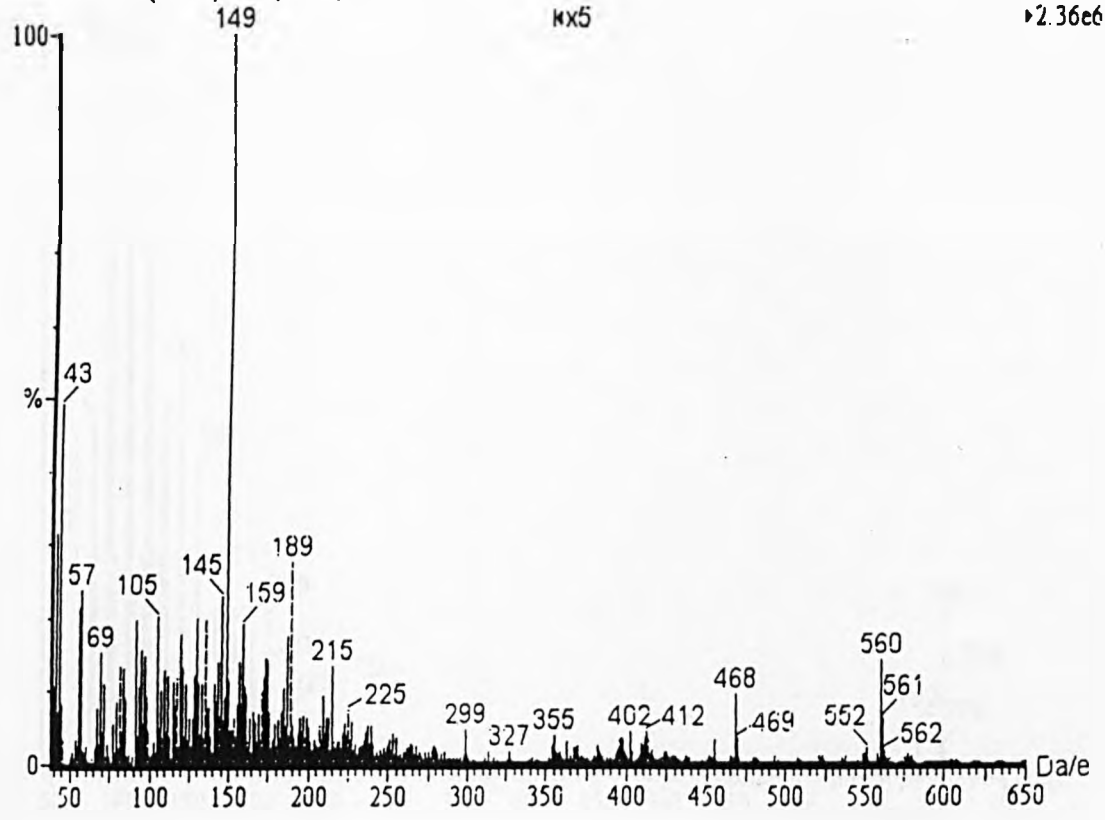

Absorbance
2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
250
300
350
400
450
500
550
600
650
Wavelength (nm)

Figure 5.5 Mass spectra of the main carotenoid pigments of the bacteria *B. linens*, *M. luteus* and *R. ruber*.
Mass spectra were obtained from carotenoids extracted and purified by TLC [2.2.26 (iii) and (iv)]. The fractions analysed from (a) *B.linens*, (b) *M. luteus* and (c) *R. ruber* are shown in Figure 5.4 and indicated in this figure with the exception of *M. luteus* where the fractions were purified by HPLC and are shown in figure 5.3 (b) [2.2.26 (ii)].

BL

CATHY16 60 (4.129) Cm (59:69)

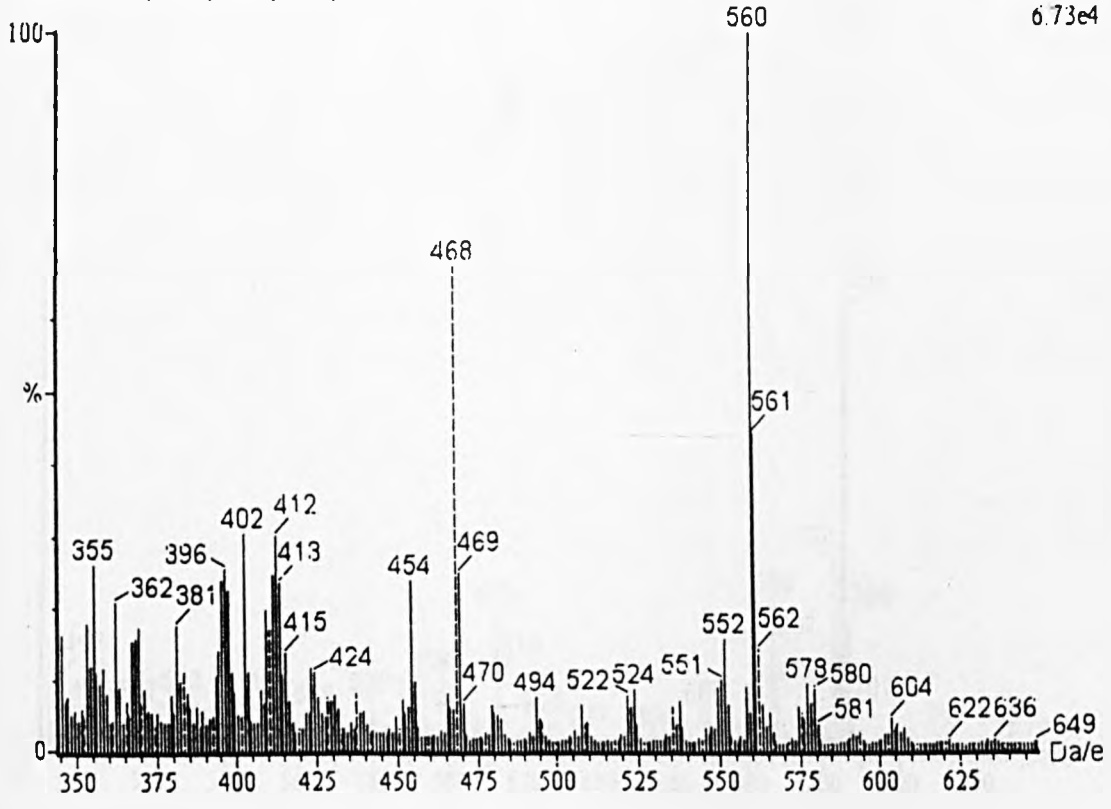
Scan EI+
2.36e6



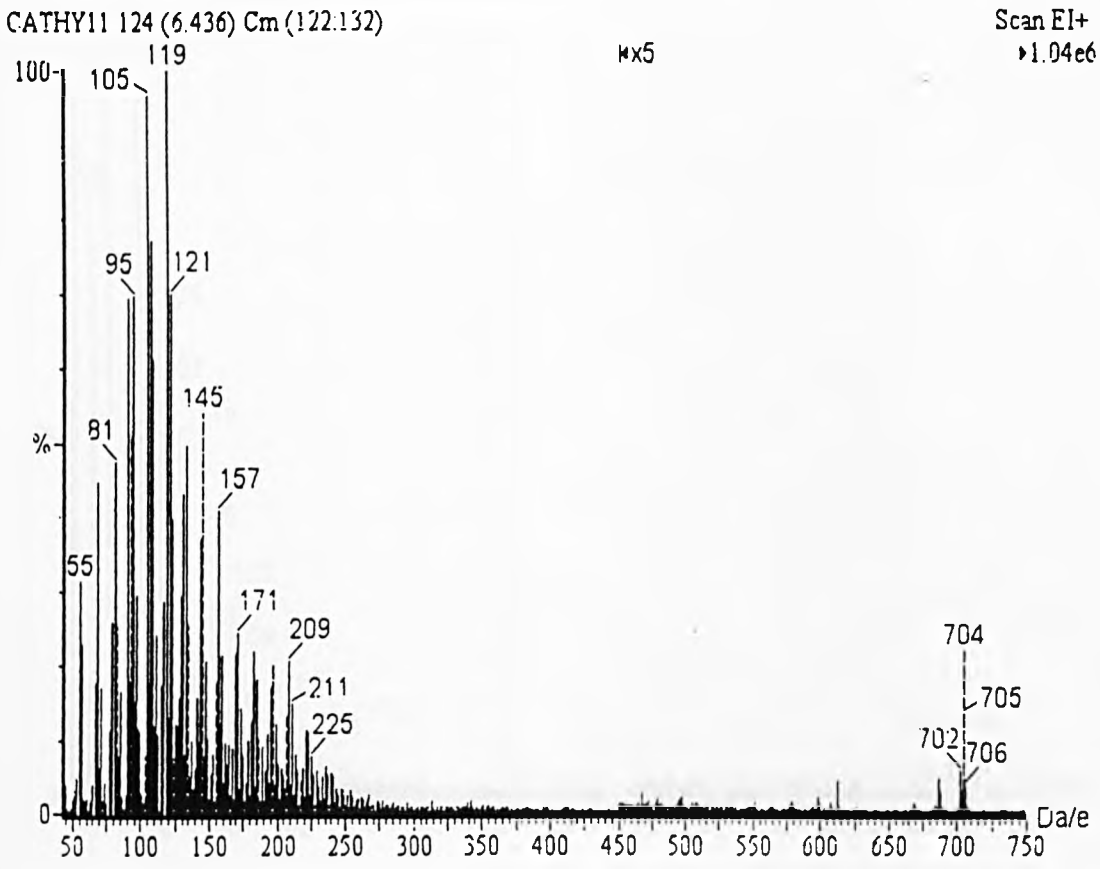
BL

CATHY16 60 (4.129) Cm (59:69)

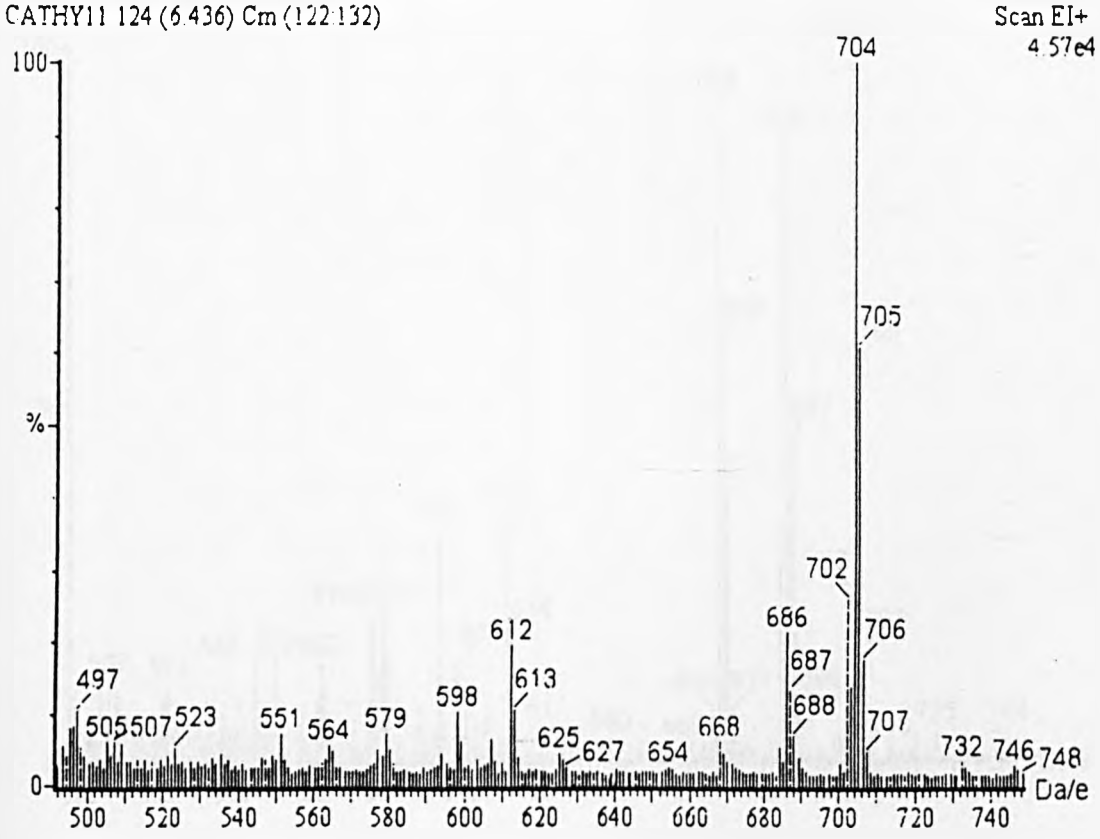
Scan EI+
6.73e4



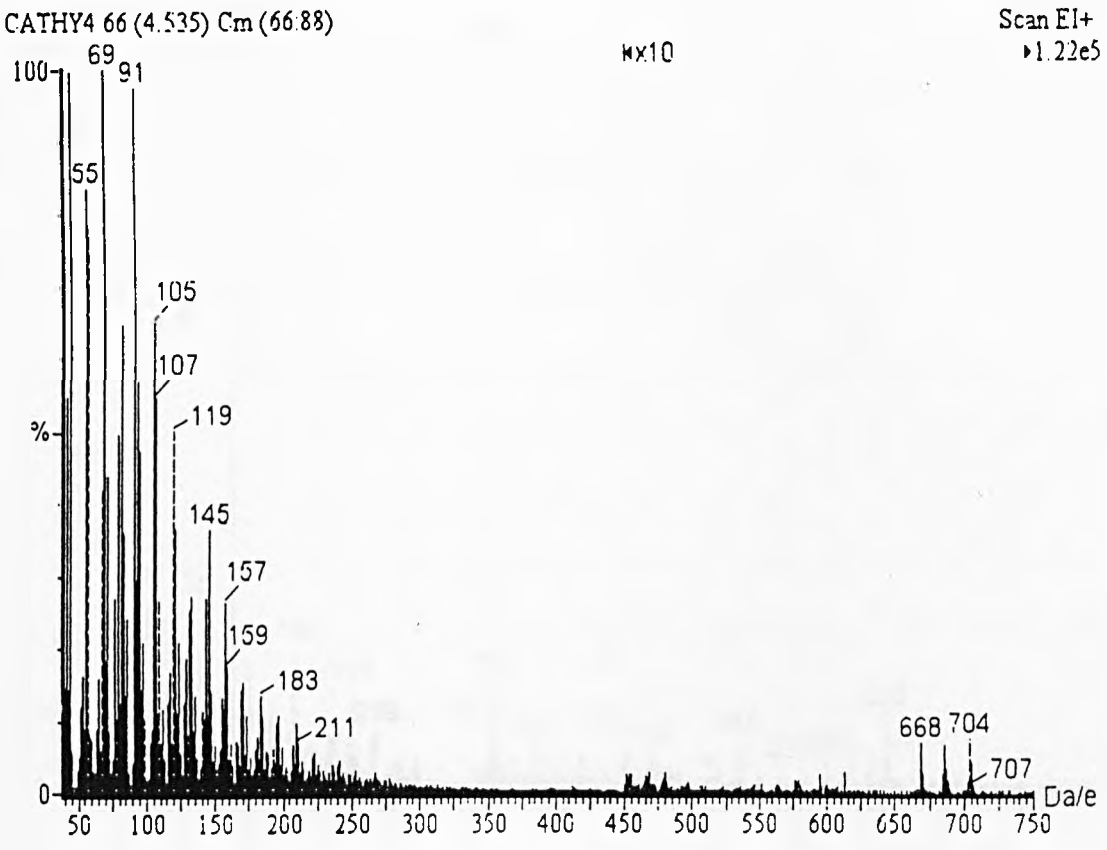
M. luteus 31/3/95



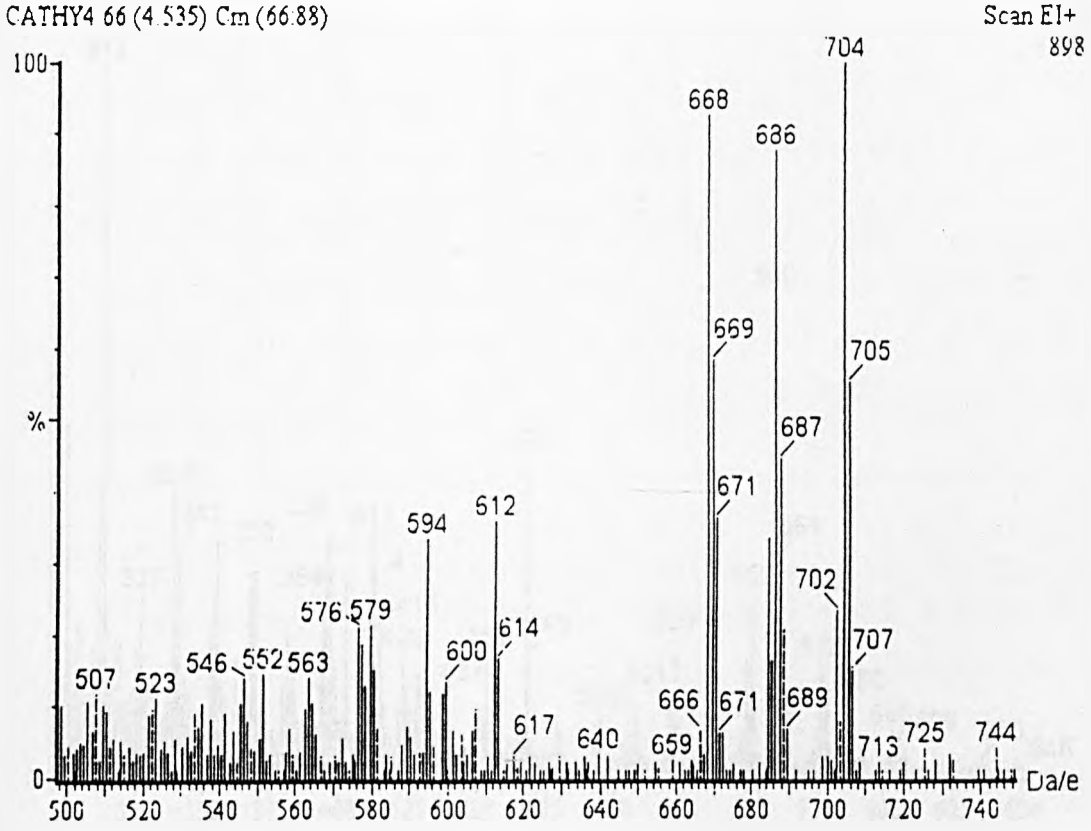
M. luteus 31/3/95



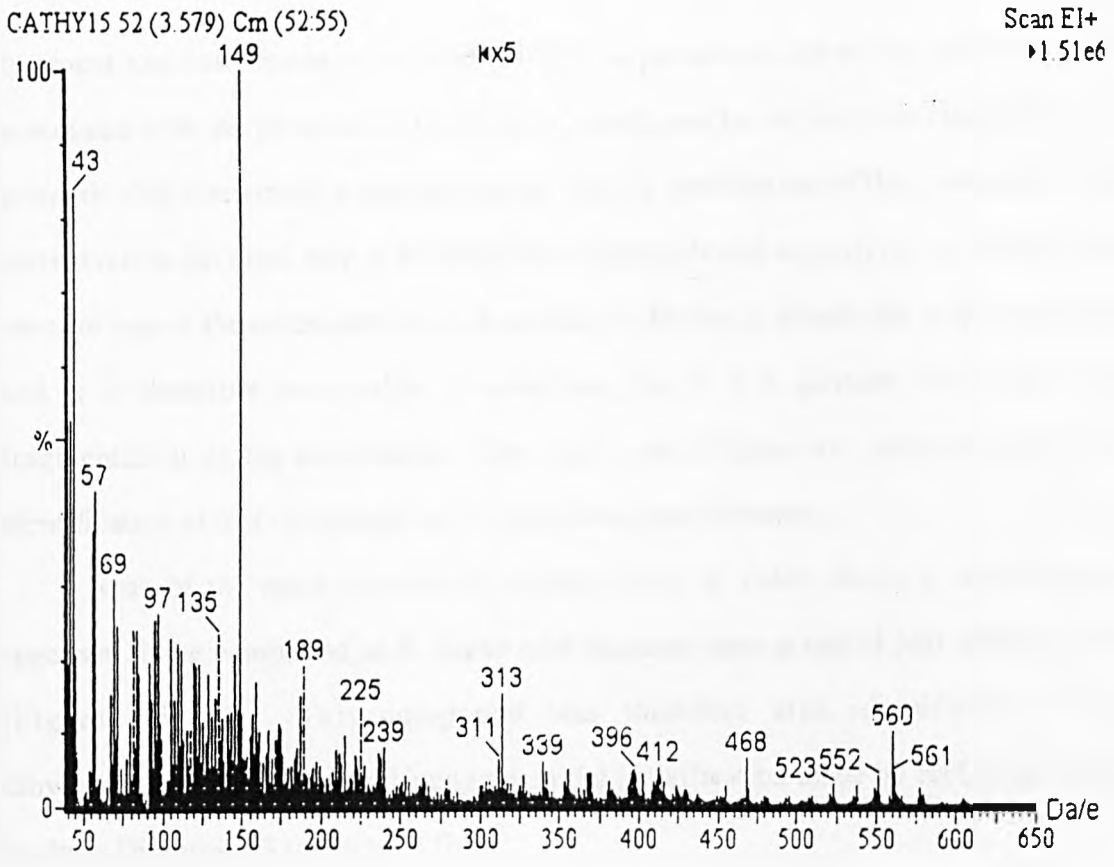
Carotenoid M.Luteus TLC fracn. 31/3/95



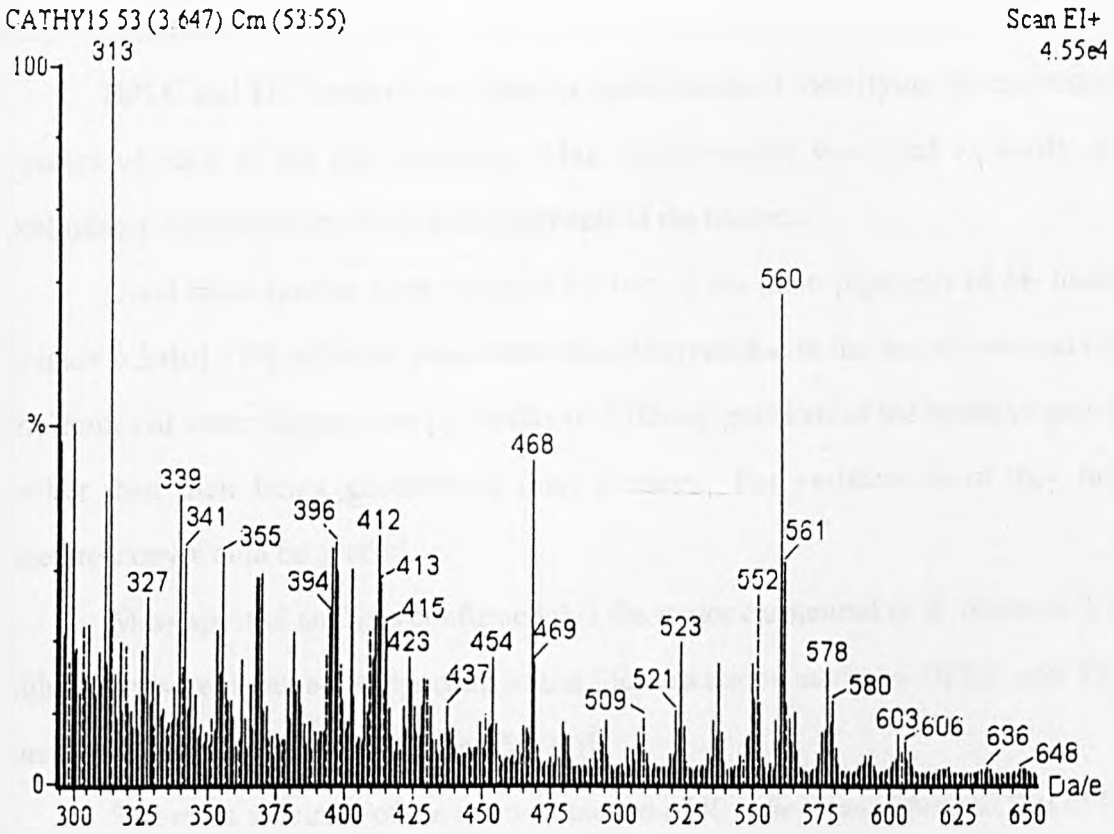
Carotenoid M.Luteus TLC fracn. 31/3/95



RR


CATHY15 52 (3.579) Cm (52:55)
Scan EI+
1.51e6
x5
149
43
57
69
97
135
189
225
239
313
311
339
396
412
468
523
552
560
561
%
0
Da/e
50
100
150
200
250
300
350
400
450
500
550
600
650

RR


CATHY15 53 (3.647) Cm (53:55)
Scan EI+
4.55e4
313
560
468
339
341
396
412
355
561
327
394
413
552
415
523
423
454
469
578
437
521
509
580
603
606
636
648
%
0
Da/e
300
325
350
375
400
425
450
475
500
525
550
575
600
625
650

fragment ion was present at *m/z* 468 (M-92). A prominent ion at *m/z* 149 frequently associated with the presence of plasticizers, could also be attributed to cleavage of an aromatic ring containing a hydroxy group. As the purification of this compound was carried out in the same way as for the other carotenoids and no such ion at *m/z* 149 was seen for any of the other samples, it is unlikely to be due to plasticizer in this one case and it is therefore reasonable to conclude that it is a genuine ion caused by fragmentation of the carotenoid. The mass spectral data are consistent with the identification of this compound as 3,3'-dihydroxyisorenieratene.

One of the main carotenoids isolated from *R. ruber* shows a similar mass spectrum to the compound of *B. linens* with fragment ions at *m/z* of 560, 468 and 149 [Figure 5.5 (c)]. This compound was therefore also identified as 3,3'-dihydroxyisorenieratene, confirming the initial identification made by HPLC and TLC analysis [Figures 5.3 (c) and 5.4 (b)].

(iv) Discussion

HPLC and TLC analysis provided an initial means of identifying the carotenoid content of each of the four bacteria. Mass spectrometry was used to verify the preliminary identifications of the main pigments of the bacteria.

Good mass spectra were obtained for two of the main pigments of *M. luteus* [Figure 5.5 (b)]. The different peak intensities observed due to the loss of one and two molecules of water suggests the possibility of differing positions of the hydroxy groups rather than their being geometrical (*cis*) isomers. For verification of this, nmr spectroscopy would be needed.

Mass spectral analysis confirmed that the major carotenoid of *B. linens* is 3,3'-dihydroxyisorenieratene, supporting initial identifications made by HPLC and TLC analysis [Figures 5.3 (a), 5.4 (a) and 5.5 (a)].

The mass spectrum of the main carotenod of *R. ruber* was similar to that of the main compound of *B. linens* [Figure 5.5 (c)]. This result was not expected as 3,3'-dihydroxyisorenieratene has previously only been identified in *Brevibacterium* sp.

Canthaxanthin has been previously thought to be the major carotenoid pigment of *R. ruber*. Verification of this result also needs to be made by nmr analysis.

Mass spectra were unobtainable for the main pigment of *S. multivorum*. This bacterium was not highly pigmented due to the presence of low levels of carotenoid. Pigments were extracted from approximately five litres of culture to obtain HPLC analysis, so more would be needed for mass spectra. This was not a feasible option. From HPLC and TLC analysis the pigment was tentatively identified as zeaxanthin [Figures 5.3 (d) and 5.4 (c)]. Mutant strains derived from this wild-type have produced substantial amounts of carotenoid allowing the characterization of (3*R*,3'*R*-) zeaxanthin as the major pigment from this strain (Bodi Hui, unpublished data).

5.4 Analysis of genomic DNA

(i) Extraction of DNA

Genomic DNA was extracted from 50ml of stationary-phase bacterial culture as described in Section 2.2.5. Typical yields were between 2 and 10mg of DNA. 10µg of genomic DNA was digested with the restriction enzymes *Bam*HI, *Eco*RI, and *Hin*dIII and double digests of *Bam*HI and *Eco*RI, *Bam*HI and *Hin*dIII and *Hin*dIII and *Eco*RI were also made (2.2.8). These enzymes were chosen on the basis of their recognition sequences. Only enzymes cutting within 6 or 8 base sequences were chosen so as to generate an even size distribution of fragments from high to low molecular weight. In principle, a 6-base recognition site should occur, assuming an even distribution of these sites, on average every 4^6 bases (Maniatis *et al.*, 1989). Fragments were separated by horizontal agarose gel electrophoresis (2.2.6) as shown in Figure 5.7 and 5.9. The gels were run slowly at 10V overnight in order to prevent smearing and to allow high resolution of the genomic fragments. DNA was transferred to Hybond N+ membranes as described in 2.2.13 (i).

(ii) Design of suitable probes

Genomic blots were used in Southern hybridization experiments with radiolabelled gene probes generated from the *E. herbicola* carotenoid biosynthesis cluster.

Initially the DNA was probed with three separate genes [radiolabelled as in 2.2.13 (ii)]. These encoded the first three enzymes of the biosynthetic pathway of *E. herbicola*. These probes hybridized non specifically to the immobilized DNA and no specific bands were observed (data not shown).

Smaller fragments (approx. 700 bp) of these genes were generated by restriction digestion. The digests were analysed by agarose gel electrophoresis (2.2.6) and the appropriate sized bands excised from the gel. They were purified (2.2.7) prior to being radiolabelled [2.2.13 (i)] and used as probes. No discernible improvement

was observed after hybridization; non-specific signals were still prominent (data not shown).

As an alternative approach the polymerase chain reaction (PCR) was used to amplify 300bp central regions of the three genes. The PCR reaction was carried out as in Section 2.2.11 with pAPU211 (100ng) as the template, at a primer annealing temperature of 50°C. The results of PCR, at magnesium concentrations of 0.5mM, 1.5mM and 2.5mM, are given in Figure 5.6. The PCR products were purified (2.2.7) before being radiolabelled [2.2.13 (ii)]. The lower band of the two products observed for the PCR of *crtE* was taken and purified.

This approach was adopted for a number of reasons. The distribution of restriction enzyme sites within the DNA would not easily allow the isolation of these regions. The use of a smaller probe would allow a more stringent hybridization with less chance of non-specific annealing. Because the reaction generated so much product a batch of DNA giving reproducible probe of high specific activity was generated. This could be used without frequent template preparation thus eliminating batch variation between probes. The primers for each of the three genes *crtE*, *crtI* and *crtB* are as follows;

crtB 5' primer; 5' - ATTACGCCCCGCATGGC - 3',

crtB 3' primer; 5' - CCCGCGCGGCATAGTTC - 3'

crtI 5' primer; 5' - TCACCACCTCGTCCAT - 3'

crtI 3' primer; 5' - TTCTGCCCCACCGGAT - 3'

crtE 5' primer; 5' - GCCGAGCTGCGCCGCGGTCA - 3'

crtE 3' primer; 5' - CGGCGACGAGGCGGAAGCAA - 3'

(iii) Hybridization experiments

Figure 5.8 shows a typical autoradiograph of *M. luteus* genomic DNA probed with the 300bp central portion of *crtI*. Bands of hybridization were observed that were

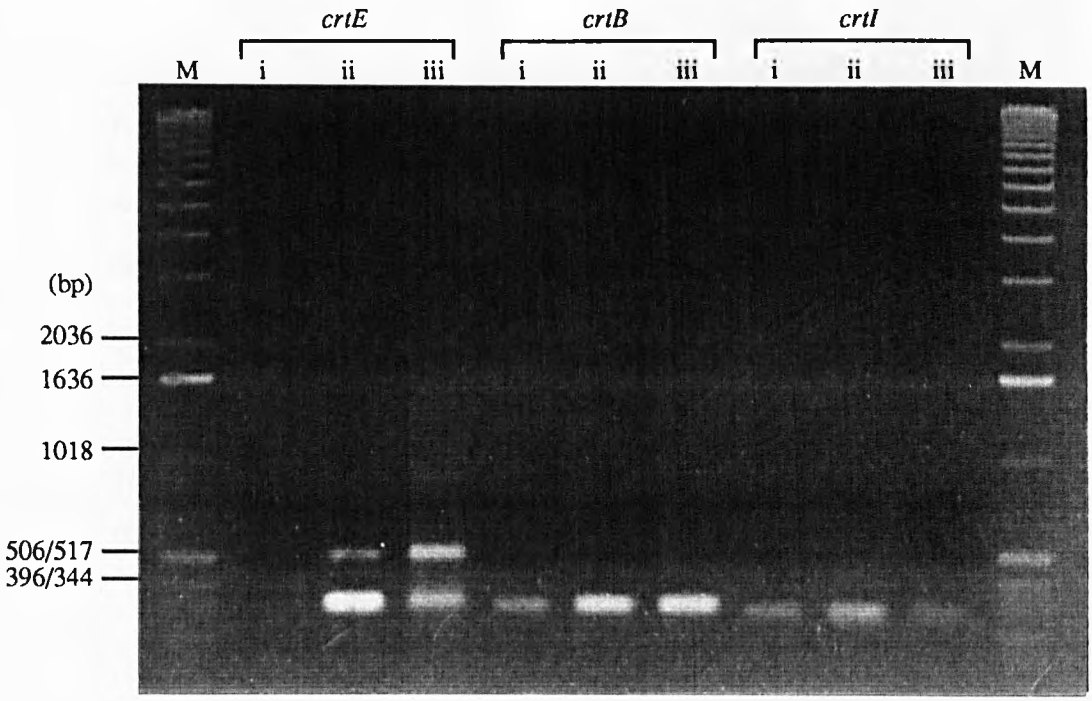


Figure 5.6 Agarose gel analysis of PCR products obtained by using primers designed for the genes *crtE*, *crtB* and *crtI*.
0.8% agarose gel analysis of PCR products obtained by using the primers described in 5.4.2 (ii) and pAPU211 as the template. PCR was carried out as described in 2.2.11 in PCR buffer with a Mg^{2+} concentration of (i) 2.5mM, (ii) 1.5mM and (iii) 0.5mM and an annealing temperature of 50°C. M - 1 kilobase ladder (Gibco-BRL).

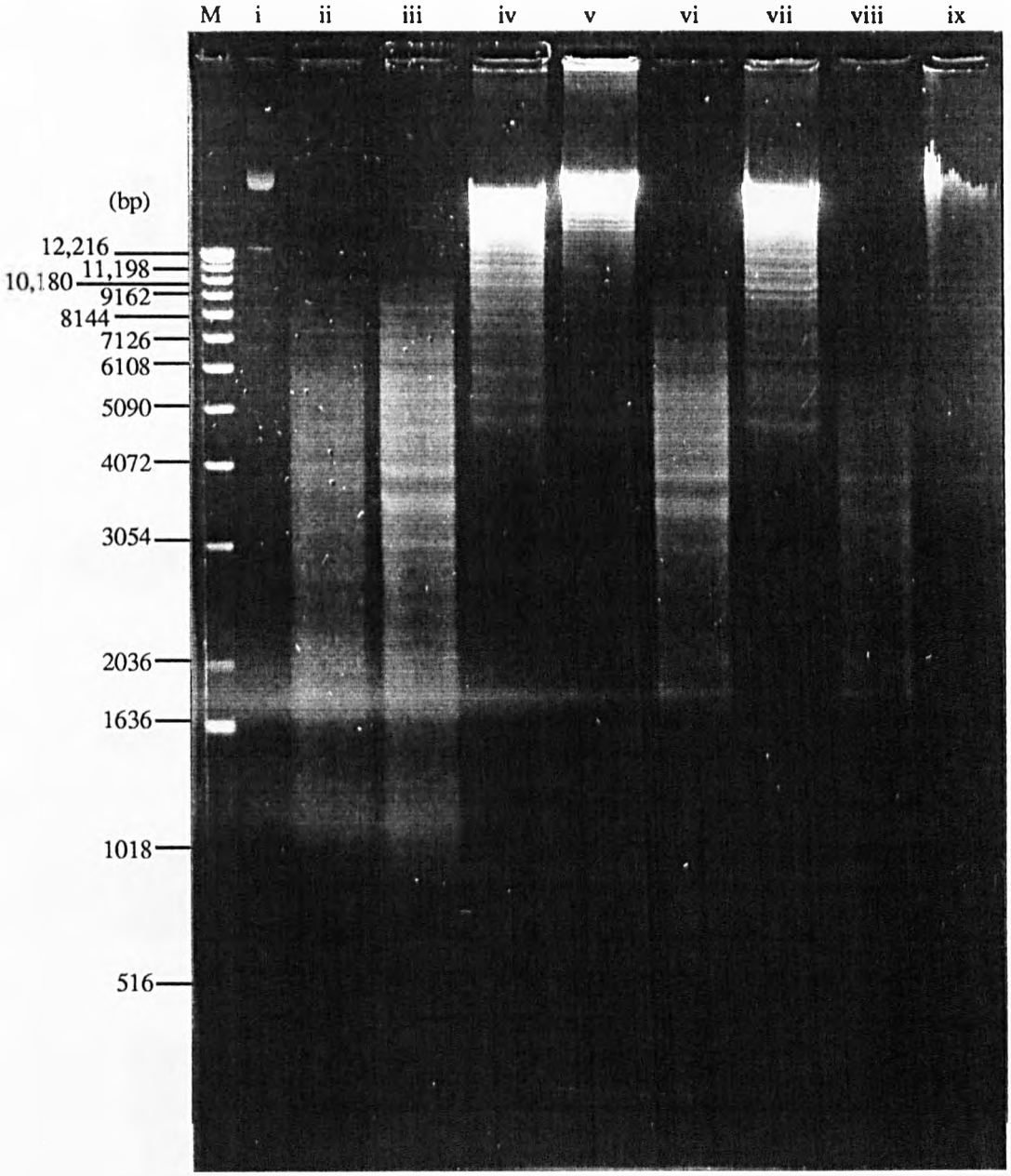


Figure 5.7 Agarose gel analysis of restriction enzyme digested *M. luteus* genomic DNA.

Restriction enzyme digests of *M. luteus* genomic DNA were analysed by agarose gel electrophoresis (0.8%). Digests performed were; (i) pAPU211 undigested, (ii) *E. coli* *Bam*HI, (iii) *Bam*HI, (iv) *Eco*RI, (vi) *Bam*HI-*Eco*RI, (vii) *Bam*HI-*Hin*dIII, viii) *Hin*dIII-*Eco*RI and (ix) undigested (2.2.8). M - 1 kilobase ladder (Gibco-BRL).

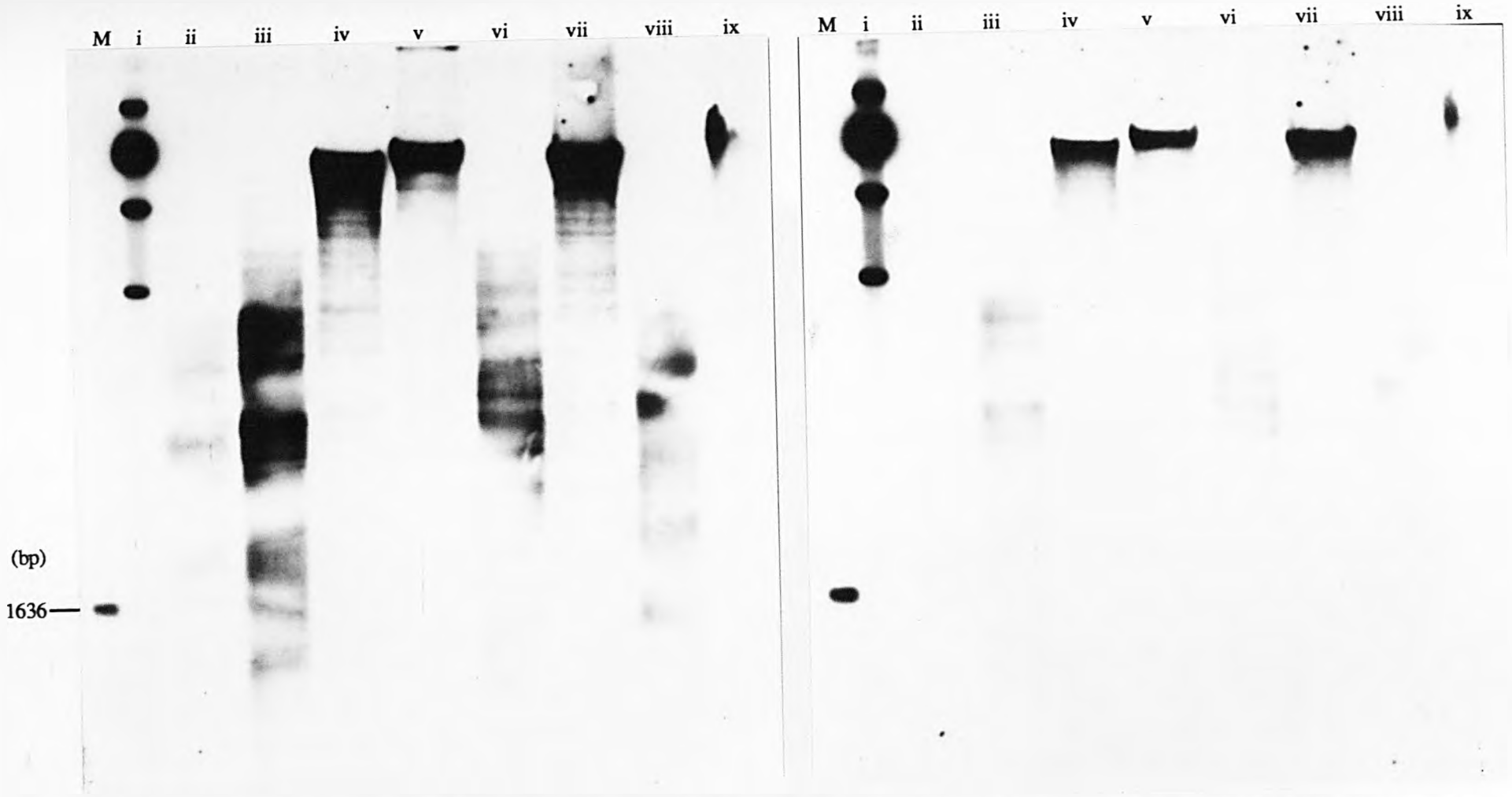


Figure 5.8 Southern blots of *M. luteus* genomic DNA probed with *crtI*. Restriction enzyme digested DNA (Fig. 5.7) was immobilised onto Hybond-N+ membranes (Amersham) and hybridised with the probe for *crtI* (2.2.13). Lanes correspond to those given in Figure 5.7. The membranes were washed as follows; (a) 1 x SSPE, 2 x 10 min, 50°C and (b) 0.1 x SSPE, 2 x 10 min, 55°C.

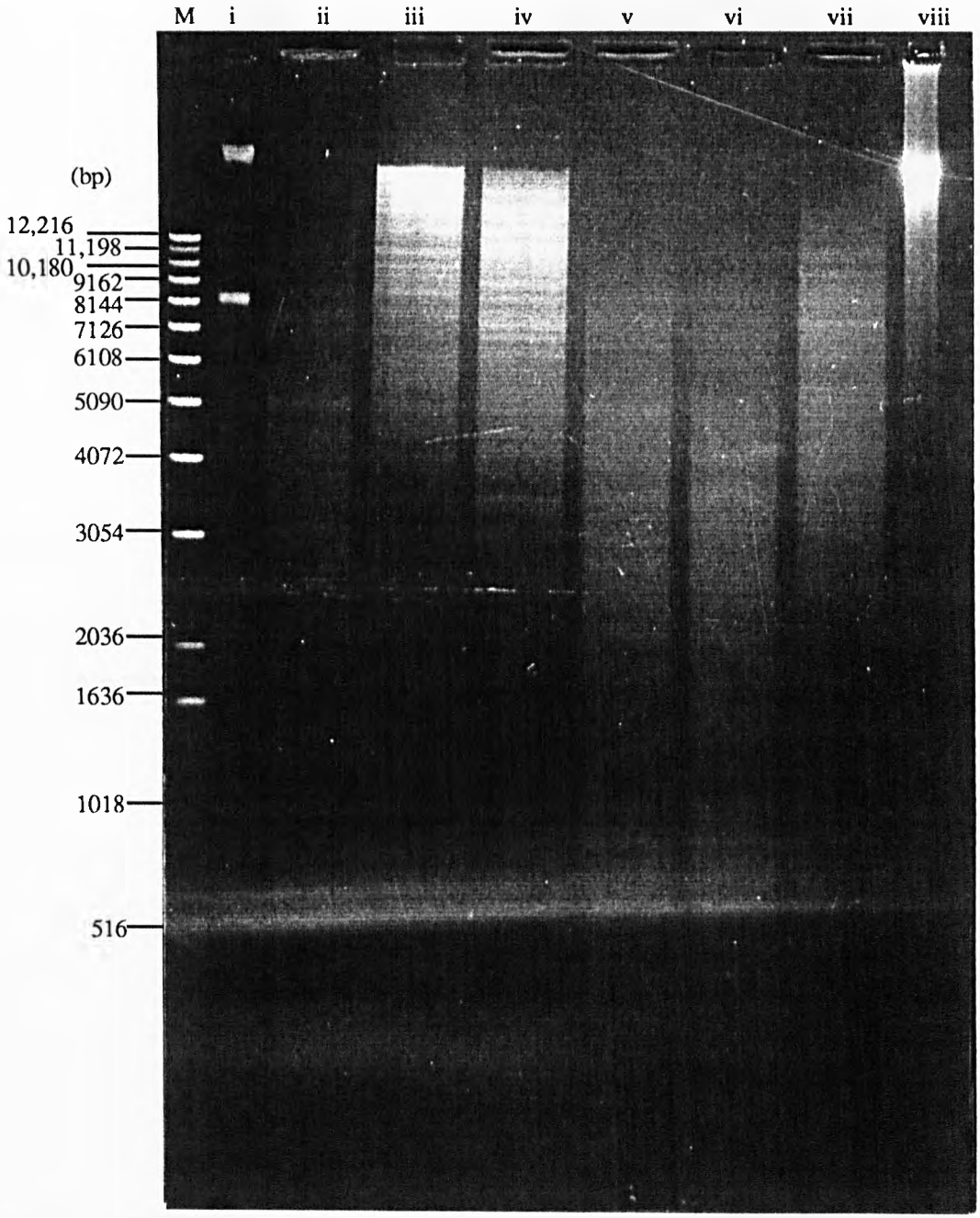


Figure 5.9 Agarose gel analysis of restriction enzyme digested *E. herbicola* genomic DNA.
Restriction enzyme digests of *E. herbicola* genomic DNA were analysed by agarose gel electrophoresis (0.8%). Digests performed were; (i) pAPU211 undigested, (ii) *Bam*HI, (iii) *Eco*RI, (iv) *Hin*dIII, (v) *Bam*HI-*Eco*RI, (vi) *Bam*HI-*Hin*dIII, (vii) *Hin*dIII-*Eco*RI and (viii) undigested. M - 1 kilobase ladder (Gibco-BRL).

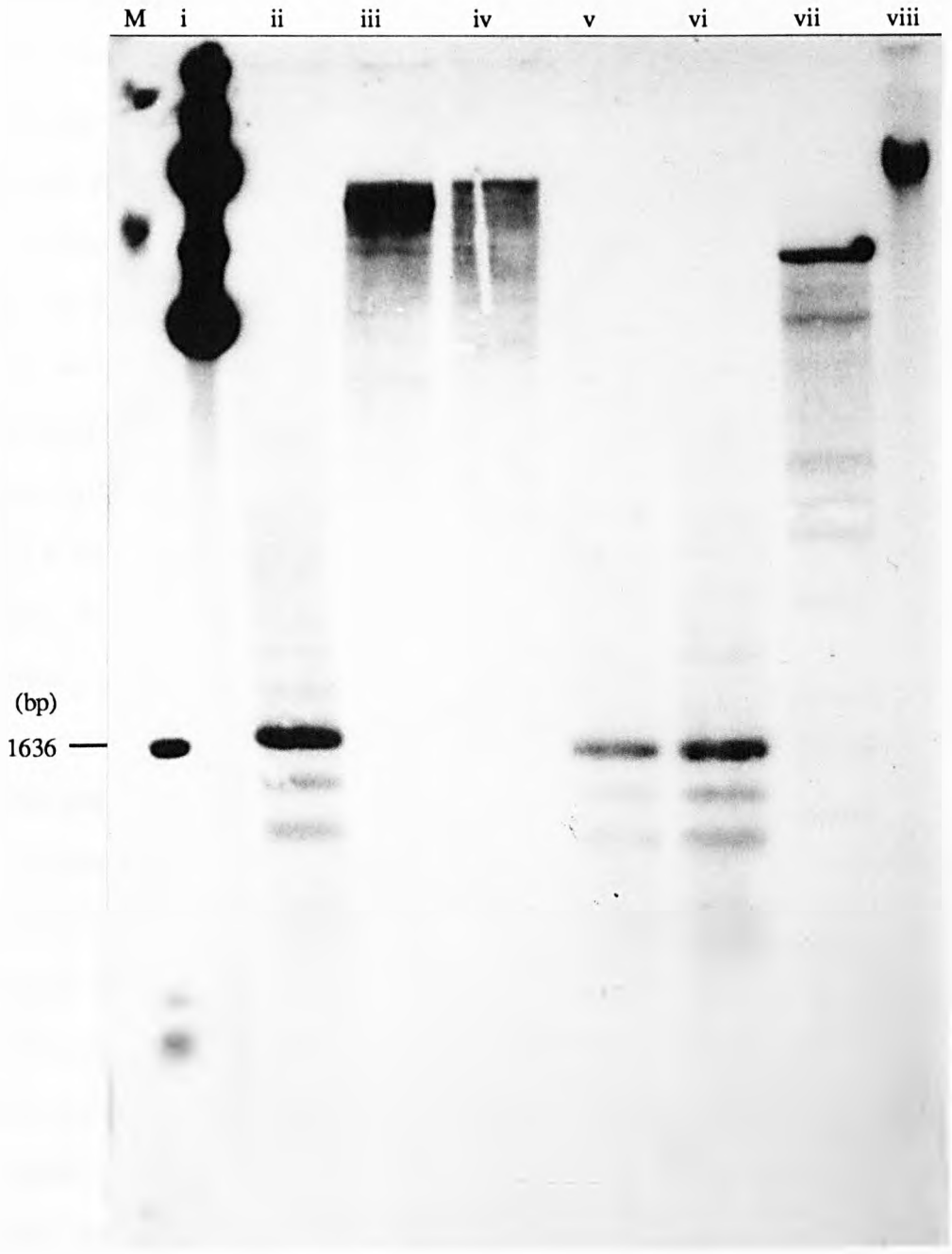


Figure 5.10 Southern blot of *E. herbicola* genomic DNA probed with *crtI*. Restriction enzyme digested DNA (Fig. 5.9) was immobilised onto Hybond-N+ membranes (Amersham) and hybridised with the probe for *crtI* (2.2.13). Lanes correspond to those given in Figure 5.9. The membrane was washed as follows; 2 x SSPE, 2 x 10 min, 50°C; 1 x SSPE, 2 x 20 min, 50°C and 0.1 x SSPE, 2 x 10 min, 60°C before exposure to x-ray film (Fuji RX-100).

still present after high stringency washing (0.1 x SSPE, 2 x 10 min., 55°C) [Figure 5.8 (b)]. The membranes were exposed overnight to X-ray film [2.2.13 (iii)]. *E. coli* genomic DNA digested with *Bam*HI was used as a negative control and the plasmid pAPU211 as a positive control. Similar patterns of hybridization were observed with the *crtE* and *crtB* probes (data not shown). The other bacterial species gave similar results with these probes (data not shown).

A positive control blot of *E. herbicola* genomic DNA, hybridized with these probes, was done in parallel. The restriction-digested DNA [restricted with the same enzymes and enzyme combinations as in 5.4 (i)] is shown in Figure 5.9. The autoradiograph of the genomic digests probed with the *crtI* 300bp fragment is shown in Figure 5.10. The membrane was washed sequentially in 2 x SSPE, 2 x 10 min.; 1 x SSPE 2 x 10 min. at 50°C; and 0.1 x SSPE, 2 x 10 min., at 60°C and exposed overnight. As was observed with the blot of *M. luteus* multiple bands of hybridization are present.

(iv) Discussion

To detect a single copy gene of mammalian genomic DNA by Southern blotting, 10µg of DNA need to be analysed (Maniatis *et al.*, 1989). Digestion of 10µg of bacterial genomic DNA thus ensures that if the *crt* genes are present as single copies they would be detected, always assuming that there is enough sequence conservation between the probe and target to allow hybridization.

Multiple bands of hybridization were observed with the probes *crtI*, *crtB* and *crtE*, in *M. luteus*. Similar results were obtained for the other three bacteria although the data are not shown.

The genomic blot of *E. herbicola* was performed as a positive control. An intense band of hybridization was seen at 1636bp. Analysis of the restriction enzyme map of pAPU211 showed that a *Bam*HI digest releases a fragment of this size containing *crtI*. However other fainter bands of hybridization were also observed.

In the genomic blot of *M. luteus* some faint hybridization was observed with *E. coli*, Figure 5.8 lane (ii) (present as a negative control), but was removed after high

stringency washing, implying that the hybridization was non-specific and the stringency of washing was adequate. The high molecular weight bands of hybridization, seen in both blots [Figs. 5.8 lanes (iv), (v), (vii) and (ix) and, 5.10 lanes (iii), (iv) and (vii)], could be disregarded as non-specific binding of probe to areas of high DNA concentration. However these bands of hybridization may also represent specific hybridizations to inadequately digested genomic DNA. Restriction enzyme digestion of genomic DNA was repeated numerous times with increasing concentrations of enzyme with no observable improvement. Multiple bands of hybridization are present [Figure 5.8 (b), lanes (iii), (vi) and (viii)] that cannot be attributed to the above phenomena and are still apparent after high stringency washing, suggesting that these bands represent genuine specific hybridization to homologous sequences. Comparison of the carotenoid biosynthesis genes of *E. herbicola* with each other, using the Alignment program of the GCG package shows regions of nucleotide homology. This would enable the sequences to hybridize with each other in a Southern blotting experiment. It is proposed that this could explain the presence of multiple bands of hybridization seen with the *E. herbicola* blot, and it is assumed that the same similarities in *crt* sequences occur in *M. luteus* thus also giving multiple bands of hybridization. Support for this theory lies in the fact that very similar patterns of hybridization were obtained when the probes *crtB* and *crtE* were used.

The amount of sequence conservation between probe and target was assessed, by the strength of hybridization signal seen after washing the membrane, (to remove non-specific hybridization) with increasing stringency. An alternative approach would be to alter the stringency of the hybridization reaction (Maniatis *et al.*, 1989). This could be performed on test, replicated membranes, by altering either the temperature of hybridization or the salt concentration of the hybridization solution. Instead of removing non-specific hybridization by washing, this procedure aims for optimum signal to background ratio negating the need for washing.

This approach of Southern blotting was undertaken as a feasibility study to discover if probing with heterologous probes was a reasonable approach in the search

for new bacterial carotenoid biosynthesis genes. As positive bands of hybridization were apparent it was concluded that the approach was possible. The probes were then used to isolate *crt* genes from genomic libraries constructed for each of the bacteria, as described in the following chapter.

CHAPTER 6

THE CONSTRUCTION AND SCREENING OF BACTERIAL GENOMIC LIBRARIES

6.1 General introduction

There are a number of approaches available for the isolation of novel carotenoid biosynthesis genes. In this study, genomic libraries were constructed and screened with previously characterized carotenoid gene probes. This chapter reports the construction of genomic libraries for *M. luteus*, *B. linens*, *R. ruber* and *S. multivorum*, by cloning 9-23Kb fragments of genomic DNA into the λ replacement vector λ GEM-11 (Promega) derived from EMBL3.

A genomic library should represent all sequences of the genome. These cloned sequences need to be large enough to contain whole genes yet small enough to enable easy analysis by restriction mapping. Preferably, cloned sequences should overlap in terminal regions to enable rescreening of the library for contiguous DNA sequences.

Numerous technical advances have allowed the rapid construction and screening of libraries. *In situ* plaque hybridization has been developed to enable the latter (Benton and Davis, 1977). Plaques are transferred to a membrane and immobilized to facilitate screening. *In vitro* packaging extracts have been developed to generate high titres of recombinant phage (Rosenberg, 1987; Hohn and Murray, 1977). The components needed to produce libraries are vector DNA, high molecular weight DNA and an *in vitro* packaging extract. Genomic DNA is partially digested to produce smaller overlapping fragments that can be ligated to a λ vector. The recombinant phage is then packaged *in vitro*. The packaged phage are grown and maintained in the appropriate *E. coli*.

(i) Bacteriophage λ

The genome of bacteriophage λ is approximately 50Kb long and is tightly packed into an icosahedral protein head structure. A protein tail, assembled separately

from the head, is added, helping to restrain the DNA in its condensed state. The tail can also determine the bacterial host.

The genome of bacteriophage λ is a linear molecule organized into three regions, these being depicted as the left and right arms and the central 'stuffer' fragment, (Figure 6.1) (Campbell, 1971). The left arm contains the genes necessary for the production of the head and tail proteins of the bacteriophage particle and the packaging of the genome. The right arm contains genes responsible for DNA replication. The central stuffer fragment contains genes involved in the lysogenic life cycle and is not required for growth or maintenance of the phage. This portion can be removed and replaced with exogenous DNA. At either end of the linear molecule there are single-stranded regions 12 base pairs long. These are known as the *cos* (*cohesive*) ends and are complementary to each other. They can anneal together to give a circular molecule or a concatenated one which is the preferred substrate for packaging.

(ii) Infection of *E. coli* cells

The tail of the bacteriophage particle is adsorbed onto the *E. coli lamB* gene product localized in the outer membrane. This gene product is expressed in the presence of maltose but not glucose. Adsorption of the bacteriophage to the cell is temperature independent and is facilitated by the presence of magnesium ions, but the insertion of the λ genome only occurs at 37°C and is a very rapid process.

(iii) Bacteriophage growth; lysis and lysogeny

Upon entering a host bacterium the cohesive termini undergo base pairing to give a circular molecule with staggered nicks. These are subsequently repaired by the ligase of the host. The closed circular form now serves as a template for transcription and replication during the early stages of infection.

Bacteriophage λ is a temperate phage, i.e. there are two different growth cycles one lytic and the other lysogenic

During lytic growth λ DNA is transcribed and translated to produce all the

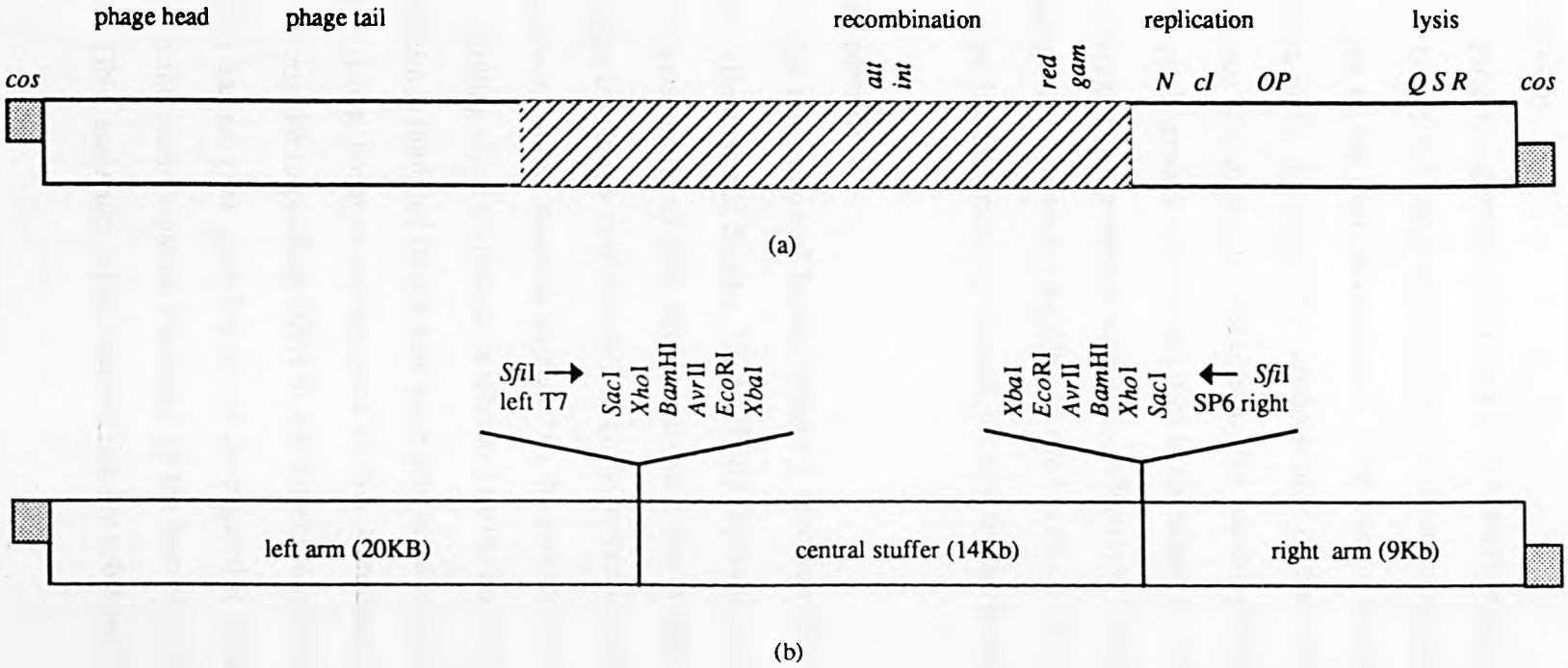


Figure 6.1 Physical and genetic map of wild-type bacteriophage λ and structural map of the λ-GEM-11 vector.
(a) Wild-type bacteriophage λ, indicating the positions of some important genes, with their general role indicated above. The hatched box indicates the area that is not essential for the propagation of λ and which can be replaced with exogenous DNA. (b) Structural map of the λ-GEM-11 vector indicating size of fragments and positions of restriction enzyme sites (Kaiser and Murray, 1986; Promega).

proteins required for replication and the production of new bacteriophage particles. Host cells can lyse within about 30 minutes after infection, releasing about 100 new phage particles.

Lysogenic growth is favoured if no bacteriophage particles are formed. The bacteriophage DNA integrates into the host chromosome and is replicated as though it were a part of the host chromosome. This can be switched to lytic growth by uv irradiation or by depleting the carbon source of the cells. The establishment and maintenance of lysogeny is controlled by the bacteriophage-encoded repressor protein cI. Lysogenic growth is favoured if cI is abundant; in early infection cI binds to all bacteriophage DNA promoter sequences, effectively blocking transcription of phage-encoded genes. The vector λ-GEM-11 used in this work lacks both the repressor and the site-specific integration genes and, because of this, is restricted to lytic growth.

(iv) Replication

The replication of bacteriophage λ involves both its own and host-encoded enzymes (Enquist and Skalka, 1973). Early replication proceeds in a closed circular form bidirectionally to give circular progeny; (theta replication), (Figure 6.2). Late replication involves a switch to rolling circle replication (also shown in Figure 6.2). A nicked chromosomal strand is displaced by the procession of continual DNA synthesis.

Rolling circle replication is inhibited by the host-encoded RecBC endonuclease. This in turn is inhibited by the *gam* gene product, allowing rolling circle replication to proceed. Long, linear or concatenated λ DNA is produced and is subsequently cleaved at every cos site to produce DNA of the correct length for packaging. The vector, λ GEM-11 has no *gam* gene but is still propagated in RecBC$^+$ host cells. The phage relies on the recombination functions of the host to produce dimeric or multimeric circular DNA molecules which are also suitable substrates for packaging.

(v) λ-GEM-11 replacement vector

Genomic DNA, partially digested with *Sau*3AI, can be cloned into the

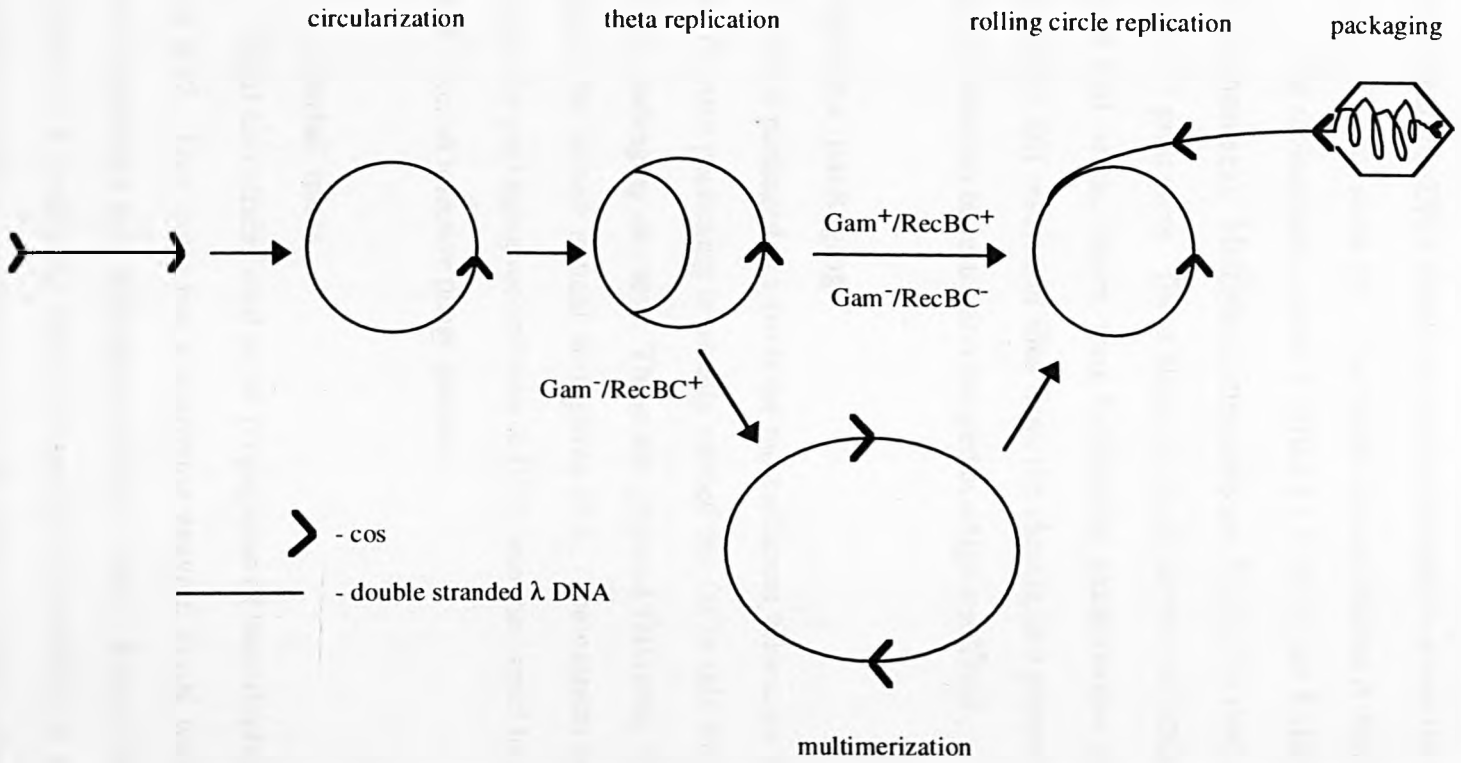


Figure 6.2 The replication pathways of bacteriopage λ. The linear phage molecule, upon entering a host cell, circularizes via its cohesive ends. Replication then proceeds bidirectionally (theta replication). Rolling circle replication proceeds if the RecBC inhibitor Gam is present (Gam is unnecessary if the RecBC is defective). This produces a linear concatenated molecule which is then cleaved at the cos sites, allowing the DNA to be packaged into an infective phage particle. Multimeric circular molecules are formed if the phage lacks the *gam* gene, but replicates in host RecBC$^+$ cells, and is a suitable substrate for packaging (Kaiser and Murray, 1986).

dephosphorylated *Bam*HI arms of the λ GEM-11 vector (Figure 6.1b). The stuffer fragment is prevented from religating to the arms by secondary digestion with *Eco*RI. The small *Eco*RI-*Bam*HI fragments are removed by selective precipitation with isopropanol. DNA fragments within the range 9-23kb can be cloned. Partially digested genomic DNA should be size fractionated to avoid the cloning of two or more fragments into the same site. The whole cloning scheme is depicted in Figure 6.3.

The replacement vector λ-GEM-11 is based on EMBL3 with several added features (Promega). Multiple cloning sites are flanked by dual opposed bacteriophage SP6 and T7 promoters. These allow synthesis *in vitro* of RNA probes corresponding to each end of the insert, thus facilitating chromosome mapping and walking. Asymmetric *Sfi*I restriction sites flank the cloning and promoter region. These sites enable the insert to be restriction mapped at a high resolution.

(vi) *In vitro* packaging

DNA packaged *in vitro* is the most efficient vehicle by which to enter bacterial hosts. *In vitro* packaging is usually carried out (as in this work) with commercially available packaging extracts. These are prepared following the induction of *E. coli* lysogenic for certain mutant derivatives of λ. The extracts contain all the proteins necessary for packaging recombinant λ DNA into the capsid head and for its assembly to the tail, i.e. an infective phage particle.

(vii) Bacterial hosts

Most host strains used in the propagation of bacteriophage λ are derivatives of *E. coli* K12. This strain has a restriction enzyme *Eco*K which cleaves DNA that contains unmodifed sites (Raleigh and Wilson, 1986). Even if vector DNA is protected by growth in a modifying strain the foreign insert DNA is not. Bacteriophage is usually propagated in a restriction-deficient strain (r_K^-) but can be grown in a modifying but not restricting strain (*hsdR*$^-$) or a strain that does not modify (*hsdS*$^-$) or restrict (*hsdR*$^-$ *hsdS*$^-$).

Another consideration in the choice of bacterial host strains is their

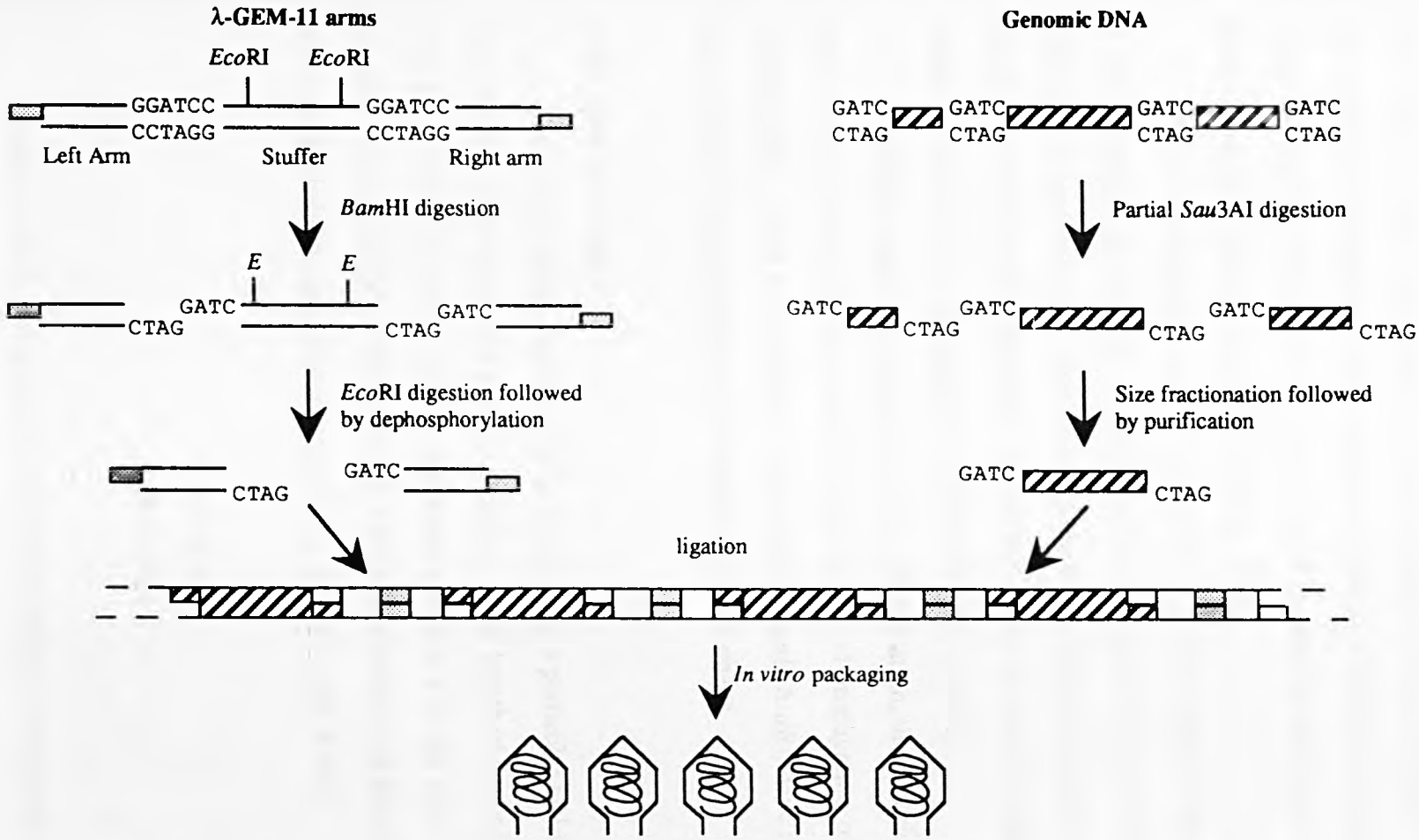


Figure 6.3 An overall scheme for the construction of genomic libraries in λ-GEM-11 *Bam*HI arms. The vector λ-GEM-11 is digested with *Bam*HI and subsequently *Eco*RI, preventing the stuffer fragment from religating. The arms are also dephosphorylated to prevent ligation to each other. Genomic DNA is partially digested with *Sau*3AI and size fractionated to give 9-23Kb fragments. Ligation of the DNA and λ arms produces a long concatenated DNA molecule which is then packaged *in vitro* by use of a commercial packaging extract (Promega).

recombination characteristics. Recombination between homologous sequences of DNA is undesirable as it can lead to duplication and deletion. This occurs rarely even in Rec^+ strains. If this is a problem the bacteriophage can be propagated in Rec^- hosts. Rec^- strains have their own difficulties in that growth of the bacteriophage is arduous due to a decreased burst size and an increased number of dead cells to which the bacteriophage still adheres (Kaiser and Murray, 1985).

The λ vectors can carry amber mutations in genes coding for the coat proteins of the bacteriophage particle. Specific hosts can carry one or two strong amber suppressors, *supE* and *supF*, that insert glutamate or tyrosine, respectively, at UAG codons and are not interchangeable. Suppressor strains do not affect the growth of λ vectors that do not carry amber mutations (Maniatis *et al.*, 1989).

When the above factors are taken into consideration, the *E. coli* strain LE392 (*supE*, *supF*, *hsdR*) is used routinely for the assay and propagation of recombinant bacteriophage. Other strains, KW251 and NM621 (*supE*, *hsdR*), have also been used in this work for the propagation of recombinant bacteriophage.

(viii) Size of library

The size of library that needs to be screened for a particular DNA sequence can be calculated, if the size of the genome and the average length of donor DNA is known (Clarke and Carbon 1976). If x is the insert size and y is the size of the haploid genome, the number (N) of clones required is calculated from the following equation, where the probability (p) of a sequence existing in single copy is 99%;

$$N = \frac{\ln (1-p)}{\ln (1- x/y)}$$

Based on the *E. coli* genome and assuming that an average size insert is 20Kb, the library required to be fully representative would contain 1000 recombinant clones.

6.2 The preparation of genomic DNA for use in constructing genomic libraries

(i) Optimizing conditions for partial digestion of bacterial genomic DNA

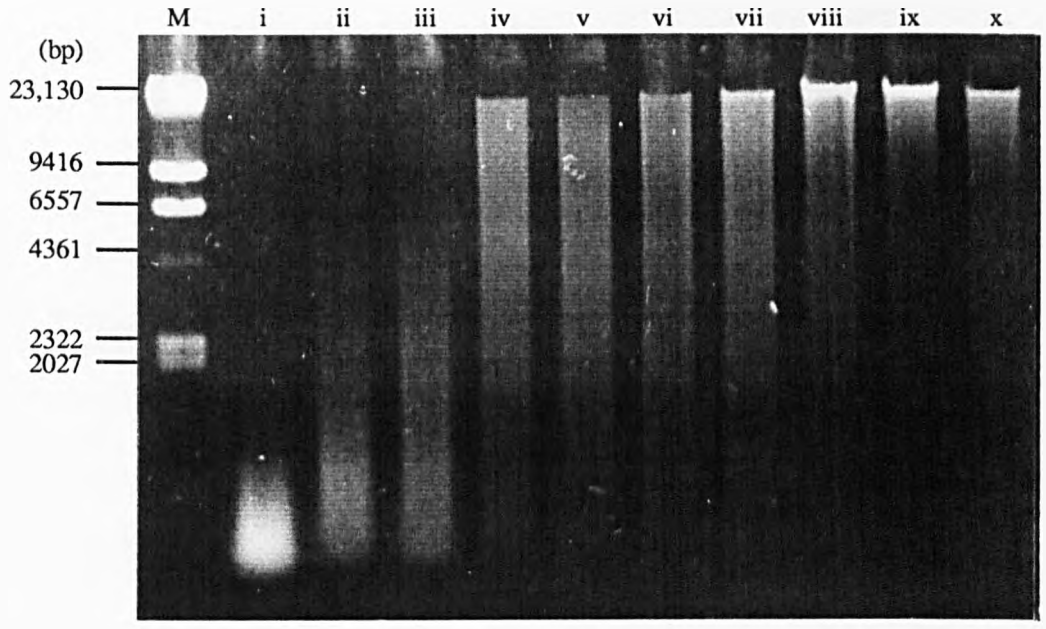
For the generation of a representative genomic library it is necessary to reduce the size range of the genomic DNA. This was done by partial restriction enzyme digestion with *Sau*3AI. Conditions were optimized in small-scale reactions before large-scale digestion of genomic DNA in preparation for a genomic library. Genomic DNA (1µg) was digested with increasing dilutions of *Sau*3AI containing from 1 to 0.0035 units of enzyme (30 minutes, 37°C) [2.2.14 (i)]. The restriction digests were analysed by electrophoresis on a 0.7% agarose gel (2.2.6). Figure 6.4 (a-d) shows the patterns obtained from the genomic DNA of the bacteria *B. linens*, *M. luteus*, *R. ruber* and *S. multivorum* digested with decreasing concentrations of *Sau*3AI. The amount of enzyme needed to produce the maximum intensity of fluorescence in the size range 9-23Kb was determined visually.

When large-scale digestions were performed, half the number of units determined in the small-scale reactions were used. This enabled the maximum number of molecules within the required size range to be obtained (Promega).

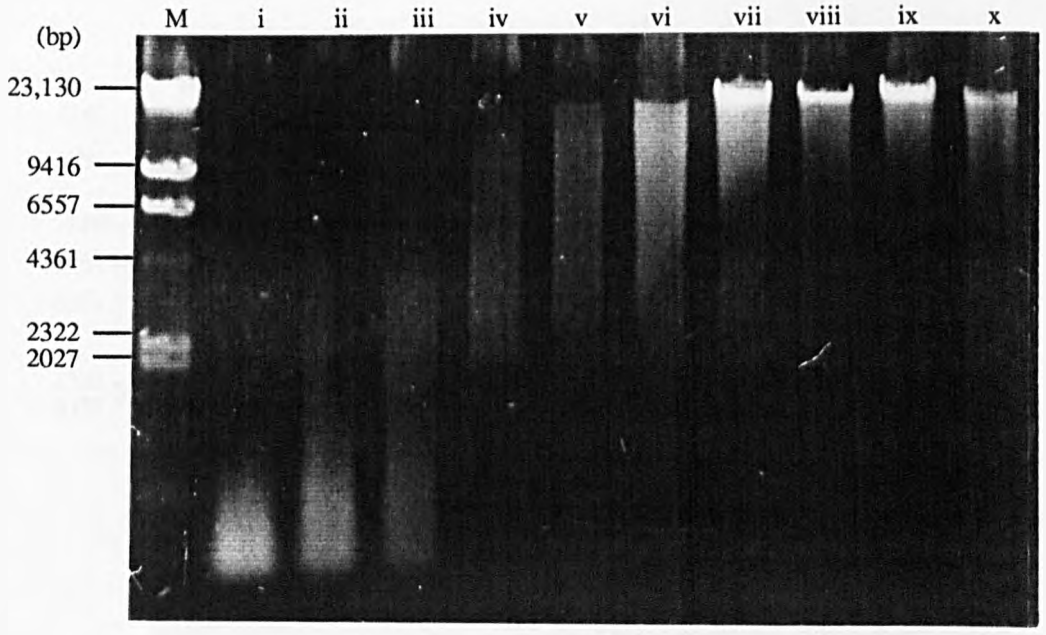
(ii) Large-scale preparation of genomic DNA

The following amounts of enzyme were used to digest partially 100µg of genomic DNA under the conditions established above

Bacterium	**U/µg**	**1/2U/µg**	**Total U/100µg**
B. linens	0.01	0.005	0.5
M. luteus	0.125	0.00625	0.625
R. ruber	0.015	0.0075	0.75
S. multivorum	0.01	0.005	0.5

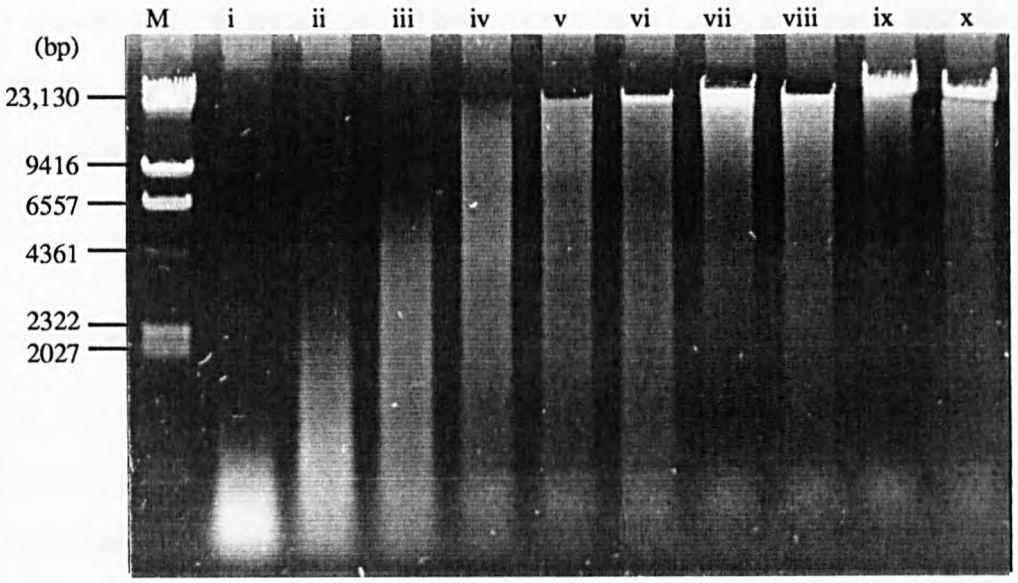


(a)

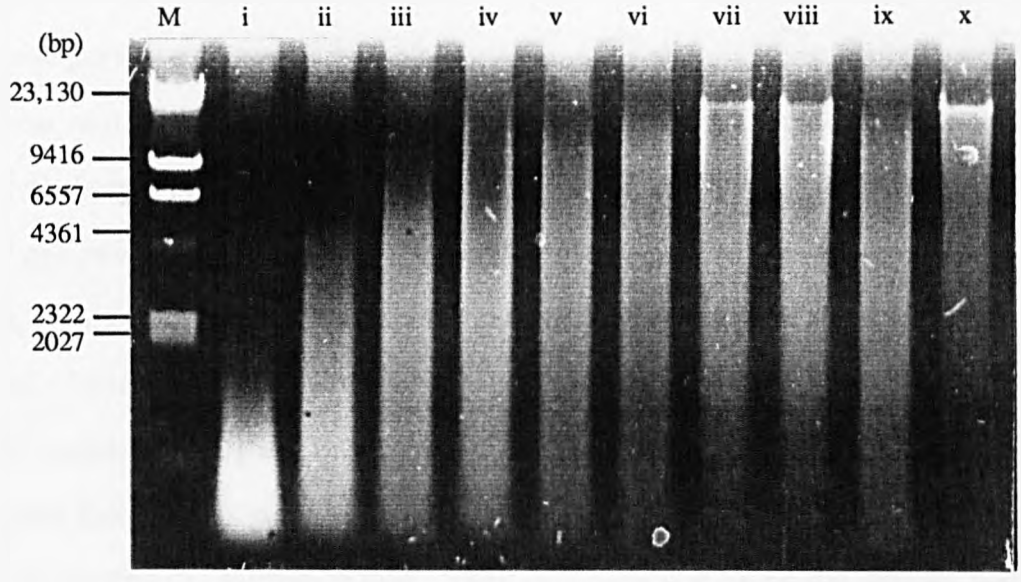


(b)

Figure 6.4 Agarose gel analysis of genomic DNA partially digested with *Sau*3AI. Restriction enzyme digests were performed as in 2.2.14 (i) on the genomic DNA of; (a) *B. linens*, (b) *M. luteus*, (c) *R. ruber* and (d) *S. multivorum* and analysed by agarose gel electrophoresis (0.8%). The units of enzyme used per μg of DNA were as follows; (i) 1, (ii) 0.1, (iii) 0.05, (iv) 0.025, (v) 0.015, (vi) 0.0125, (vii) 0.01, (viii) 0.0085, (ix) 0.005 and (x) 0.0035. M - λ *Hin*dIII markers (Promega).



(c)



(d)

Figure 6.4 (continued)

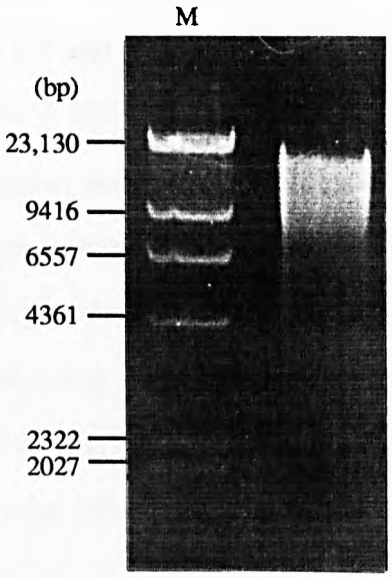
The DNA concentration, time of digestion and temperature were identical to those established in the small-scale reactions. After 30 minutes a small aliquot (0.5µg) was removed and analysed on a 0.8% agarose gel (2.2.6); the results are shown in Figure 6.5 (a) for *M. luteus*. If digestion was adequate the restriction enzyme fragments were precipitated with ethanol as in 2.2.4 and resuspended in 100µl of sterile H_2O.

(iii) Size fractionation of partially digested genomic DNA

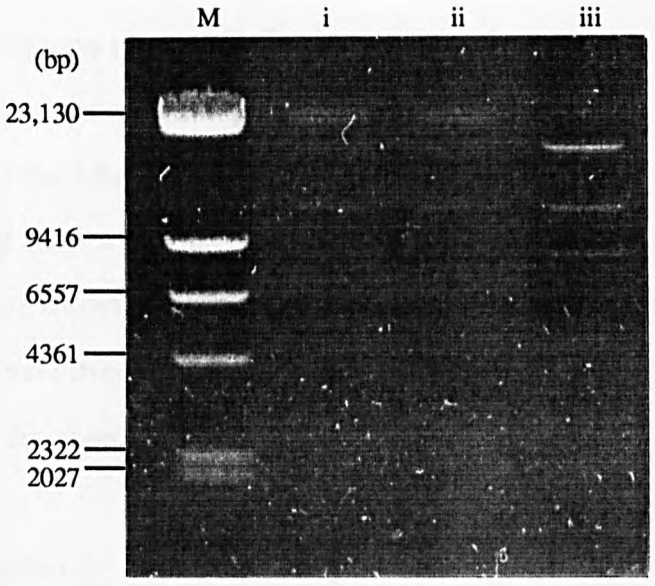
The restricted DNA was separated on an agarose gel (0.8%). Electrophoresis was carried out with a Biometra uv band elution system at 15V until an even smear of the DNA fragments was observed (usually overnight). This system enabled the DNA to be seen during electrophoresis by means of a uv transilluminator directly under the gel tray. When sufficient separation was observed electrophoresis was stopped. The gel was sliced across its entire width between the 9 and 23Kb bands of the λ *Hin*d III markers. The gel slice was removed and placed in a fresh tray where a new 0.8% agarose gel was cast around it with a slot-forming comb positioned approximately 2cm in front of it. The trapping slots (made by the slot-forming comb) were flushed with several changes of fresh electrophoresis buffer and then electrophoresis was restarted at 50V and monitored by means of the integral uv transilluminator. Electrophoresis was stopped when the DNA was just in front of the trapping slot. The slot was filled with 250µl of sterile focusing solution [10% (w/v) sucrose; 10% (v/v) 3M Na-acetate; 1 x TAE buffer, pH 8.0] and electrophoresis was resumed until the DNA entered the trapping slot. At this point, electrophoresis was terminated and the focusing solution was transferred to a microfuge tube. 500µl of ethanol was added and the tube was held at -20°C for 30 minutes before centrifugation (12000g, 10 min.). The DNA pellet was resuspended in sterile H_2O to a final concentration of 0.5mg/ml.

6.3 Ligation of genomic DNA to λ-GEM-11 *Bam*HI arms

To determine the optimal ratio of vector arms to genomic DNA in order to produce the greatest number of packagable molecules, test ligations were first



(a)



(b)

Figure 6.5 Agarose gel analysis of size fractionated, partially digested *M.luteus* genomic DNA and its ligation to λ-GEM-11 *Bam*HI arms.
(a) An aliquot of size fractionated genomic DNA was analysed by agarose gel electrophoresis (0.6%). (b) The size fractionated DNA was ligated to λ-GEM-11 *Bam*HI arms in a ratio of DNA:arms of (i) 3:1 and (ii) 1:1, (iii) λ-GEM-11 *Bam*HI arms. An aliquot of the ligation was analysed by agarose gel electrophoresis (0.4%). M - λ *Hin*dIII markers (Promega).

performed on *M. luteus*. The ligation conditions that were established were then used for the other three bacteria.

Ligation reactions were performed as in 2.2.14 (iii). The molar ratios of arms:insert tested were 1:1 and 1:3. The 16Kb plasmid pTI11 predigested with *Bam*HI, supplied with the λ GEM-11 *Bam*HI arms cloning system, was used as a positive control for the ligation reaction. The background levels of religated arms were also determined by carrying out the ligation reactions with no insert DNA. A sample (0.5µl, approximately 50ng) was taken out of the ligation reaction and analysed on a 0.4% agarose gel [Figure 6.5 (b)]. Both ligation conditions were deemed successful as high molecular weight DNA was observed in both lanes. Similar ligation results were also observed for the other bacteria when these ratios were used (the gel photograph is not shown).

6.4 Packaging and titration of recombinant λ-GEM-11 phage

All the ligation reactions were packaged as described in 2.2.14 (iv). Lambda c1857 *sam*7 DNA was used as a positive control for the packaging reaction. Bacterial plating cells for the strains LE392, KW251 and NM621 were prepared as described [2.2.14 (iv)] and used to titre the packaged recombinant phage library.

Table 6.1 shows the number of pfu/µg determined for the different ligation conditions and the different plating cells. The results of the ligations for *B. linens*, *R. ruber* and *S. multivorum* are also given in this table.

6.5 Amplification of the libraries

Because of time constraints only the *M. luteus* genomic library was screened initially and then it was amplified along with the others. All the recombinant phages were harvested by scraping off the top agar containing the phage and bacterial cells. The agar was suspended in phage buffer and chloroform was added to lyse the cells. The phage was allowed to diffuse into the buffer (1 hour, 4°C), and then the agar and cells were pelleted by centrifugation and removed. The supernatant, containing the

Table 6.1 Recombinant λ-GEM-11 phage titres given in pfu/μg. The host strain LE392 was used in all control ligations and packaging reactions.

HOST STRAIN;-	LE392		KW251		NM621	
INSERT:ARMS;-	3:1	1:1	3:1	1:1	3:1	1:1
LIBRARY;						
B. linens	$1.5x10^5$	$2x10^5$	$2.45x10^5$	$5x10^5$	$2x10^5$	$2.1x10^5$
M. luteus	$1x10^5$	$1.75x10^5$	$1.83x10^5$	$2.35x10^5$	$1.75x10^5$	$2.31x10^5$
R. ruber	$1.3x10^5$	$1.41x10^5$	$1.3x10^5$	$1.35x10^5$	$1.4x10^5$	$2.1x10^5$
S. multivorum	$2.8x10^5$	$3.1x10^5$	$1.9x10^5$	$3x10^5$	$2.65x10^5$	$3x10^5$

positive ligation control	$2.4x10^6$
λ-arms negative ligation control	17
negative packagene control	52
packagene control	$7.2x10^7$

phage, was stored at 4°C (after addition of a drop of chloroform) to produce a stable amplified library.

Representative libraries for the four bacteria *B. linens*, *M. luteus*, *R. ruber* and *S. multivorum* were constructed in the replacement vector λ-GEM-11 and packaged *in vitro* using the Promega Packagene system. The libraries contained in the order of 10^5 clones with an average insert between 9-23Kb.

6.6 Discussion

A number of crucial criteria must be satisfied during the construction of a fully representative genomic library. The first is to obtain genomic DNA of the highest possible molecular weight. The average genomic DNA fragment available for ligation needs to be in the size range 9-23Kb. This range is generated by partial digestion with *Sau*3AI. If genomic DNA is less than 50Kb in length the partial *Sau*3AI digestion will generate fragments with non-cohesive termini. These fragments will inhibit the ligation, by not generating concatenated phage DNA as only one λ arm will be able to ligate. As concatenated phage DNA is the usual substrate for packaging *in vitro* this reaction will also be inhibited. The best method discovered after evaluating commercial kits was a simple phenol/chloroform extraction after cell lysis and treatment with Proteinase K and RNase A. The phenol/chloroform was gently mixed in, the mixture centrifuged and then wide-bore pipette tips were used to transfer the aqueous phase. Although some loss did occur due to mechanical shearing in the extraction processes, these methods preserved the DNA better than previously. The DNA was spooled around a glass rod after ethanol precipitation to reduce the risk of shearing compared with that during centrifugation.

Once high molecular weight DNA was obtained and digested the DNA was size fractionated, so that the majority of fragments were in the range 9-23Kb. This was achieved by the use of electrophoresis equipment from Biometra. Other methods of size fractionation, e.g. sucrose density-gradient centrifugation, NaCl density-gradient centrifugation and other preparative gel electrophoresis techniques, were attempted but proved to be time consuming and gave poor yields. This technique was faster and although the yield was still low it was substantially higher than that obtained by the previous protocols. Other commercially available genomic cloning λ vectors are available which negate the need for size fractionation. The vector λ-GEM-11 is digested with *Xho*I and then the first two nucleotides of this site are filled in, before dephosphorylation. These arms are compatible for ligation with genomic DNA that has been partially digested with *Sau*3AI or *Mbo*I and the first two nucleotides filled in. Since insert-to-insert ligation is prevented, the only ligation products will be those of

single genomic insert to the arms. This procedure, however, does have limitations in that genomic fragments smaller than 9Kb and larger than 23Kb will also ligate to the arms, resulting in substrates that are unsuitable for *in vitro* packaging.

Commercially available genomic cloning packages are optimized to produce the highest number of recombinant phage molecules with the lowest non-recombinant background. This enables the production of larger libraries with less screening to filter out non-recombinants. *Spi* selection against stuffer fragments religating with λ arms is also unnecessary. Phage replication is inhibited in P2 lysogens where functional *red* and *gam* genes (present on the stuffer fragment) are present. This is termed sensitivity to P2 interference or *Spi* selection. Replacing the stuffer DNA with exogenous DNA allows replication to proceed. To use *Spi* selection *recBC*$^+$ hosts must be used and this may lead to the deletion of potential clones by recombination. By using the commercially available genomic cloning vector λ-GEM-11, *Bam*HI, and dephosphorylated arms, these potential problems are avoided.

The most efficient way to package a concatamerized λ molecule is by the use of *in vitro* packaging extracts. The packaging extract was extremely efficient as observed by the high titre of the positive control. The background due to endogenous phage within the packaging extracts was low. This indicates that most phage in the library is the result of ligations between genomic inserts and λ arms.

The ligation reaction was also tested with the necessary positive and negative controls. The ligation reaction was tested by using the positive control insert supplied with the arms. A high titre was observed indicating that the ligation conditions were adequate to produce recombinant phage. Background due to self ligation of the λ arms was low, ensuring that nearly all phage in the libraries contained an insert. If the background was high, the number of clones required to be fully representative of the genome would be increased. The titre of the positive ligation reaction is a thousand-fold higher than the titres of the bacterial genomic libraries. The lower titre observed could be due to a population of 'rag-ended' donor fragments, mentioned earlier, acting as chain stoppers in the ligation and packaging reactions.

Reactions were performed with two different ligation ratios, 3:1 and 2:1 (insert:arms). Both ligation ratios produced high titres with the ratio of 1:1 (insert: arms) producing slightly higher titres than 3:1.

With the exception of the titres for *R. ruber* the most accommodating bacterial host for the genomic libraries was KW251, although the differences in titres seen with the different hosts was marginal. However the slightly higher titres could be obtained because the recombinant phage is propagated in a *recD*$^-$ strain.

Taking into account the above, fully representative libraries were obtained for the four bacterial species *B. linens*, *M. luteus*, *R. ruber* and *S. multivorum*. The library of *M. luteus* was screened using *crtI* as a probe but no positive duplicate signals were observed. The libraries were amplified and are available for future screening by laboratory members and academic/industrial collaborators.

Genomic libraries were constructed to isolate carotenoid biosynthesis genes as this method was perceived to hold numerous advantages over other methods of isolation. Screening on the basis of nucleic acid hybridization (once accepted as a feasible approach, see 5.3) facilitates the screening of large numbers of recombinant clones simultaneously. The sequence target does not have to be entire (once isolated it can be used as a homologous probe to isolate the entire sequence), nor does it have to express a biological product, as is the case with screening expression libraries with antibodies. Larger numbers of recombinant clones have to be screened with antibodies to account for incorrect reading frame and orientation of genes. This is not applicable with nucleic acid screening and so consequently smaller numbers of recombinants can be screened. Although this implies that screening expression libraries with antibodies is more difficult, *crt* genes have been successfully isolated from *Capsicum annuum* by the use of antibodies (Hugueney *et al.*, 1992; Bouvier *et al.*, 1994). However the cDNA library was constructed from RNA expressed in the early stages of fruit ripening which is when *crt* RNA transcripts are prevalent. This would enrich the library with *crt* sequences, and so screening with antibodies would prove more worthwhile. The same

workers have also isolated phytoene synthase from *C. annuum* by using a tomato cDNA as a heterologous probe (Romer *et al.*, 1993), after failing to isolate this gene using their antibody.

Many other *crt* genes have been isolated from libraries by the use of heterologous probes as described in Chapter 1. As the libraries constructed were considered to be fully representative, and hybridization between probe and possible targets has been assessed as viable, it should not be too long in the future before new *crt* genes are isolated.

CHAPTER 7

SCREENING A GENOMIC LIBRARY FROM *RHODOPSEUDOMONAS ACIDOPHILA*

7.1 General introduction

Rhodopseudomonas acidophila is a purple non-sulphur photosynthetic bacterium. It was first characterized in 1969 and was named because of its ability to grow optimally at pH 5.2 (Pfennig, 1969). The main carotenoid pigment of *R. acidophila* is rhodopin which occurs together with its glucoside derivative (Appendix I).

The photosynthetic unit of purple, non-sulphur, phototrophic bacteria, is made up of pigment-protein complexes that are located on the intracytoplasmic membrane. Two types of antenna complex are present: those that are intimately associated with the photochemical reaction centre and the peripheral light-harvesting complex. The antenna complex consist of bacteriochlorophyll *a* (bchl *a*) and carotenoids that are non-covalently bound to low molecular weight polypeptides. The complexes are identified by the absorption maxima of bchl *a*.

The light-harvesting pigment-protein complexes of three wild-type strains of *R. acidophila* have been studied in detail. Four different types of antenna complexes (B880, B800-830 and two types of B800-850) were isolated, purified and studied. The bchl *a* and carotenoids were bound to low molecular weight peptides within the complexes (Cogdell *et al.*, 1983). The amino acid sequence of the antenna polypeptides showed areas of homology to corresponding sequences from the bacteria *R. sphaeroides* and *R. capsulatus* (Bissig *et al.*, 1988). The efficiency of energy transfer between carotenoid and bacteriochlorophyll *a* was independent of temperature and varied depending upon the type of antenna complex (Angerhofer *et al.*, 1986). Decreased light intensity resulted in the B800-820 complex predominating over B800-850 in strains 7050 and 7750. The carotenoid content of strain 7050 also altered according to light intensity. Strain 10050 was found to contain only one type of

complex (B800-850) that did not vary according to light intensity (Gardiner *et al.*, 1993). The crystal structure of the light-harvesting antenna complex from strain 10050 has been elucidated and studied in detail (McDermott *et al.*, 1995).

This chapter describes the screening of the genomic library of strain 10050 (kindly provided by R. Cogdell) for the carotenoid biosynthesis genes responsible for rhodopin production.

7.2 Design of a suitable probe

(i) Degenerate PCR

Degenerate PCR was used in attempts to obtain a homologous probe to screen the *R. acidophila* library. PCR using degenerate primers has been used in the cloning of uncharacterized genes belonging to known families (Lee *et al.*, 1988; Girgis *et al.*, 1988; Mack and Sninsky, 1988). A number of factors have to be considered in the design of degenerate primers. The first is the degeneracy of the genetic code. Whilst the amino acids methionine and tryptophan are encoded by a single codon, others, for example arginine and leucine, can be encoded by four or six codons respectively. Areas of sequence containing such amino acids were avoided if possible. Single base mismatches can stop extension and so degeneracy at the 3' end is best avoided. Once PCR products are analysed by sequencing they can be used as probes in hybridization reactions against fully representative libraries.

Sequences of a number of bacterial *crtI* genes show conserved areas. Degenerate primers were designed based on the areas of sequence conservation of CrtI of the bacteria *E. herbicola*, *E. uredovora*, *R. capsulatus*, *R. sphaeroides* and *M. xanthus*. These sequences were aligned by use of the Pileup program of the GCG software package. Figure 7.1 shows regions of the alignments that were highly conserved. The degenerate primers are shown in Figure 7.1 below the chosen conserved areas with their relative positions along the *crtI* gene. The restriction enzyme sites *Eco*RI and *Xba*I were incorporated into the 5' and 3' primers, respectively. This facilitates the directional cloning of any product of the PCR into the vector pUC19 (Appendix II). These sites were not present within any previously cloned *crtI* gene and so were less likely to be present in *R. acidophila crtI*.

The success of degenerate PCR is dependent on a compromise between specificity and efficiency of the primers. When the PCR was started, low non-stringent annealing temperatures were used to allow the correct primer sequence to anneal. Once a few cycles were completed the annealing temperature was increased based on the annealing temperature of the entire primer duplex (Dumas Milne Edwards *et al.*, 1995).

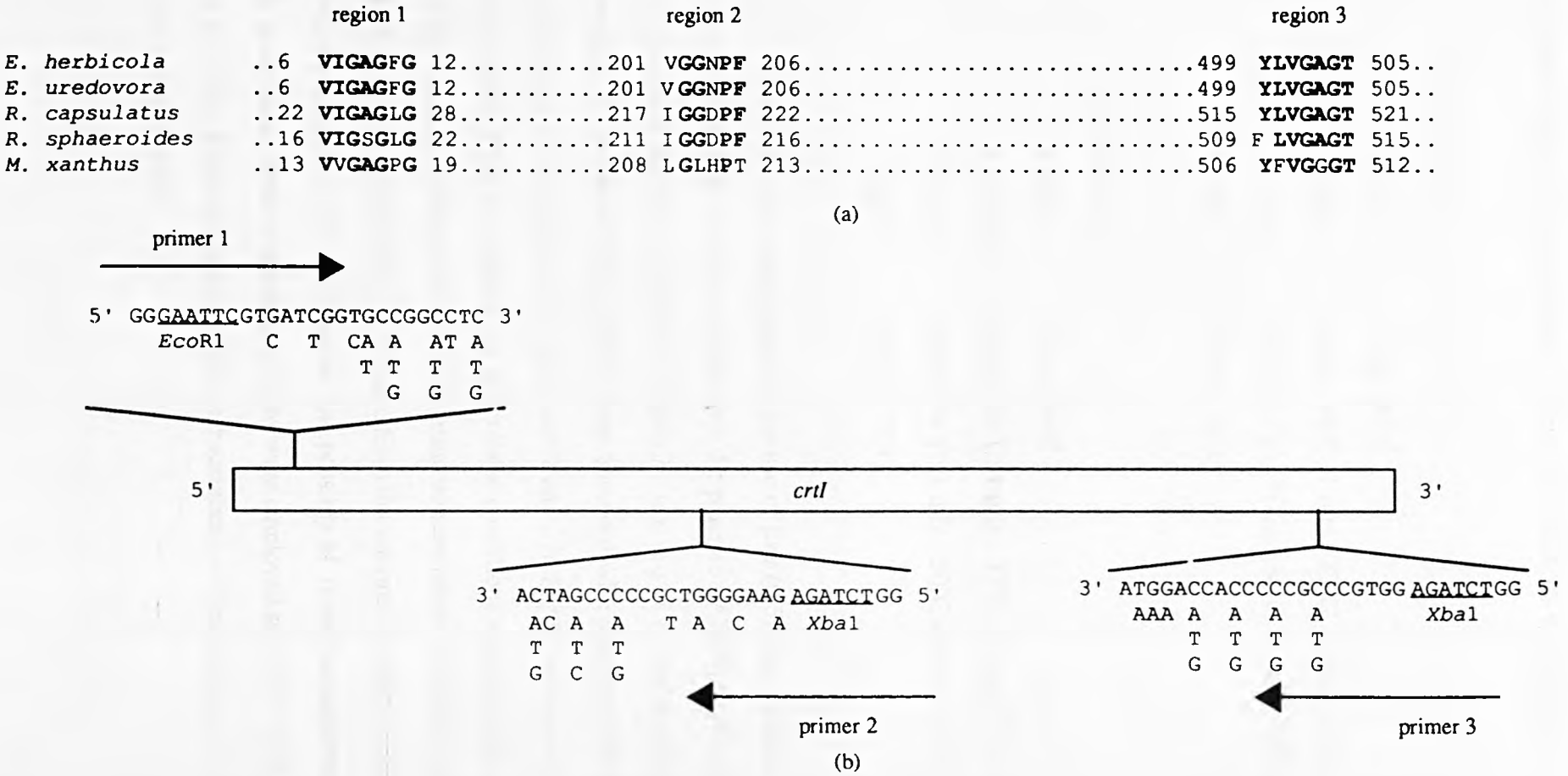


Figure 7.1 Alignment of CrtI sequences and degenerate primers designed from the alignment.
(a) Crt I protein sequences were aligned by use of the Pileup program of the GCG software package. Conserved residues are given in bold. (b) From the three conserved regions indicated, degenerate primers were designed.

A slow ramp time between the annealing and extension temperature was also used to aid in the specific primer annealing. Two PCR conditions were chosen:

PCR Program A;

1 cycle;	3 min., 94°C
5 cycles;	30 sec., 94°C; 1 min., 37°C; 1 min., 72°C (4.2 ramp)
25 cycles;	30 sec., 94°C; 30 sec., 50°C; 1 min., 72°C (1 ramp)
1 cycle;	10 min., 72°C

PCR Program B;

1 cycle;	3 min., 94°C
5 cycles;	30 sec., 94°C; 1 min., 37°C; 2 min., 72°C (4 ramp)
30 cycles;	30 sec., 94°C; 1 min., 50°C; 2 min., 72°C (1 ramp)
1 cycle;	10 min., 72°C

The PCR was carried out with the use of low (0.5mM), medium (1.5mM) and high (2.5mM) Mg^{2+} concentrations with the plasmid pAPU211 (100ng) as template. The results of the PCR are given in Figure 7.2 (a). No PCR products were visible with program A. Putative PCR products were observed with program B in lanes ii and iii with primer combinations of 1 and 3 and 1 and 2 at Mg2+ concentrations of 1.5mM and 2.5mM. PCR was carried out under these conditions with bacterial genomic DNA as the template (100ng), ten times the usual concentration of primer and dNTPs were used, the extra volume being subtracted from the volume of water. Products were seen with the templates of *S. multivorum* and possibly *M. luteus* but unfortunately not with *R. acidophila*. Other reaction conditions were employed such as touchdown PCR (Don *et al.*, 1991; Kidd and Ruano, 1995) and variations of the above cycles but no product was ever observed.

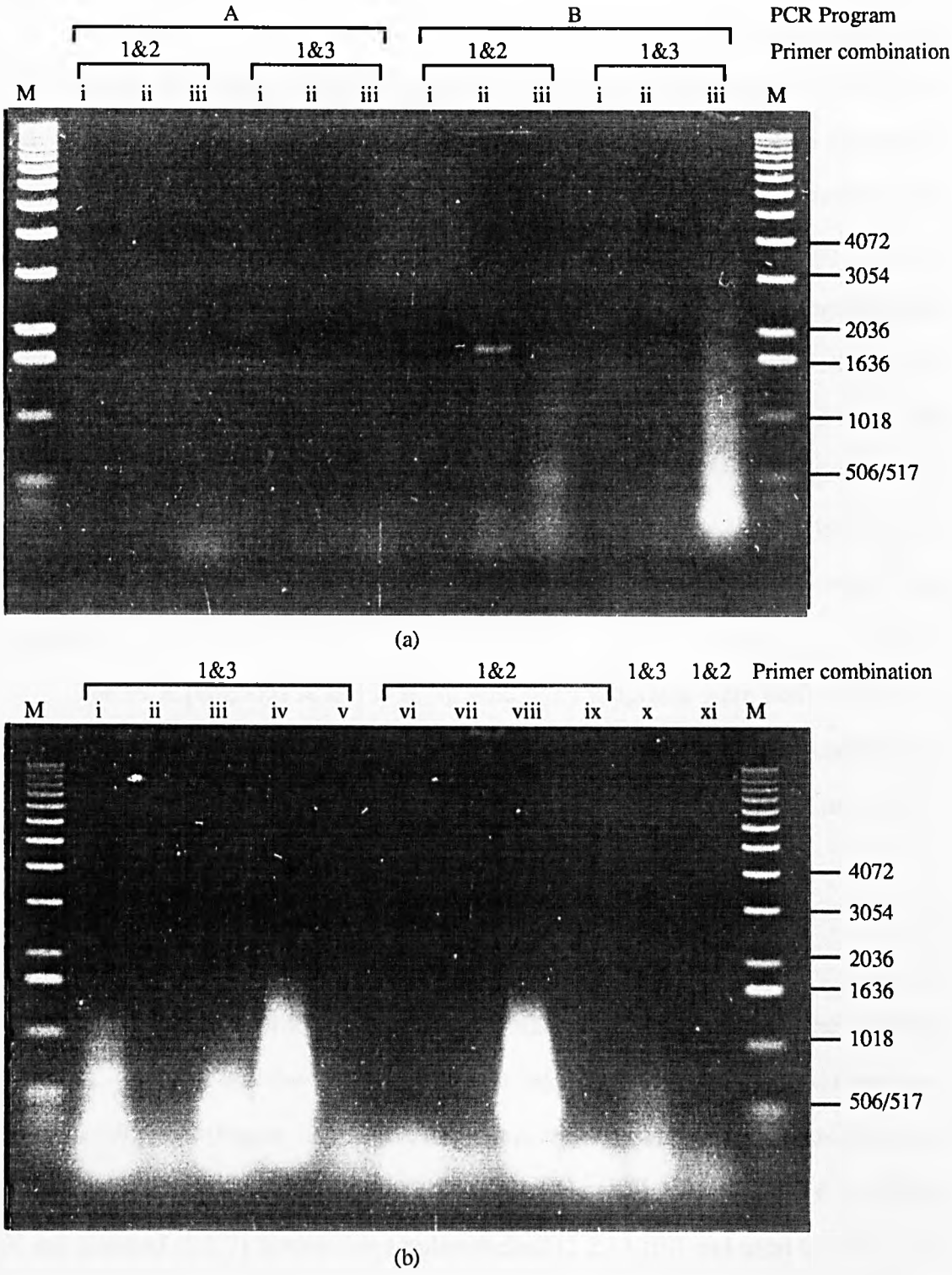


Figure 7.2 Agarose gel analysis of PCR products obtained from degenerate primers. (a) Agarose gel (0.8%) analysis of PCR products obtained from the degenerate primers shown in Figure 7.1 and pAPU211 as the template. PCR was carried out as described in 7.2 (i), using PCR buffer with a final Mg^{2+} concentration of (i) 0.5mM, (ii) 1.5mM and (iii) 2.5mM. (b) Agarose gel (0.8%) analysis of PCR products using the degenerate primers shown in Figure 7.1 and the following templates; (i) pAPU211, (ii) and (vi) *B. linens*, (iii) and (vii) *M. luteus*, (iv) and (viii) *S. multivorum*, (v) and (ix) *R. ruber* and (x) and (xi) *R. acidophila*. PCR program B [7.2 (i)] was used with a final Mg^{2+} concentration of 1.5mM. M - 1 kilobase ladder (Gibco-BRL).

(ii) Redesigned degenerate primers

New primers were designed to improve the results already obtained (Figure 7.2). The CrtI sequences of the four bacteria *E. herbicola*, *E. uredovora*, *R. capsulatus* and *R. sphaeroides* were aligned as previously [7.2.1 (i)]. Alternative conserved regions were chosen. Region three (Figure 7.3) showed a highly conserved area between all four bacteria. The new primers were primarily based on the areas of conservation between the two photosynthetic bacteria which would be expected to be most closely related to *R. acidophila*. In areas where differences in amino acid sequence were seen between non-photosynthetic and photosynthetic bacterium, the photosynthetic bacteria showed amino acid conservation with each other as did the non-photosynthetic sequences. The primers could therefore be designed with a lower level of degeneracy, so that specificity to *crtI* of *R. acidophila* would be expected to be increased.

The PCR programs A and B along with other programs were performed with the new primers (Figure 7.3). The templates and variations in magnesium ion concentration were the same as previously. There was no improvement in the results obtained previously.

(iii) Other possible probes

Two alternative probes to screen the library were obtained from the plasmid pAPU211. The first was the 1Kb fragment of a *Sma*I digest containing portions of the *crtI* and *crtB* genes (Figure 7.4). DNA fragments of the *Sma*I digest were analysed by agarose gel electrophoresis (2.2.6). The 1Kb fragment was then recovered from the gel and purified (2.2.7) before being radiolabelled [2.2.13 (ii)] and used to probe the library [2.2.14 (v)].

The other probe used in screening was a PCR product encompassing a 300bp central region of *crtI* (Figure 7.4). The primers and reaction conditions for the PCR are described in 5.4.2 (ii). The PCR product was purified and radiolabelled as above before being used as a probe. Neither probe obtained from pAPU211 produced any duplicated positive signals.

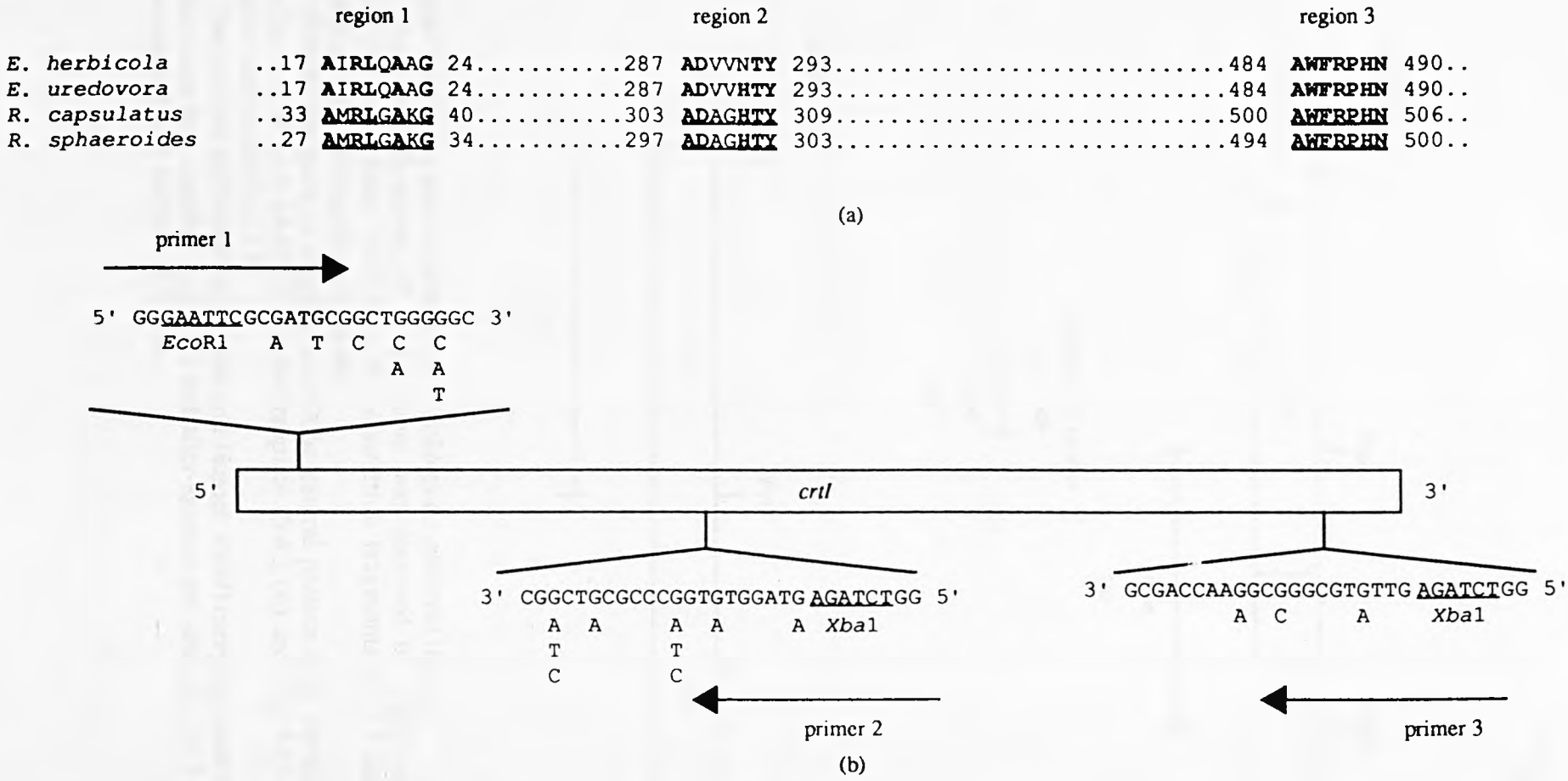


Figure 7.3 Alignment of CrtI sequences and new degenerate primers designed from the alignment.
(a) CrtI sequences were aligned as previously (Fig. 7.1). (b) Different conserved regions were chosen and degenerate primers were designed based on the conserved domains of *R. capsulatus* and *R. sphaeroides* (underlined).

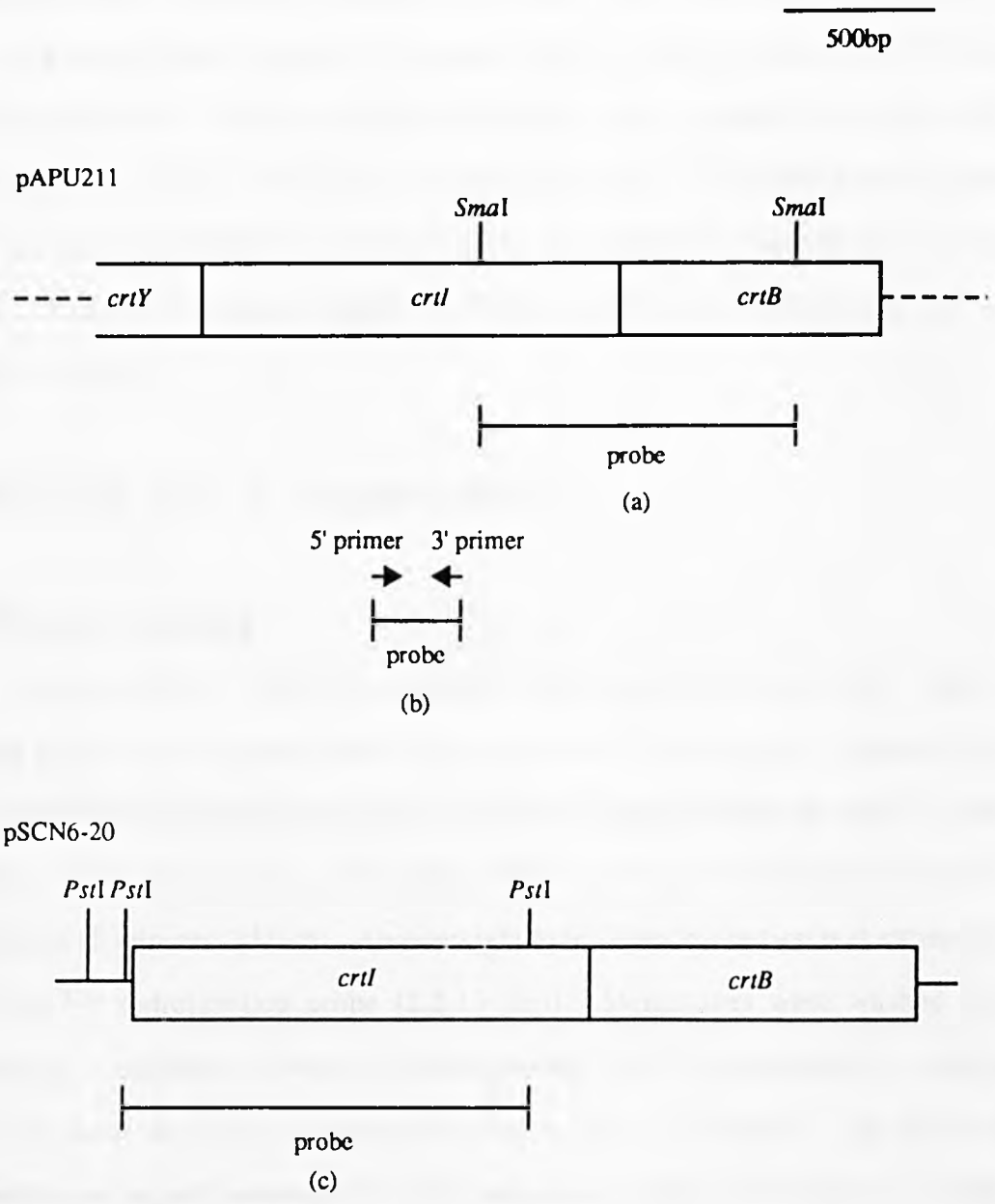


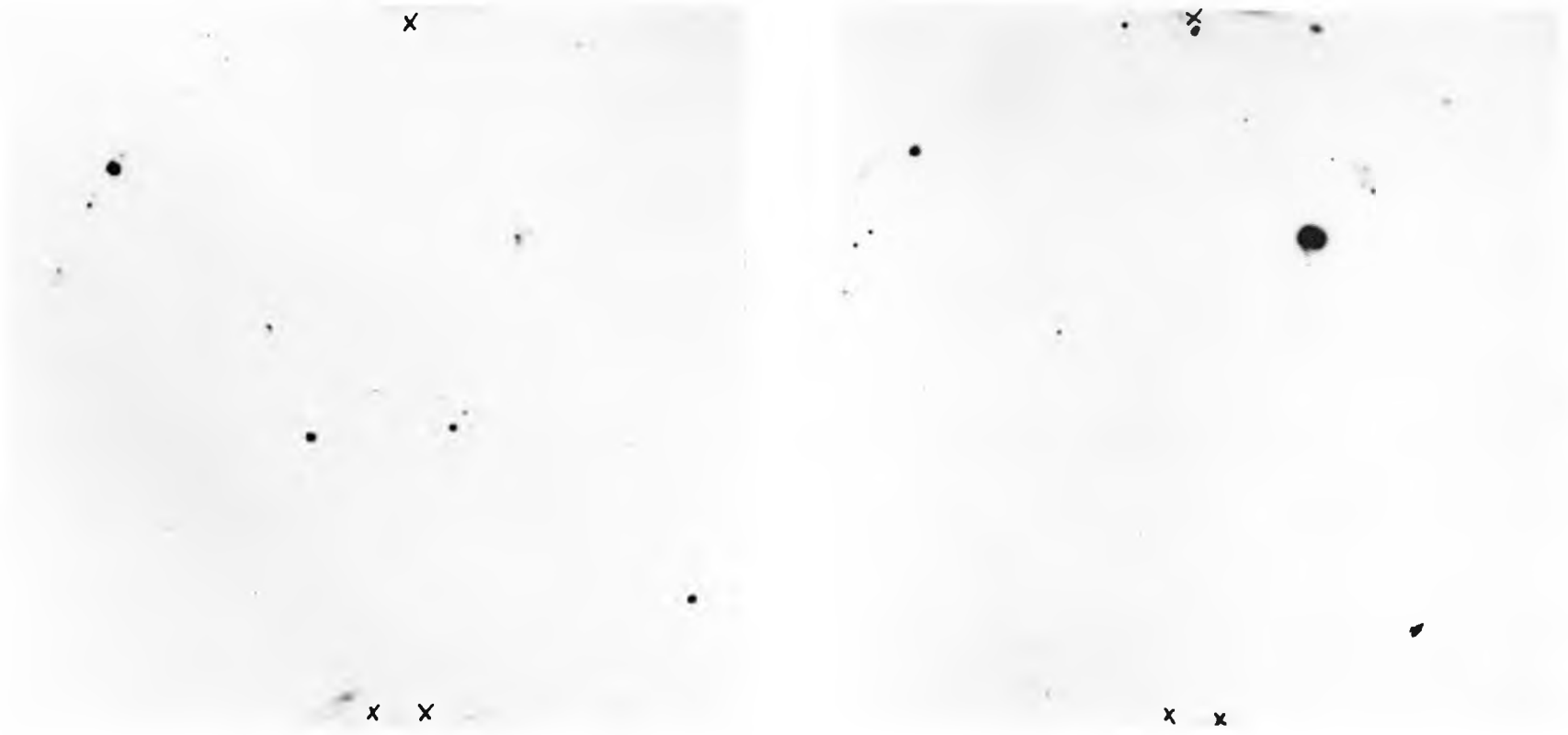
Figure 7.4 Probes used to screen the *R. acidophila* genomic library.
(a) The *crtI/B* fragment of *E. herbicola* was obtained by digesting the plasmid pAPU211 with *Sma*I, analysing the restriction fragments on an agarose gel and purifying the 1Kb fragment (2.2.6-8).
(b) Primers designed to amplify the 300bp central protion of *E. herbicola crtI* were used in a PCR with pAPU211 as the template (5.4.2 (ii) and Fig. 5.6). The resulting product was purified (2.2.7).
(c) The plasmid pSCN6-20 (Coomber and Hunter, 1989) carrying some *crt* genes of *R. sphaeroides* was digested with *Pst*I and after agarose gel analysis the 1.4Kb fragment containg *crtI* was purified (2.2.6-8).

The third option was to use digested fragments of the plasmid pSCN6-20 which contained some *crt* genes of *R. acidophila* (Coomber and Hunter, 1989). Nearly all of the *crtI* gene could be obtained by digestion with *Pst*I. The plasmid pSCN6-20 was digested with *Pst*I and the resulting fragments were separated by agarose gel electrophoresis (2.2.6). DNA bands corresponding to the *crtI* fragment were excised from the gel and purified (2.2.7) before being radiolabelled for subsequent use as a probe. Positive duplicated signals were observed by use of this probe and are described below.

7.3 Screening of the *R. acidophila* library

(i) Primary screening

Approximately 100,000 recombinant phage were plated out using NM621 plating cells on a 90mm round petri dish as described in 2.2.14 (iv). Duplicate lifts were made from the plate and orientation marks were made on both the membrane and the dish. After denaturation of the phage, DNA was fixed to the membranes by uv irradiation (5 minutes, 254nm). An overnight hybridization reaction was carried out with the ^{32}P radiolabelled probe [2.2.13 (iii)]. Membranes were washed with increasing stringency to remove unincorporated label and non-specific signals. Between each wash the membranes were exposed overnight. By checking hybridization signal intensity after each stringency wash, putative positives were identified. Signals observed after the initial high salt wash either faded or were still present as the stringency was increased. Signals that faded were disregarded. Signals still present after the high stringency wash were assumed to be the result of high specificity binding between immobilized DNA and probe. Figure 7.5 shows the duplicate membranes after washing at high stringency. Two signals were observed that were duplicated and were still present after this wash. As it was impossible to pick a single plaque responsible for the signal, the area surrounding was taken and subjected to a further round of screening.



(a) (b)

Figure 7.5 Primary screening of *R. acidophila* genomic library with the *crtI* gene from *R. sphaeroides*.
Approximately 10^5 recombinant phage were plated on 90mm dishes and plaque lifts were performed in duplicate (a) first lift, (b) second lift as in 2.2.14 (v). The membranes were subsequently hybridised with the probe for *crtI* (2.2.13) and then washed as follows; 2 x 10 min, 2 x SSPE, 50°C; 2 x 10 min, 1 x SSPE, 50°C; 1 x 10 min, 0.1 x SSPE, 65°C, before exposure to X-ray film.

(ii) Secondary screening

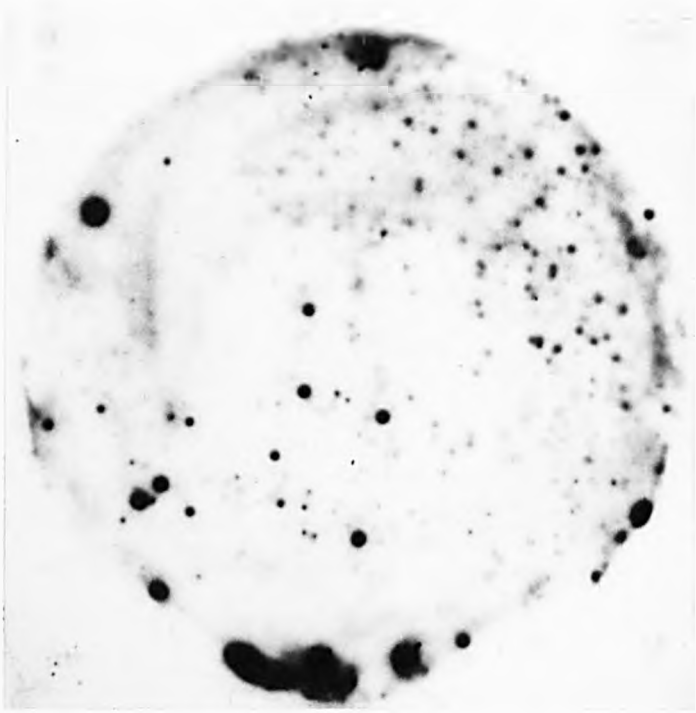
Phage picked from the primary screen were allowed to diffuse into SM buffer before being plated out at a density of approximately 1000 per 60mm dish [2.2.14 (vi)]. Hybridization experiments with the *R. sphaeroides crtI* fragment were performed exactly as previously. Of the two areas picked, only one still gave positive signals that were duplicated and still present after high stringency washing. The other primary signal was possibly due to non-specific hybridization which coincidentally gave duplicate signals. Figure 7.6 shows the membrane after the high stringency wash. Radiolabelling of the probe was not as efficient as before as indicated by the intense dark circles corresponding to unincorporated label. Positive signals that were duplicated and present at the high stringency wash were aligned to individual plaques. Six of these plaques were picked as before to undergo a further round of screening.

(iii) Tertiary screening

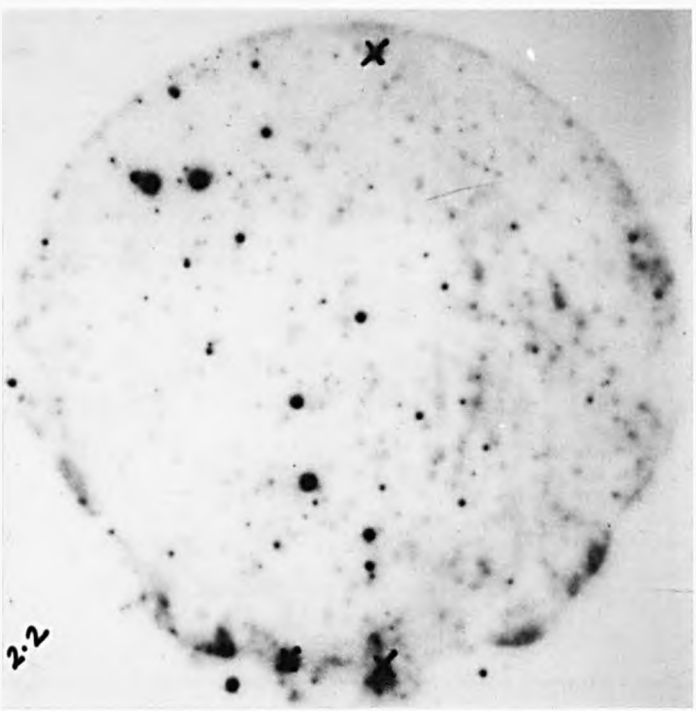
After the picked phage were allowed to diffuse into SM buffer they were plated out at a density of approximately 100 per 60mm petri dish. Hybridisation and membrane washing was performed as before. Figure 7.7 shows the duplicate membranes of one putative positive of the six isolated from secondary screening. All signals were duplicated and were aligned to individual plaques. The other five putative positives picked from secondary screening gave similar results (data not shown). A phage DNA stock was made from one plaque on each of the six plates [2.2.13 (v)].

7.4 Characterization of putative positive isolates

The putative positive isolates designated RA1-6 were passed on to other lab members for further characterization.

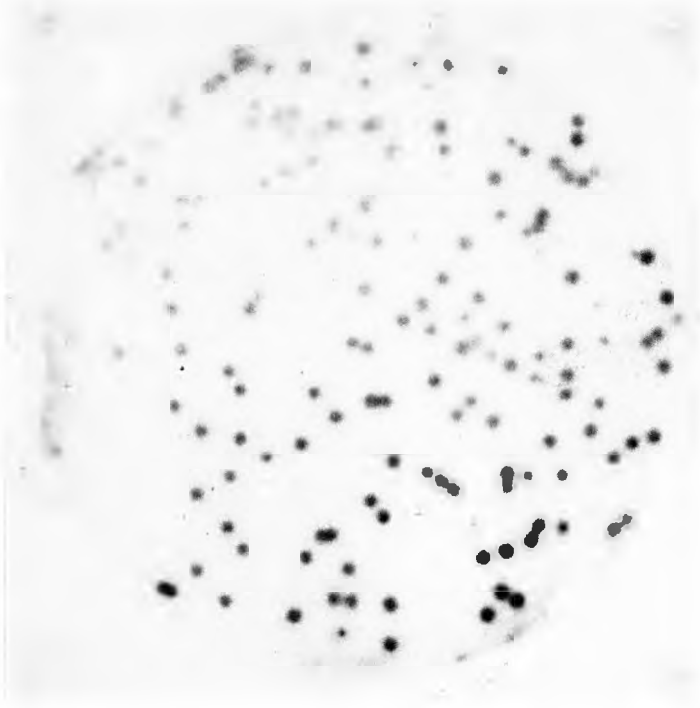


(a)

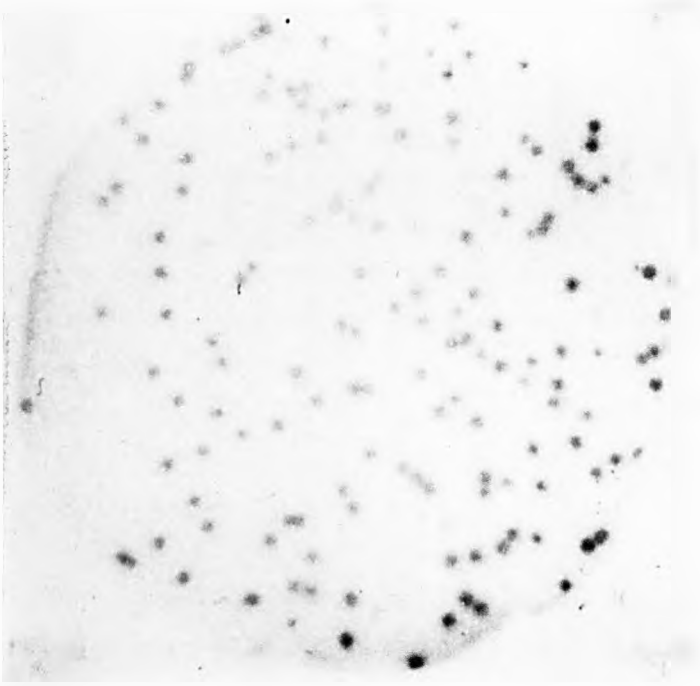


(b)

Figure 7.6 Secondary screen of *R. acidophila* genomic library.
Putative positives selected from the primary hybridisation were screened in a secondary hybridisation experiment (2.2.14 (vi)). Approximately 10^3 recombinant phage were plated on 60mm dishes and plaque lifts, hybridisation with *crtI* and subsequent washes were performed as previously (Fig. 7.5).



(a)



(b)

Figure 7.7 Tertiary screen of *R. acidophila* genomic library.
Putative positives selected from the secondary hybridisation were screened in a tertiary hybridisation experiment (2.2.14 (vi)). Approximately 100 recombinant phage were plated on 60mm dishes and plaque lifts, hybridisation with *crtI* and subsequent washes were performed as previously (Fig. 7.5).

7.5 Discussion

Various techniques have been employed in the isolation of *crt* genes. These have included the use of antibodies in screening expression libraries, heterologous complementation assays, transposon mutagenesis and screening genomic and cDNA libraries with heterologous probes.

Degenerate PCR is another adaptable technique, that could be very useful in the search for new carotenoid genes. However experience suggests it will take some time to optimize conditions.

Because the bacterial CrtI sequences show areas of sequence homology, designing degenerate primers from these areas seemed a reasonable approach in the search for CrtI of *R. acidophila*. As mentioned earlier, for successful PCR using degenerate primers a compromise needs to be reached between the specificity of the primers and the efficiency of the amplification reaction (Compton, 1990). In practice this means that the primers must be degenerate enough to encompass the unknown sequence, but not so degenerate that actual primer annealing and subsequent extension efficiency is compromised to the point that the reaction is not feasible.

The primers were designed from areas of sequence conservation between the bacteria, to give the best possible representation of the chosen domains. Primer 1, in both designs, includes some of the highly conserved βαβ structure characteristic of CrtI enzymes, thought to be involved in binding FAD or NAD(P) cofactors (Armstrong *et al.*, 1993). The initial set of primers was designed with the help of Dr. H. Lang and Prof. C. N. Hunter (University of Sheffield) and were degenerate at the 3' end. In retrospect, this was not the best strategy as chain termination can occur with degeneracy at the 3' end. This could explain why no specific products were observed with the *R. acidophila* template. Products seen with the plasmid pAPU211 are probably due to the presence of a higher copy number of template. Overall the first set of primers was substantially more degenerate than the second set. Decreasing the degeneracy should allow the specificity of the primers to increase, but may also reduce the chances of encompassing the specific primer sequence. By separating photosynthetic and non-photosynthetic bacteria into 'sub-classes', areas of minimum ambiguity were observed,

implying decreased degeneracy. The second set of primers was also not degenerate at the 3' end, however the results of the PCR were no different from those obtained with the first set.

Improvements can be made to generate successful PCR with degenerate primers. The number of amino acids used in the design of degenerate primers could be reduced to just two or three plus a restriction enzyme site. The corresponding DNA sequence would be six or nine bases long with an extra six nucleotides that are not degenerate. This would increase the efficiency of amplification as the stability of the priming duplex would be vastly increased, compared to numerous duplexes obtained with longer and subsequently more degenerate sequences of primer. Sequences of amino acids used at the carboxyl end of the protein could incorporate stop codons, thus helping to reduce the degeneracy. Deoxyriboinosine is a nucleotide that will pair with all four bases, and has been successfully used, within degenerate primers to amplify rare sequences in PCR's (Knoth *et al.*, 1988), and in sequencing products obtained with highly degenerate primers (Shen *et al.*, 1993). By using this base in replacement of the others the degeneracy would improve dramatically, but care should be taken as the sequence of the product would be altered (Martin and Castro, 1985). Codon usage tables (Wada *et al.*, 1992) could be used to determine specific codon bias of different species. In the same vain an intercodon dinucleotide frequencies table could also be used as studies have shown that these are not random (Smith *et al.*, 1983).

No specific product was obtained with either set of primers under the PCR conditions described. Further empirical testing of the primers involving alterations in the degree of degeneracy and size of the primers and taking into account the above could prove successful in future applications.

Other sequences were available with which to screen the library and are depicted in Figure 7.4. Each probe was used to screen the library. No positive hybridization signals were obtained with the two probes obtained from *E. herbicola*. This could be because of the differences between sequences in photosynthetic and non-photosynthetic organisms. The choice of probe with which to screen non-homologous libraries with involves a compromise between length and specificity. If it is too long, the chances of

hybridization are decreased as more regions will not be conserved. Shorter sequences may not encompass enough conservation and so hybridization efficiency is also impaired. As the *crtI* sequence was available from *R. sphaeroides*, this was used as an alternative and produced duplicated positive results (Figure 7.5).

The primary positive signals were very faint but, as they were duplicated, screening was taken through to another round. The signals were stronger on tertiary screening due to the larger plaque size, as growth of these plaques was not inhibited by close contact with other plaques. Every plaque corresponded to a hybridization signal in tertiary screening. It is quite possible that the *crtI* gene of *R. acidophila* has been isolated. To verify this further characterization needs to be performed.

Time was not available to characterise the positive signal obtained and this was passed on for further study. Initially Southern blotting experiments need to be performed, using exactly the same conditions as with screening the library, to identify which region/s of the isolate are responsible for hybridization with the probe. The putative positive regions then need to be restriction mapped. Eventually the putative positive needs to be sequenced, a process which would involve subcloning. Expression studies would follow, enabling further understanding of carotenoid biosynthesis in bacteria. The *crtI* probe of *R. acidophila* could also be used to screen other bacterial genomic libraries for *crt* sequences.

CHAPTER 8

SUMMARY

The work detailed in this thesis has demonstrated the difficulties associated with the overproduction of carotenoids above a given level. Use of the powerful T7-RNA polymerase dependent transcription system, while increasing RNA levels did not result in significant increases in enzyme or carotenoid levels, suggesting effective post-transcriptional controls are present.

Localization experiments suggest that the enzymes are cytoplasmic, but do not preclude a transient membrane association/complex formation.

Antibody was generated against a region of CrtI and this allowed the identification of CrtI in whole cell lysates. A monospecific monoclonal antibody will probably be required to enable *in situ* immunocytochemical analysis of the precise cellular location of the Crt proteins.

Southern hybridization revealed multiple bands hybridising to *crt* genes in the four carotenoid producing strains *B. linens*, *M. luteus*, *R. ruber* and *S. multivorum*, suggesting the presence of a number of *crt* genes, identified due to homology at the DNA level between *crt* genes.

The carotenoids produced by the four bacteria *B. linens*, *M. luteus*, *R. ruber* and *S. multivorum*, were extracted and analysed. *R. ruber* was found to produce the carotenoid dihydroxyisorenieratene rather than canthaxanthin as previously suggested. The isomers of sarcinaxanthin, from *M. luteus*, were shown not to be configarational, and implied differing positions of the hydroxyl group in each isomer. The carotenoids dihydroxyisorenieratene and zeaxanthin biosynthesized by *B. linens* and *S. multivorum* were verified.

Fully representative genomic libraries were generated for future use in identifying carotenoid genes from the four bacteria.

Lastly, although a degenerate PCR approach was unsuccessful in identifying *R. acidophila crt* genes, a *crtI* probe from *R. sphaeroides* was successfully used to

identify putative plaques from the *R. acidophila* phage library. This is currently being further characterized and the results are awaited with interest.

FINAL COMMENTS

From experiences gained throughout this work a multidisciplinary approach is best employed in the search for and manipulation of new carotenoid biosynthesis genes. A combination of techniques would prove far more successful than the optimization of individual methods.

Applications of techniques already in practice in the isolation of other non-*crt* genes could also be used. For example a cDNA library could be constructed from a population of RNA molecules known to be synthesized under light growth conditions. Degenerate primers could be used as primers for the reverse transcription reaction, enriching the library again for *crt* genes. Bands of digested DNA showing hybridization with *crt* probes could be excised from agarose gels and cloned, resulting in an enriched environment to search for *crt* genes. Transposon mutagenesis, resulting in the loss of pigment production, is also another technique which has been successful in isolating *crt* genes.

REFERENCES

Acevedo, P. and J.S. Bertram (1995) Liarozole potentiates the cancer chemopreventive activity of and the up-regulation of gap junctional communication and connexin43 expression by retinoic acid and β-carotene in 10T1/2 cells. Carcinogenesis **16**, 2215-2222.

Albrecht, M., A. Klein, P. Hugueney, G. Sandmann and M. Kuntz (1995) Molecular cloning and functional expression in *E. coli* of a novel plant enzyme mediating ζ-carotene desaturation. FEBS Letts. **372**, 199-202.

Albrecht, M. and G. Sandmann (1994) Light-stimulated carotenoid biosynthesis during transformation of maize etioplasts is regulated by increased activity of isopentenyl pyrophosphate isomerase. Plant Physiol. **105**, 529-534.

Angerhofer, A., R.J. Cogdell and M.F. Hipkins (1986) A spectral characterization of the light harvesting pigment-protein complexes from *Rhodopseudomonas acidophila*. Biochim. Biophys. Acta **848**, 333-341.

Aracri, B., G.E. Bartley, P.A. Scolnik and G. Giuliano (1994) Sequence of the phytoene desaturase locus of tomato. Plant Physiol. **106**, 789.

Aragon, C.M.G., F.J. Murillo, M.D. de la Guardia and E. Cerda-Olmedo (1976) An enzyme complex for the dehydrogenation of phytoene in *Phycomyces*. Eur. J. Biochem. **63**, 71-75.

Armstrong, G.A. (1994) Eubacteria show their true colours: Genetics of carotenoid pigment biosynthesis from microbes to plants. J. Bacteriol. **176**, 4795-4802.

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

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

Armstrong, G.A., D.N. Cook, D. Ma, M. Alberti, D.H. Burke and J.E. Hearst (1993) Regulation of carotenoid and bacteriochlorophyll biosynthesis genes and identification of an evolutionary conserved gene required for bacteriochlorophyll accumulation. J. Gen. Microbiol. **139**, 897-906.

Armstrong, G.A., B.S. Hundle and J.E. Hearst (1993) Evolutionary conservation and structural similarities of carotenoid biosynthetic gene products from photosynthetic and non-photosynthetic organisms. Meth. Enzymol. **214**, 297-311.

Armstrong, G.A., A. Schmidt, G. Sandmann and J.E. Hearst (1990) Genetic and biochemical characterisation of carotenoid biosynthesis mutants of *Rhodobacter capsulatus*. J. Biol. Chem. **265**, 8329-8338.

Ausich, R.L. (1994) Production of carotenoids by recombinant DNA technology. Pure Appl. Chem. **66**, 1057-1062.

Baba, T. and C.M. Allen (1980) Prenyl transferases from *Micrococcus luteus*: characterisation of undecaprenyl pyrophosphate synthetase. Arch. Biochem. Biophys. **200**, 474-484.

Baima, S., A. Carattoli, G. Macino and G. Morrelli (1992) Photoinduction of *albino-3* gene expression in *Neurospora crassa* conidia. J. Photchem. Photobiol. B:Biol. **15**, 233-238.

Baima, S., G. Macino and G. Morelli (1991) Photoregulation of the *albino-3* gene in *Neurospora crassa*. J. Photochem. Photobiol. B:Biol. **11**, 107-115.

Bartley, G.E., T.J. Schmidhauser, C. Yanofsky and P.A. Scolnik (1990) Carotenoid desaturases from *Rhodobacter capsulatus* and *Neurospora crassa* are structurally and functionally conserved and contain domains homologous to flavoprotein disulfide oxidoreductases. J. Biol. Chem. **265**, 16020-16024.

Bartley, G.E. and P.A. Scolnik (1989) Carotenoid biosynthesis in photosynthetic bacteria. Genetic characterization of the *Rhodobacter capsulatus* Crt I protein. J. Biol. Chem. **264**, 13109-13113.

Bartley, G.E. and P.A. Scolnik (1993) cDNA cloning, expression during development and genome mapping of *PSY2* a second tomato gene encoding phytoene synthase. J. Biol. Chem. **268**, 25718-257721.

Bartley, G.E., P.A. Scolnik and G. Giuliano (1994) Molecular biology of carotenoid biosynthesis in plants. Ann. Rev. Plant Physiol. Plant Mol. Biol. **45**, 287-301.

Bartley, G.E., P.V. Viitanen, K.O. Bacot and P.A. Scolnik (1992) A tomato gene expressed during fruit ripening encodes an enzyme of the carotenoid biosynthesis pathway. J. Biol. Chem. **267**, 5036-5039.

Bartley, G.E., P.V. Viitanen, I. Pecker, D. Chamovitz, J. Hisrchberg and P.A. Scolnik (1991) Molecular cloning and expression in photosynthetic bacteria of a soybean cDNA coding for phytoene desaturase, an enzyme of the carotenoid biosynthesis pathway. Proc. Natl. Acad. Sci. USA **88**, 6532-6536.

Bendich, A. (1989) Carotenoids and the immune response. J. Nutr. **119**, 112-115.

Bendich, A. (1994) Recent advances in clinical research involving carotenoids. Pure Appl. Chem. **66**, 1017-1024.

Benton, W.D. and R.W. Davis (1977) Screening λ gt recombinant clones by hybridization to single plaques *in situ*. Science **196**, 180.

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

Billing, E. and L.A.E. Baker (1963) Characteristics of *Erwinia*-like organisms found in plant material. J. Appl. Bacteriol. **26**, 58-65.

Bird, C.R., J.A. Ray, J.D. Fetcher, J.M. Boniwell, A.S. Bird, C. Teeulieres, I. Blain, P.M. Bramley and W. Schuch (1991) Using antisense RNA to study gene function: inhibition of carotenoid biosynthesis in transgenic tomatoes. Biotechnology **9**, 635-639.

Birnboim, H.C. and J. Doly (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucl. Acids Res. **7**, 1519-1523.

Bissig, I., R.A. Brunisholz, F. Suter, R.J. Cogdell and H. Zuber (1988) The complete amino acid sequences of the B800-850 antenna polypeptides from *Rhodopseudomonas acidophila* strain 7750. Z. Naturforsch **43c**, 77-83.

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

Bouvier, F., P. Hugueney, A. d'Harlingue, M. Kuntz and B. Camara (1994) Xanthophyll biosynthesis in chromoplasts: isolation and molecular cloning of an enzyme catalyzing the conversion of 5,6-epoxycarotenoid into ketocarotenoid. The Plant J. **6**, 45-54.

Bramley, P.M. (1985) The *in vitro* biosynthesis of carotenoids. Adv. Lipid Res. **21**, 243-279.

Bramley, P.M. and A. Mackenzie (1988) Regulation of carotenoid biosynthesis. Curr. Topics in Cell. Reg. **29**, 291-343.

Bramley, P.M., C. Teulieres, I. Blain, C. Bird and W. Schuch (1992) Biochemical characterization of transgenic tomato plants in which carotenoid synthesis has been inhibited through the expression of antisense RNA to pTOM5. Plant J. **2**, 343-349.

Britton, G. (1976). Biosynthesis of Carotenoids. p. 362-327 *In* T.W. Goodwin (ed). Chemistry and Biochemistry of Plant Pigments. Academic Press, New York.

Britton, G. (1983). The Biochemistry of the Natural Pigments. Cambridge Univ. Press, Cambridge, U.K.

Britton, G. (1988) Biosynthesis of carotenoids. p. 262-327. *In* T.W. Goodwin (ed). Chemistry and Biochemistry of Plant Pigments. Academic Press, New York.

Brown, D.J., G. Britton and T.W. Goodwin (1975) Carotenoid biosynthesis by a cell-free preparation from a Flavobacterium species. Biochem. Soc. Trans. **3**, 741-742.

Buckner, B., T.L. Kelson and D.S. Robertson (1990) Cloning of the *y1* locus of maize, a gene involved in the biosynthesis of carotenoids. Plant Cell **2**, 867-876.

Buckner, B. and D.S. Robertson (1993) Cloning of carotenoid biosynthetic genes from maize. Meth. Enzymol. **214**, 311-323.

Burton, G.W. (1989) Antixoidant action of carotenoids. J. Nutr. **119**, 109-111.

Calo, P., T. de Miguel, J.B. Velazquez and T.G. Villa (1995) Mevalonic acid increases *trans*-astaxanthin and carotenoid biosynthesis in *Phaffia rhodozyma*. Biotec. Letts. **17**, 575-578.

Camara, B. and O. Dogbo (1986) Demonstration and solubilisation of lycopene cyclase from *Capsicum* chromoplast membranes. Plant Physiol. **80**, 172-174.

Campbell, A. (1971). Genetic structure in the bacteriophage lambda. Cold Spring Harbour Laboratory, Cold Spring Harbor, New York.

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

Carattoli, A., C. Cogoni, G. Morelli and G. Macino (1994) Molecular characterisation of upstream regulatory sequences controlling the photo-induced expression of the *albino-3* gene of *Neurospora crassa*. Mol. Microbiol. **13**, 787-795.

Carattoli, A., N. Romano, P. Ballario, G. Morelli and G. Macino (1992) The *Neurospora crassa* carotenoid biosynthesis gene (albino 3) reveals highly conserved regions among prenyl transferases. J. Biol. Chem. **266**, 5854-5859.

Chamovitz, D., N. Misawa, G. Sandmann and J. Hirschberg (1992) Molecular cloning and expression in *Escherichia coli* of a cyanobacterial gene coding for phytoene synthase, a carotenoid biosynthesis enzyme. FEBS Lett. **296**, 305-310.

Chamovitz, D., I. Pecker and J. Hirschberg (1991) The molecular basis of resistance to the herbicide norflurazon. Plant Mol. Biol. **16**, 967-974.

Chamovitz, D., G. Sandmann and J. Hirschberg (1993) Molecular and biochemical characterization of herbicide resistant mutants of cyanobactteria reveals that phytoene desaturation is a rate-limiting step in carotenoid biosynthesis. J. Biol. Chem. **268**, 17348-17353.

Chatterjee, A.K. and L.N. Gibbins (1971) Induction of non-pigmented variants of *Erwinia herbicola* by incubation at supraoptimal temperatures. J. Bacteriol. **105**, 107-112.

Cheniclet, C., F. Rafia, A. Saint-Guily, A. Verna and J.P. Carde (1992) Loacalization of the enzyme GGPP synthase in *Capsicum* friuts by immunogold cytochemistry after conventional chemical fixation or quick-freezing followed by freeze-substitution. Labelling evolution during fruit ripening. Biol. Cell. **75**, 145-154.

Clarke, L. and J. Carbon (1976) A colony bank containing synthetic Col E1 hybrid plasmids representative of the entire *E. coli* genome. Cell **9**, 91.

Cogdell, R.J. and H.A. Frank (1987) How carotenoids function in photosynthetic bacteria. Biochim. Biophys. Acta **895**, 63-79.

Cogdell, R.J., I. Durant, J. Valentine, I.G. Lindsay and K. Schmidt (1983) The isolation and partial characterisation of the light-harvesting pigment-protein complement of *Rhodopseudomonas acidophila*. Biochim. Biophys. Acta **722**, 427-435.

Cogdell, R.J. (1988) The function of pigments in chloroplasts. p. 183-230 *In* T.W. Goodwin (ed). Chemistry and Biochemistry of Plant Pigments. Academic Press, New York.

Cogoni, C., N. Romano and G. Macino (1994) Suppression of gene expression by homologous transgenes. Antonie van Leeuwenhoek **65**, 205-209.

Compton, T. (1990) Degenerate primers for DNA amplfication. p. 39-45. *In* Innis, M.A., Gelfand, D.H., Sninsky, J.J. & White, T.J. (eds). PCR Protocols, a guide to methods and applications. Academic Press, San Diego.

Coomber, S.A., M. Chaudri, A. Connor, G. Britton and C.N. Hunter (1990) Localized transposon Tn5 mutagenesis of the photosynthetic gene cluster of *Rhodobacter sphaeroides*. Mol. Microbiol. **4**, 977-989.

Coomber, S.A. and C.N. Hunter (1989) Construction of a physical map of the 45kb photosynthetic gene cluster of *Rhodobacter sphaeroides*. Arch. Microbiol. **151**, 454-458.

Crowe, J.S., H.J. Cooper, M.A. Smith, M.J. Sims, D. Parker and D. Gewert (1991) Improved cloning efficiency of polymerase chain reaction (PCR) products after proteinase K digestion. Nucleic Acids Res. **19**, 184.

Cunningham, F.X., D. Chamovitz, N. Misawa, E. Gantt and J. Hirschberg (1993) Cloning and functional expression in *Escherichia coli* of a cyanobacterial gene for lycopene cyclase, the enzyme that catalyses the biosynthesis of b-carotene. FEBS Lett. **328**, 130-138.

Cunningham, F.X., Z. Sun, D. Chamovitz, J. Hirsschberg and E. Gantt (1994) Molecular structure and enzymaticc function of lycopene cyclase from the cyanobacterium *Synechococcus* sp. Strain PCC7942. Plant Cell **6**, 1107-1121.

Dahl, T.A., W.R. Midden and P.E. Hartman (1989) Comparison of killing of Gram-negative and Gram-positive bacteria by pure singlet oxygen. J. Bacteriol. **171**, 2188-2194.

Devereux, J., P. Haeberli and O. Smithies (1984) A comprehensive set of sequence analysis programs for the VAX. Nuc. Acids Res. **12**, 387-395.

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

Don, R.H., P.T. Cox, B.J. Wainwright, K. Baker and J.S. Mattick (1991) 'Touchdown' PCR to circumvent spurious priming during gene amplification. Nuc. Acids Res. **19**, 4008.

Dubendorf, J.W. and F.W. Studier (1991) Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with *lac* repressor. J. Mol. Biol. **219**, 45-59.

Dumas Milne Edwards, J.B., P. Ravassard, C. Icard-Lipkalns and J. Mallet (1995) cDNA cloning by RT-PCR. p. 89-118 *In* McPherson, M.J., Hames, B.D. & Taylor, G.R. (eds). PCR2; A practical approach. IRL Press, Oxford.

Ehrenshaft, M. and M.E. Daub (1994) Isolation, sequence and characterisation of the *Cereospora nicotianae* phytoene dehydrogenase gene. Appl. Environ. Micobiol. **60**, 2766-2771.

Enquist, L.W. and A. Skalka (1973) Replication of bacteriophage λ DNA dependent on the function of host and viral genes. I. Interaction of *red*, *gam* and *rec*. J. Mol. Biol. **75**, 185-.

Esterbauer, H., G. Wag and H. Puhl (1993) Lipid peroxidation and its role in atherosclerosis. Br. Med. Bull. **49**, 566-576.

Finnerty, W.R. (1992) The biology and genetics of the genus *Rhodococcus*. Ann. Rev. Microbiol. **46**, 193-218.

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

Fraser, P.D. and P.M. Bramley (1994) The purification of phytoene dehydrogenase from *Phycomyces blakesleeanus*. Biochem. Biophys. Acta **1212**, 59-66.

Fraser, P.D., H. Linden and G. Sandmann (1993) Purification and reactivation of recombinant *Synechococcus* phytoene desaturase from an overexpressing strain of *Escherichia coli*. Biochem. J. **291**, 687-692.

Fraser, P.D., N. Misawa, H. Linden, S. Yamano, K. Kobayashi and G. Sandmann (1992) Expression in *Esscherichia coli*, purification and reactivation of the recombinant *Erwinia uredovora* phytoene desaturase. J. Biol. Chem. **267**, 19891-19895.

Fraser, P.D. and G. Sandmann (1992) *In vitro* assays of three carotenogenic membrane-bound enzymes from *E. coli* transformed with different *crt* genes. Biochem. Biophys. Res. Comm. **185**, 9-15.

Fraser, P.D., M.R. Truesdale, C.R. Bird, W. Schuch and P.M. Bramley (1994) Carotenoid biosynthesis during tomato fruit development. Plant Physiol. **105**, 405-413.

Fujii, H., H. Sagami, T. Koyama, K. Ogura and S. Seto (1980) Variable product specificity of solanesyl pyrophosphate synthetase. Biochem. Biophys. Res. Comm. **96**, 1648-1653.

Gardiner, A.T., R.J. Cogdell and S. Takaichi (1993) The effect of growth conditions on the light-harvesting apparatus in *Rhodopseudomonas acidophila*. Photosynth. Res. **38**, 159-167.

Gari, E., J.C. Toledo, I. Gibert and J. Barbe (1992) Nucleotide sequence of the methyoxyneurosporene dehydrogenase gene from *Rhodobacter sphaeroides*: Comparison with other bacterial carotenoid dehydrogenases. FEMS Microbiol. Lett. **93**, 103-108.

Gassmann, M., P. Thommes, T. Weiser and U. Hubscher (1990) Efficient production of chicken egg yolk antibodies against a conserved mammalian protein. FASEB J. **4**, 2528-2532.

Gey, K.F. (1993) Prospects for the prevention of free radical disease, regarding cancer and cardiovascular disease. Br. Med. Bull. **49**, 679-699.

Girgis, S.I., M. Alevizaki, P. Denny, G.J.M. Ferrier and S. Legon (1988) Generation of DNA probes from peptides with highly degenerate codons using mixed primer PCR. Nuc. Acids Res. **16**, 10371.

Goodwin, T.W. (1980) The Biochemistry of the carotenoids Volume I Plants. Chapman and Hall, London.

Goodwin, T.W. (1983) Developments in carotenoid biochemistry over 40 years. Biochem. Soc. Trans. **11**, 473-483.

Griffin, T.J. and R.D. Kolodner (1990) Purification and preliminary characterization of the *Escherichia coli* K-12 RecF protein. J. Bacteriol. **172**, 6291-6299.

Griffiths, M. and R.Y. Stainier (1956) Some mutational changes in the photosynthetic pigment system of *Rhodopseudomonas sphaeroides*. J. Gen. Microbiol. **14**, 698-715.

Guiliano, G., G.E. Bartley and P.A. Scolnik (1993) Regulation of carotenoid biosynthesis during tomato development. Plant Cell **5**, 379-387.

Guiliano, G., D. Pollock and P.A. Scolnik (1986) The gene *crtI* mediates the conversion of phytoene into coloured carotenoids in *Rhodopseudomonas capsulata*. J. Biol. Chem. **261**, 12925-12929.

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

Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166**, 557-580.

Hart, S., R. Kirby and D.R. Woods (1990) Structure of a *Rhodococcus* gene encoding pigment production in *Escherichia coli*. J. Gen. Microbiol. **136**, 1357-1363.

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

Hohn, B. and K. Murray (1977) Packaging recombinant DNA molecules into bacteriophage particles *in vitro*. Proc. Natl. Acad. Sci. USA **74**, 3259.

Holmes, D.S. and M. Quigley (1981) A rapid method for the preparation of bacterial plasmids. Anal. Biochem. **114**, 193-197.

Hoshino, T., R. Fujii and T. Nakahara (1993) Molecular cloning and sequence analysis of the *crtB* gene of *Thermus thermophilus* HB27, an extreme thermophile producing carotenoid pigments. Appl. Environ. Microbiol. **59**, 3150-3153.

Hoshino, T., R. Fujii and T. Nakahara (1994) Overproduction of carotenoids in *Thermus thermophilus*. J. Fermen. Bioeng. **77**, 423-424.

Hoshino, T., Y. Yoshino, E.D. Guevarra, S. Ishida, T. Hiruta, R. Fujii and T. Nakahara (1994) Isolation and partial characterization of carotenoid underproducing and overproducing mutants from an extremely thermophilic *Thermus thermophilus*. J. Fermen. Bioeng. **77**, 131-136.

Houssaini-Iraqui, M., H.L. David, S. Clava-Seres, F. Hilali and N. N. Rastogi (1993) Characterization of *carα*, *carLep* and *crtI* genes cotrolling biosynthesis of carotenes in *Mycobacterium aurum*. Curr. Microbiol. **27**, 317-322.

Houssaini-Iraqui, M., R. Lazraq, S. Clavel-Seres, N. Rastogi and H.L. David (1992) Cloning and expression of *Mycobacterium aurum* carotenogenesis genes in *Mycobacterium smegmatis*. FEMS Microbiol. Lett. **90**, 239-244.

Hugueney, P., S. Romer, M. Kuntz and B. Camara (1992) Characterization and molecular cloning of a flavoprotein catalyzing the synthesis of phytofluene and z-carotene in *Capsicum* chromoplasts. Eur. J. Biochem. **209**, 399-407.

Hundle, B.S., M. Alberti, V. Nievelstein, P. Beyer, H. Kleinig, G.A. Armstrong, D.H. Burke and J.E. Hearst (1994) Functional assignment of *Erwinia herbicola* Eho 10 carotenoid genes expressed in *Escherichia coli*. Mol. Gen. Genet. **245**, 406-416.

Hundle, B.S., P. Beyer, H. Kleinig, G. Englert and J.E. Hearst (1991) Carotenoids of *Erwinia herbicola* and an *E. coli* HB101 strain carrying the *Erwinia herbicola* carotenoid gene cluster. Photochem. Photobiol. **54**, 89-93.

Hundle, B.S., D.A. O'Brien, M. Alberti, P. Beyer and J.E. Hearst (1992) Functional expression of zeaxanthin glucosyltransferase from *Erwinia herbicola* and a proposed uridine diphosphate binding site. Proc. Natl. Acad. Sci. USA **89**, 9321-9325.

Hundle, B.S., D.A. O'Brien, P. Beyer, H. Kleinig and J.E. Hearst (1993) *In vitro* expression and activity of lycopene cyclase and b-carotene hydroxylase from *Erwinia herbicola*. FEBS Lett. **315**, 329-334.

Hunter, C.N., B.S. Hundle, J.E. Hearst, H.P. Lang, A.T. Gardiner, S. Takaichi and R.J. Cogdell (1994) Introduction of new carotenoids into the bacterial photosynthetic apparatus by combining the carotenoid biosynthetic pathways of *Erwinia herbicola* and *Rhodobacter sphaeroides*. J. Bacteriol. **176**, 3692-3697.

Ichhiyama, S., K. Shimokata and M. Tsukamaura (1989) Carotenoid pigments of genus *Rhodococcus*. Microbiol. Immunol. **33**, 503-508.

IUPAC Commission on the Nomenclature of Organic Chemistry and IUPAC-IUB Commision on Biochemical Nomenclature (1971) Rules for the nomenclature of carotenoids. p. 851-864. *In* O. Isler (ed). Carotenoids. Birkhauser, Basel.

Jacques, P.F., S.C. Hartz, L.T. Chylack, R.B. McGandy and J.A. Sadowski (1988) Nutritional status in persons with and without senile cataract: blood vitamin and mineral levels. Am. J. Clin. Nutr. **48**, 152-158.

Jagannadham, M.V., V.J. Rao and S. Shivaji (1991) The major carotenoid pigment of a psychrotrophic *Micrococcus roseus* strain: purification, structure and interaction with synthetic membranes. J. Bacteriol. **73**, 7911-7917.

Kaiser, K. and N.E. Murray (1985) The use of phage lambda replacement vectors in the construction of representative genomic DNA libraries. p. 1-47. *In* D.M. Glover (ed). DNA Cloning; A practical Approach. IRL Press, Oxford.

Kajiwara, S., T. Kazikono, T. Saito, K. Kondo, T. Ohtani, N. Nishio, S. Nagai and N. Misawa (1995) Isolation and functional identification of a novel cDNA for astaxanthin biosynthesis from *Haematococcus pluvialis*, and astaxanthin synthesis in *Escherichia coli*. Plant Mol. Biol. **29**, 343-352.

Kandutsch, A.A. (1969) Geranylgeranyl pyrophosphate synthetase of *Micrococcus lysodeikticus*. Meth. Enzymol. **15**, 490-494.

Karvouni, Z., I. John, J.E. Taylor, C.F. Watson, A.J. Turner and D. Grierson (1995) Isolation and characterisation of a melon cDNA clone encoding phytoene synthase. Plant Mol. Biol. **27**, 1153-1162.

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

Kaufmann, N., H. Hudig and G. Drews (1984) Transposon Tn5 mutagenesis of genes for the photosynthetic apparatus in *Rhodopseudomonas capsulata*. Mol. Gen. Genet. **198**, 153-158.

Kidd, K.K. and G. Ruano (1995) Optimizing PCR. p. 1-21 *In* M.J. McPherson, B.D. Hames, and G.R. Taylor, (eds). PCR2; A practical approach. IRL Press, Oxford.

Kleinig, H. (1975) On the utilization *in vivo* of lycopene and phytoene as precursors for the formation of carotenoid glucoside ester and on the regulation of carotenoid biosynthesis in *Myxococcus fulvus*. Eur. J. Biochem. **57**, 301-308.

Knoth, K., S. Roberds, C. Poteet and M. Tamkun (1988) Highly degenerate, inosine containing primers specifically amplify rare cDNA using the polymerase chain reaction. Nucleic Acids Res. **16**, 10932.

Kritsky, M.S., V.Y. Sokolovsky, T.A. Belozersskaya and E.K. Cherysheva (1982) Relationship between cyclic AMP level and accumulation of carotenoid pigments in *Neurospora crassa*. Arch. Microbiol. **133**, 206-208.

Kumagai, M.H., J. Donson, G. Della-Cioppa, D. Harvey, K. Hanley and L.K. Grill (1995) Cytoplasmic inhibition of carotenoid biosynthesis with virus-derived RNA. Proc. Natl. Acad. Sci. USA **92**, 1679-1683.

Kuntz, M., S. Romer, C. Surie, P. Hugueney, J.H. Weil, R. Schantz and B. Camara (1992) Identification of a cDNA for the plastid-located GGPP synthase from *Capsicum annuum*: correlative increase in enzyme activity and transcript level during fruit ripening. The Plant J. **2**, 25-34.

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

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **277**, 680-685.

Lang, H.P., R.J. Cogdell, A.T. Gardner and C.N. Hunter (1994) Early steps in carotenoid biosynthesis: sequences and transcriptional analysis of the *crtI* and *crtB* genes of *Rhodobacter Sphaeroides* and overexpression and reactivation of *crtI* in *Escherichia coli* and *R. Sphaeroides*. J. Bacteriol. **176**, 3859-3869.

Lang, H.P., R.J. Cogdell, S. Takaichi and C.N. Hunter (1995) Complete DNA sequence, specific Tn5 insertion map and gene assignment of the carotenoid biosynthesis pathway of *Rhodobacter sphaeroides*. J. Bacteriol. **177**, 2064-2073.

Lang, H.P. and C.N. Hunter (1994) The relationship between carotenoid biosynthesis and the assembly of the light-harvesting LH2 complex in *Rhodobacter sphaeroides*. Biochem. J. **298**, 197-205.

Lee, C.C., X. Wu, R.A. Gibbs, R.G. Cook, D.M. Muzny and C.T. Caskey (1988) Generation of cDNA probes directed by amino acid sequence: cloning of urate oxidase. Science **239**, 1288-1291.

Lee, L.Y. and S.T. Liu (1991) Characterisation of the yellow pigment genes of *Erwinia herbicola*. Mol. Microbiol. **5**, 217-224.

Lerner, R.A. (1984) Antibodies of pre-determined specificity in biology and medicine. Adv. Immunol. **36**, 1.

Li, C. and T.J. Schmidhauser (1995) Developmental and photoregulation of *al-1* and *al-2*, structural genes for two enzymes essential for carotenoid biosynthesis in *Neurospora*. Dev. Biology **169**, 90-95.

Li, Z.H., P.D. Matthews, B. Burr and E.T. Wurtzel (1996) Cloning and characterization of a maize cDNA encoding phytoene desaturase, an enzyme of the carotenoid biosynthetic pathway. Plant Mol. Biol. **30**, 269-279.

Lin, T.P., E.H. Lai, K.L. To, Y.S. Chang and S.H. Liu (1994) Transcriptional activation of flanking sequences by Tn*1000* insertion. Mol. Gen. Genet. **245**, 417-423.

Linden, H., M.M. Lucas, M.R. de Felipe and G. Sandmann (1993) Immunogold localization of phytoene desaturase in higher plant chloroplasts. Physiologia Plantarum **88**, 229-236.

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

Linden, H., A. Vioque and G. Sandmann (1993) Isolation of a carotenoid biosynthesis gene coding for ζ-caroptene desaturase from *Anabaena* PCC 7120 by heterologous complementation. FEMS Microbiol. Letts. **106**, 99-104.

Liu, S.T. (1993) Carotenoid biosynthesis genes as a genetic marker for the purpose of gene cloning. Bioch. Biophys. Res. Comm. **195**, 259-263.

Looman, A.C., J. Bodlaender, M. de Gruyter, A. Vogelaar and P.H. van Knippenberg (1986) Secondary structure as primary determinant of the efficiency of ribosomal binding sites in *Escherichia coli*. Nucleic Acids Res. **14**, 5481-5497.

Lotan, T. and J. Hirschberg (1995) Cloning and expression in *Escherichia coli* of the gene encoding β-C-4-oxygenase, that converts β-carotene to the ketocarotenoid canthaxanthin in *Haematococcus pluvialis*. FEBS Letts. **364**, 125-128.

Ma, D., D.N. Cook, D.A. O'Brien and J.E. Hearst (1993) Analysis of the promoter and regulatory sequences of an oxygen-regulated *bch* operon in *Rhodobacter capsulatus* by site-directed mutagenesis. J. Bacteriol. **175**, 2037-2045.

Mack, D. and J.J. Sninsky (1988) A sensitive method for the identification of uncharacterised viruses related to known virus groups: hepadnavirus model system. Proc. Natl. Acad. Sci. USA **85**, 6977-6981.

Maniatis, T., E.F. Fritsch and J. Sambrook (1989) Molecular cloning: A laboratory Manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Mann, V., I. Pecker and J. Hirschberg (1994) Cloning and characterisation of the gene for phytoene desaturase (*Pds*) from tomato (*Lycopersicon esculentum*). Plant Mol. Biol. **24**, 429-434.

Mares-Perlman, J., W.E. Brady, R. Klein, B.E.K. Klein, P. Bowen, M. Stacewicz-Sapuntzakis and M. Palta (1995) Serum antioxidants and age related macular degeneration in a population-based case-control study. Arch. Ophtalmol. **113**, 1518-1523.

Marrs, B. (1981) Mobilization of the genes for photosynthesis from *Rhodopseudomonas capsulata* by a promiscuous plasmid. J. Bacteriol. **146**, 1003-1012.

Martin, F.H. and M. Castro (1985) Base pairing involving deoxyinosine: implications for probe design. Nucl. Acids Res. **13**, 8927.

Martinez-Ferez, I., B. Fernandez-Gonzalez, G. Sandmann and A. Vioque (1994) Cloning and expression in *Escherichia coli* of the gene coding for phytoene synthase from the cyanobacterium *Synechocystis* sp. PCC 6803. Biochim. Biophys. Acta **1218**, 145-152.

Martinez-Ferez, I.M. and A. Vioque (1992) Nucleotide sequence of the phytoene desaturase gene from *Synechocystis* sp. PCC6803 and characterisation of a new mutation which confers resistance to the herbicide norflurazon. Plant Mol. Biol. **18**. 981-983.

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

Martinez-Laborda, A., M. Elias, R. Ruiz-Vazquez and F.J. Murillo (1986) Insertions of Tn5 linked to mutations affecting carotenoid synthesis in *Myxococcus xanthus*. Mol. Gen. Genet. **205**, 107-114.

Math, S.K., J.E. Hearst and C.D. Poulter (1992) The *crtE* gene in *Erwinia herbicola* encodes geranylgeranyl diphosphate synthase. Proc. Natl. Acad. Sci. USA **89**, 6761-6764.

Mathews, M.M. (1963) Comparitive study of lethal photosensitisation of *Sarcina lutea* by 8-methoxypsoralen and by toluidene blue. J. Bacteriol. **85**, 322-328.

Mathews-Roth, M.M. (1986) Beta-carotene therapy for erythropoietic protoporphyria and other photosensitivty diseases. Biochimie **68**, 875-884.

Mathews-Roth, M.M. and N.I. Krinsky (1970) Failure of conjugated octaene carotenoids to protect a mutant of *Sarcina lutea* against lethal photosensitisation. Photochem. Photobiol. **11**, 555-557.

Mathews-Roth, M.M. and N.I. Krinsky (1970) Studies on the protective function of the carotenoid pigments of *Sarcina lutea*. Photochem. Photobiol. **11**, 419-428.

Mathews-Roth, M.M., T. Wilson, E. Fujimori and N.I. Krinsky (1974) Carotenoid chromophore length and protection against photosensitisation. Photochem. Photobiol. **19**, 217-222.

McDermott, G., S.M. Prince, A.A. Freer, A.M. Hawthornwaite-Lawless, M.Z. Papiz, R.J. Cogdell and N.W. Isaacs (1995) Crystal structure of an integral membrane light-harvesting complex from photosynthetic bacteria. Nature **374**, 517-521.

McDermott, J.C.B., D.J. Brown, G. Britton and T.W. Goodwin (1974) Alternative pathways of zeaxanthin biosynthesis in a *Flavobacterium* species. Biochem. J. **144**, 231-243.

McGowan, S.J., H.C. Gorham and D.A. Hodgson (1993) Light-induced carotenogenesis in *Myxococcus xanthus*: DNA sequence analysis of the *carR* region. Mol. Microbiol. **10**, 713-735.

Meagher, R.B. (1977) Protein expression in E. coli minicells by recombinant plasmids.. Cell **10**, 521-536.

Meyer, P.S. and J.C. Du Preez (1994) Photo-regulated astaxanthin production by *Phaffia rhodozyma* mutants. System. Appl. Microbiol. **17**, 24-31.

Michalowski, C.B., W. Loffelhardt and H.J. Bohnert (1991) An ORF323 with homolgy to *crtE*, specifying prephytoene pyrophosphate dehydrogenase,

is encoded by cyanelle DNA in the eukaryotic alga *Cyanophora paradoxa*. J. Biol. Chem. **266**, 11866-11870.

Misawa, N., S. Kajiwara, K. Kondo, A. Yokoyama, Y. Satomi, T. Saito, W. Miki and T. Ohtani (1995) Canthaxanthin biosynthesis by the conversion of methylene to keto groups in a hydrocarbon β-carotene by a single gene. Biochem. Biophys. Res. Comm. **209**, 867-876.

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

Misawa, N., K. Masamoto, T. Hori, T. Ohtani, P. Boger and G. Sandmann (1994) Expression of an *Erwinia* phytoene desaturase gene not only confers mulitple resistance to herbicides interfering with carotenoid biosynthesis but also alters xanthophyll metabolism in transgenic plants. Plant J. **6**, 481-489.

Misawa, N., Y. Satomi, K. Kondo, A. Yokoyama, S. Kajiwara, T. Saito, T. Ohtani and W. Miki (1995) Structure and functional analysis of a marine bacterial gene cluster and astaxanthin biosynthetic pathway proposed at the gene level. J. Bacteriol. **177**, 6575-6584.

Misawa, N., M.R. Truesdale, G. Sandmann, P.D. Fraser, C. Bird, W. Schuch and P.M. Bramley (1994) Expression of a tomato cDNA coding for phytoene synthase in *Escherichia coli*, phytoene formation *in vivo* and *in vitro* and functional analysis of the various truncated gene products. J. Biochem. **116**, 980-985.

Misawa, N., S. Yamano and H. Ikenaga (1991) Production of β-carotene in *Zymomonas mobilis* and *Agrobacterium tumefaciens* by introduction of the biosynthesis genes from *Erwinia uredovora*. Appl. Environ. Microbiol. **57**, 1847-1849.

Misawa, N., S. Yamano, H. Linden, M.R. de Felipe, M. Lucas, H. Ikenaga and G. Sandmann (1993) Functional expression of the *Erwinia uredovora* carotenoid biosynthesis gene *crtI* in transgenic plants showing an

increase of β-carotene biosynthesis activity and resistance to the bleaching herbicide norflurazon. Plant J. **4**, 833-840.

Mosqueda-Cano, G. and J.F. Gutierrez-Corona (1995) Environmental and developmental regulation of carotenogenesis in the dimorphic fungus *Mucor rouxii*. Curr. Microbiol. **31**, 141-145.

Nakagawa, M. and N. Misawa (1991) Analysis of carotenoid glycosides produced in Gram-negative bacteria by introduction of the *Erwinia uredovora* carotenoid biosynthesis genes. Agric. Biol. Chem. **55**, 2147-2148.

Nakamura, T. (1936) A new bacterioxanthophyll the yellow pigment of *Sarcina lutea*. Bull. Chem. Soc. **11**, 176-179.

Nelis, H.J. and A.P. De Leenheer (1989) Reinvestigation of *Brevibacterium* sp. strain KY-4313 as a source of canthaxanthin. Appl. Environ. Microbiol. **55**, 2505-2510.

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

Nevin, D.E. and J.M. Pratt (1991) A coupled *in vitro* transcription-translation system for the exclusive synthesis of polypeptides expressed from the T7 promoter. FEBS **291**, 259-263.

Nicolas, F.J., R.M. Ruiz-Vazquez and F.J. Murillo (1994) A genetic link between light response and multicellular development in the bacterium *Myxococcus xanthus*. Genes Dev. **8**, 2375-2387.

Norris, S.R., T.R. Barrette and D. DellaPenna (1995) Genetic dissection of carotenoid synthesis in arabidopsis defines plastoquinone as an essential component of phytoene desaturation. The Plant Cell **7**, 2139-2149.

Ohnuma, S., T. Koyama and K. Ogura (1991) Purification of solanesyl-diphosphate synthase from *Micrococcus luteus*. J. Biol. Chem. **266**, 23706-23713.

Ohnuma, S., T. Koyama and K. Ogura (1992) Chain lengtth distribution of the products formed in solanesyl-diphosphate synthase reaction. J. Biochem. **112**, 743-749.

Owens, T.G. (1992) Dynamics and mechanism of singlet energy transfer between carotenoids and chlorophylls: light harvesting and non-photochemical fluorescence quenching. Photosynthesis Res. **34**, symposium 2-2.

Pecker, I., D. Chamovitz, H. Linden, G. Sandmann and J. Hirschberg (1992) A single polypeptide catalyzing the conversion of phytoene to ζ-carotene is transcriptionally regulated during tomato fruit ripening. Proc. Natl. Acad. Sci. USA **89**, 4962-4966.

Pemberton, J.M. and C.M. Harding (1986) Cloning of carotenoid biosynthesis genes from *Rhodopseudomonas sphaeroides*. Curr. Microbiol. **14**, 25-29.

Pemberton, J.M. and C.M. Harding (1987) Expression of *Rhodopseudomonas sphaerodes* carotenoid photopigment genes in phylogenetically related nonphotosynthetic bacteria. Curr. Microbiol. **15**, 67-71.

Penfold, R.J. and J.M. Pemberton (1991) A gene from the photosynthetic gene cluster of *Rhodobacter sphaeroides* induces *trans* suppression of bacteriochlorophyll and carotenoid levels in *R. sphaeroides* and *R. capsulatus*. Curr. Microbiol. **23**, 259-263.

Penfold, R.J. and J.M. Pemberton (1994) Sequencing, chromosomal inactivation and functional expression in E. coli of *ppsR* a gene which represses carotenoid and bacteriochlorophyll synthesis in *Rhodobacter sphaeroides*. J. Bacteriol. **176**, 2869-2876.

Perry, K.L., T.A. Simonitch, K.J. Harrison-Lavoie and S.T. Liu (1986) Cloning and regulation of *Erwinia herbicola* pigment genes. J. Bacteriol. **168**, 607-612.

Pfennig, N. (1969) *Rhodopseudomonas acidophila*, sp. a new species of the budding purple nonsulfur bacteria. J. Bacteriol. **99**, 597-602.

Polson, A., B. von Wechmar and G. Fazakesley (1980) Antibodies from yolk of immunized hens. Immunol. Commun. **9**, 495-514.

Ponnampalam, S.N., J.J. Buggy and C.E. Bauer (1995) Characterization of an aerobic repressor that coordinately regulates bacteriochlorophyll, carotenoid and light harvesting-II expression in *Rhodobacter capsulatus*. J. Bacteriol. **177**, 2990-2997.

Pratt, J.M. (1984) Coupled Transcription-Translation in Prokaryotic Cell-Free Systems. p. 179-209. *In* B.D. Hames and S.J. Higgins (eds). Transcription and Translation: a practical approach. IRL Press, Oxford.

Raleigh, E.A. and G. Wilson (1986) *Escherichia coli* K-12 restricts DNA containing 5-methylcytosine. Proc. Natl. Acad. Sci. USA **83**, 9070.

Randall, L.L., S.J.S. Hardy and J.R. Thom (1987) Export of Protein: A Biochemical View. Ann. Rev. Microbiol. **41**, 507-541.

Ray, J., C. Bird, M. Maunders, D. Grierson and W. Schuh (1987) Sequence of pTOM5 a ripening related cDNA from tomato. Nuc. Acids Res. **15**, 10587.

Ray, J., P. Moureau, C. Bird, A. Bird, D. Grierson, M. Maunders, M. Truesdale, P. Bramley and W. Schuch (1992) Cloning and characterisation of a gene involved in phytoene synthesis from tomato. Plant Mol. Biol. **19**, 401-404.

Reith, M. and J. Munholland (1993) A high-resolution gene map of the chloroplast genome of the red algae *Porphyra purpurea*. Plant Cell **5**, 465-475.

Rock, C.D. and J.A.D. Zeevaart (1991) The *aba* mutant of *Arabidopsis thaliana* is impaired in epoxy-carotenoid biosynthesis. Proc. Natl. Acad. Sci. USA **88**, 7496-7499.

Romano, N. and G. Macino (1992) Quelling: transient inactivation of gene expression in *Neurospora crassa* by transformation with homologous sequences. Mol. Microbiol. **6**, 3343-3353.

Romer, S., P. Huguerney, F. Bouvier, B. Camara and M. Kuntz (1993) Expression of the genes encoding the early carotenoid biosynthesis enzymes in *Capsicum annuum*. Biochem. Biophys. Res. Comm. **196**, 1414-1421.

Rosenberg, S.M. (1987) Improvment *in vitro* packaging of λ DNA. Meth. Enzymol. **153**, 95.

Ruiz-Vazquez, R., M. Fontes and F.J. Murillo (1993) Clustering and co-ordinated activation of carotenoid genes in *Myxococcus xanthus* by blue light. Mol. Microbiol. **10**, 25-34.

Ruizzo, M.A., R. Bertekap and M.L. Mishkind (1992) Consequences of herbicide-induced pigment deficiencies on thylakoid membrane proteins of *Chlamydomonas reinhardtii*. Plant Science **81**, 13-20.

Sagami, H., K. Ogura and S. Seto (1977) Solanesyl pyrophosphate synthetase from *Micrococcus lysodeikticus*. Biochemistry **16**, 4616-4622.

Sandmann, G. (1991) Light-dependent switch from formation of poly-cis carotenes to all-trans carotenoids in the *Scenedesmus* mutant C-6D. Arch. Microbiol. **155**, 229-233.

Sandmann, G. and P.M. Bramley (1985) The *in vitro* biosynthesis of b-cryptoxanthin with Aphanocapsa membranes. Biochim. Biophys. Acta. **843**, 73-77.

Sandmann, G. and P.D. Fraser (1993) Differential inhibition of phytoene desaturases from diverse origins and analysis of resistant cyanobacterial mutants. Z. Naturforsch. **48c**, 307-311.

Sandmann, G. and S. Kowalczyk (1989) *In vitro* carotenogenesis and characterization of the phytoene desaturase reaction in *Anacystis*. Biochem. Biophys. Res. Commun. **163**, 916-21.

Sandmann, G. and N. Misawa (1992) New functional assignment of the carotenogenic genes *crtB* and *crtE* with constructs of these genes from *Erwinia* species. FEMS Microbiol. Lett. **90**, 253-258.

Sandmann, G., N. Misawa, M. Wiedemann, P. Vittorioso, A. Carattoli, G. Morelli and G. Macino (1993) Functional identification of *al-3* from neurospora crassa as the gene for geranylgeranyl pyrophosphate synthase by complementation with *crt* genes, *in vitro* characterization of the gene product and mutant analysis. J. Photochem. Photobiol. B: Biol. **18**, 245-251.

Sandmann, G., W.S. Woods and R.W. Tuveson (1990) Identification of carotenoids in *Erwinia herbicola* and in a transformed *E. coli* strain. FEMS Microbiol. Lett. **71**, 77-82.

Schmidhauser, T.J., F.R. Lauter, V.E. Russo and C. Yanofsky (1990) Cloning, sequence and photoregulation of *al-1*, a carotenoid biosynthetic gene of *Neurospora crassa*. Mol. Cell Biol. **10**, 5064-5070.

Schmidhauser, T.J., F.R. Lauter, M. Schumacher, W. Zhou, V.E.A. Russo and C. Yonofsky (1994) Characterization of *al-2*, the phytoene synthase gene of *Neurospora crassa*. J. Biol. Chem. **269**, 12060-12066.

Schmidt, A. and G. Sandmann (1990) Cloning and nucleotide sequence of the *crtI* gene encoding phytoene dehydrogenase from the cyanobacterium *Aphanocapsa* PCC6714. Gene **91**, 113-117.

Schmidt, A., G. Sandmann, G.A. Armstrong, J.E. Hearst and P. Beyer (1989) Immunological detection of phytoene desaturase in algae and higher plants using an antiserum raised against a bacterial fusion gene construct. Eur. J. Biochem. **184**, 375-378.

Schnurr, G., N. Misawa and G. Sandmann (1996) Expression, purification and properties of lycopene cyclase from *Erwinia uredovora*. Biochem. J. **315**, 869-874.

Schwender, J., M. Seeman, H.K. Lichtenthaler and M. Rohmer (1996) Biosynthesis of isoprenoids (carotenoids, sterols, prenyl side-chains of chlorophylls and plastoquinone) via a novel pyruvate/glyceraldehyde 3-phosphate non-mevalonate pathway in the green alga *Scenedesmus obliquus*. Biochem. J. **316**, 73-80.

Scolnik, P.A. and G.E. Bartley (1993) Phytoene desaturase from *Arabidopsis*. Plant Physiol. **103**, 1475.

Scolnik, P.A. and G. Guiliano (1993) Regulation of carotenoid biosynthesis genes during plant development. Pure and Appl. Chem. **66**, 1063-1068.

Scolnik, P.A., M.A. Walker and B.L. Marrs (1980) Biosynthesis of carotenoids derived from Neurosporene in *Rhodopseudomonas capsulatus*. J. Biol. Chem. **255**, 2427-2432.

Serrano, A., P. Gimenez, A. Schmidt and G. Sandmann (1990) Immunocytochemical localisation and functional determination of phytoene desaturase in photo-autotrophic prokaryotes. J. Gen. Microbiol. **136**, 2465-2469.

Shen, Z., J. Liu, R.L. Wells and M.M. Elkind (1993) Cycle sequencing using degenerate primers containing inosines. Biotechniques **15**, 82-84.

Sies, H., W. Stahl and A.R. Sundquist (1992) Antioxidant functions of vitamins. Ann. N. Y. Acad. Sci. **669**, 7-20.

Smith, T.F., M.S. Waterman and J.R. Sadler (1983) Statistical characterization of nucleic acid sequence functional domains. Nuc. Acids Res. **11**, 2205.

Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**, 503-517.

Sporn, M.B. and A.B. Roberts (1983) Role of retinoids in differentiation and carcinogenesis. Cancer Res. **43**, 3034-3040.

Stanssens, P., E. remaut and W. Fiers (1985) Alterations upstream from the Shine-Dalgarno region and their effect on bacterial gene expression. Gene **36**, 211-223.

Starr, M.P. (1981). The genus *Erwinia*. p. 1260-1271 *In* M.P. Starr, H.G. Truper, A. Balows and H.G. Schlegel (ed). The prokaryotes. Springer-Verlag Berlin.

Stoker, N.G., J.M. Pratt and I.B. Holland (1984) *In vivo* gene expression systems in prokaryotes. p. 153-176. *In* B.D. Hames and S.J. Higgins (eds). Transcription and Translation; a practical approach. IRL Press, Oxford.

Studier, F.W. and B.A. Moffatt (1986) Use of bacteriophage T7 polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. **189**, 113-130.

Tabata, K., S. Ishida, T. Nakahara and T. Hoshino (1994) A carotenogenic gene cluster exists on a large plasmid in *Thermus thermophilus*. FEBS Lett. **341**, 251-255.

Tabor, S. and C.C. Richardson (1985) A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc. Natl. Acad. Sci. USA **82**, 1074-1078.

Takaichi, S., J. Ishidsu, T. Seki and S. Fukada (1990) Carotenoid pigments from Rhodococcus rhodochrous RNMSI: Two monocyclic carotenoids, a carotenoid monoglycoside and carotenoid glycoside monoesters. Agric. Biol. Chem. **54**, 1931-1937.

Takaichi, S. and J. Isshidsu (1993) Influence of growth temerature on compositions of carotenoids and fatty acids from carotenoid glucoside ester and from cellular lipids in *Rhodococcus rhodochrous* RNMS1. Biosci. Biotech. Biochem. **57**, 1886-1889.

Takatsuji, H., T. Nishino, I. Miki and H. Katsuki (1983) Studies on isoprenoid biosynthesis with bacterial intact cells. Biochem. Biophys. Res. Comm. **110**, 187-193.

Taylor, D.P., S.N. Cohen, W.G. Clark and B.L. Marrs (1983) Alignment of genetic and restriction maps of the photosynthesis region of the *Rhodopseudomonas capsulata* chromosome by a conjugation mediated marker rescue technique. J. Bacteriol. **154**, 580-590.

Thirkell, D. (1969) Growth and pigmentation of *Micrococcus radiodurans*. J. Gen. Microbiol. **55**, 337-340.

Thirkell, D. and M.I.S. Hunter (1969) Carotenoid-glycoprotein of *Sarcina flava* membrane. J. Gen. Microbiol. **58**, 289-292.

Thirkell, D. and M.I.S. Hunter (1969) The polar carotenoid fraction from *Sarcina flava*. J. Gen. Microbiol. **58**, 293-299.

Thirkell, D. and R.H.C. Strang (1967) Analysis and comparison of the carotenoids of *Sarcina flava* and *S. lutea*. J. Gen. Microbiol. **49**, 53-57.

To, K.Y., E.M. Lai, L.Y. Lee, T.P. Lin, C.H. Hung, C.L. Chen, Y.S. Chang and S.T. Liu (1994) Analysis of the gene cluster encoding carotenoid biosynthesis in *Erwinia herbicola* Eho 13. Microbiology **140**, 331-339.

Towbin, H., T. Staehelin and J. Gordon (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. Proc. Natl. Acad. Sci. USA **76**, 4350-4354.

Tuveson, R.W., R.A. Larson and J. Kagan (1988) Role of cloned carotenoid genes expressed in E. coli in protecting against inactivation by near-UV light and specific phototoxic molecules. J. Bacteriol. **170**, 4675-4680.

Vogelstein, B. and D. Gillespie (1979) Preparitive and analytical purification of DNA from agarose. Proc. Natl. Acad. Sci. USA **76**, 615-619.

Wada, K.N., Y. Wada, F. Ishibashi, T. Gojobori and T. Ikemura (1992) Codon usage tabulated from the GenBank genetic sequence data. Nuc. Acids Res. **20**, 2111-2118.

Weidemann, M., N. Misawa and G. Sandmann (1993) Purification and enzymatic characterization of the geranylgeranyl pyrophosphate synthase from *Erwinia uredovora* after expression in *Escherichia coli*. Arch. Biochem. Biophys. **306**, 152-157.

Wellington, C.L. and J.T. Beatty (1991) Overlapping mRNA transcripts of photosynthesis gene operons in *Rhodobacter capsulatus*. J. Bacteriol. **173**, 1432-1443.

Wellington, C.L., A.K.P. Taggart and J.T. Beatty (1991) Functional significance of overlapping transcripts of *crtEF*, *bchCA* and *puf* photosynthesis gene operons in *Rhodobacter capsulatus*. J. Bacteriol. **173**, 2954-2961.

Wetzel, C.M., C.Z. Jiang, L.J. Meehan, D.F. Voytas and S.R. Rodermel (1994) Nuclear-organelle interactions: the *immutans* variegation mutant of *Arabidopsis* is plastid autonomous and impaired in carotenoid biosynthesis. The Plant J. **6**, 161-175.

Wieland, B., C. Feil, E. Gloria-Maecker, G. Thumm, M. Lechner, J.M. Bravo, K. Poralla and F. Gotz (1994) Genetic and biochemical analyses of the biosynthesis of the yellow carotenoid 4,4'-diaponeurosporene of *Staphylococcus aureus*. J. Bacteriol. **176**, 7719-7726.

Windhovel, U., B. Geiges, G. Sandmann and P. Boger (1994) Engineering cyanobacterial models resistant to bleaching herbicides. Pest. Biochem. Physiol. **49**, 63-71.

Windhovel, U., B. Geiges, G. Sandmann and P. Boger (1994) Expression of *Erwinia uredovora* phytoene desaturase in *Synechococcus* PCC7942 leading to resistance against a bleaching herbicide. Plant Physiol. **104**, 119-125.

Wise, A., R. Brems, V. Ramakrishnan and M. Villarejo (1996) Sequences in the -35 region of *Escherichia coli rpoS*-dependent genes promote transcription by $E\sigma^s$. J. Bacteriol. **178**, 2785-2793.

Woodall, A.A. (1994) Carotenoids and the protection of membranes against oxidative damage. Ph D Thesis. The University of Liverpool.

Yanisch-Perron, C., J. Vieira and J. Messing (1985) Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**, 103-119.

Yen, H.C. and B. Marrs (1976) Map of genes for carotenoid and bacteriochlorophyll biosynthesis in *Rhodopseudomonas capsulata*. J. Bacteriol. **126**, 619-629.

Yokoyama, A. and W. Miki (1995) Composition and presumed biosynthetic pathway of carotenoids in the astaxanthin-producing bacterium *Agrobacterium aurantiacum*. FEMS Microbiol. Letts. **128**, 139-144.

Young, D.A., C.E. Bauer, J.A.C. Williams and B.L. Marrs (1989) Genetic evidence for superoperonal organization of genes for photosynthetic pigments and pigment-binding proteins in *Rhodobacter capsulatus*. Mol. Gen. Genet. **218**, 1-12.

Young, D.A., M.B. Rudzik and B.L. Marrs (1992) An overlap between operons involved in carotenoid and bacteriochlorophyll biosynthesis in *Rhodobacter capsulatus*. FEMS Microbiol. Lett. **95**, 213- 218.

Zeigler, R.G. (1989) A review of epidemiologic evidence that carotenoids reduce the risk of cancer. J. Nutr. **119**, 116-122.

Zeigler, R.G. (1993) Carotenoids, cancer and clinical trials. Ann. N. Y. Acad. Sci. **691**, 110-119.

Zeigler, R.G., A.F. Subar, N.E. Craft, G. Ursin, B.H. Patterson and B.I. Graubard (1992) Does β-carotene explain why reduced cancer risk is associated with vegetable and fruit intake. Cancer Res. Suppl. **52**, 2060s-2066s.

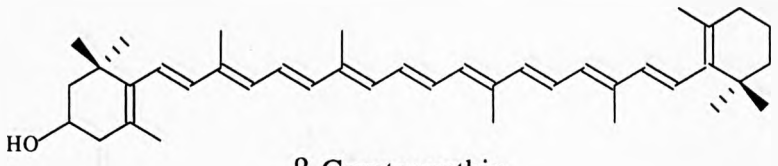
Zhang, L.X., R.V. Cooney and J.S. Bertram (1991) Carotenoids enhance gap junctional communication and inhibit lipid peroxidation in C3H/10T1/2 cells: relationship to their cancer chemopreventitive action. Carcinogenesis **12**, 2109-2114.

Zhang, L.X., R.V. Cooney and J.S. Bertram (1992) Carotenoids up-regulate *Cinnexin43* gene expression independant of their provitamin A or antioxidant properties. Cancer Res. **52**, 5707-5712.

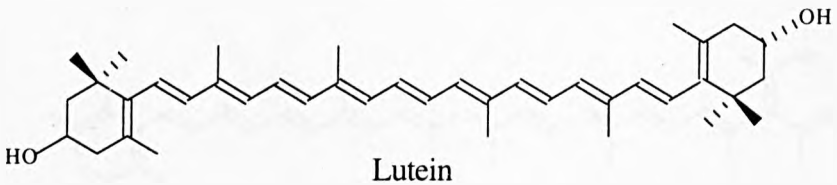
Zhu, Y.S., D.N. Cook, F. Leach, G.A. Armstrong, M. Alberti and J.E. Hearst (1986) Oxygen regulated mRNA's for light-harvesting and reaction center complexes and for bacteriochlorophyll and carotenoid biosynthesis in *Rhodobacter capsulatus* during the shift from anaerobic to aerobic growth. J. Bacteriol. **168**, 1180-1188.

Zhu, Y.S. and J.E. Hearst (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. Proc. Natl. Acad. Sci. USA **83**, 7613-7617.

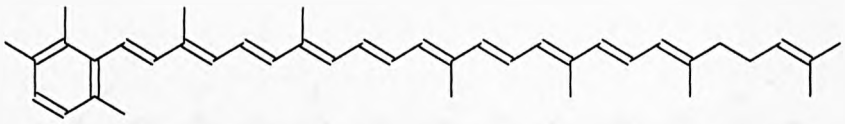
Zsebo, K.M. and J.E. Hearst (1984) Genetic-physical mapping of a photosynthetic gene cluster from *Rhodobacter capsulata*. Cell **37**, 937-947.



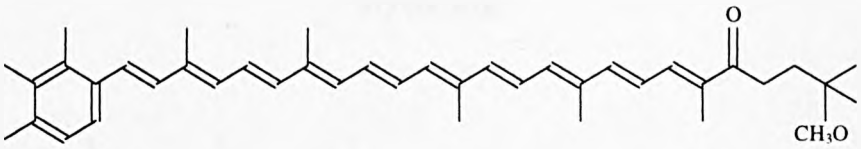
β-Cryptoxanthin



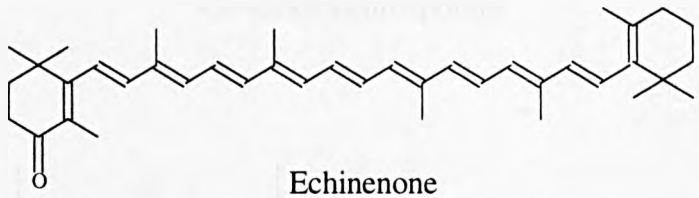
Lutein



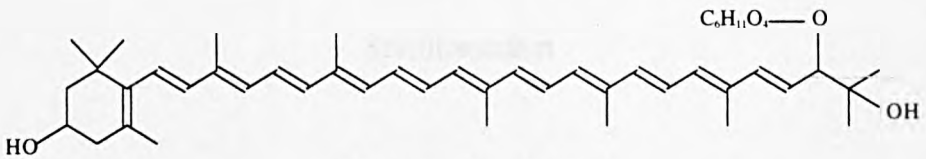
Chlorobactene



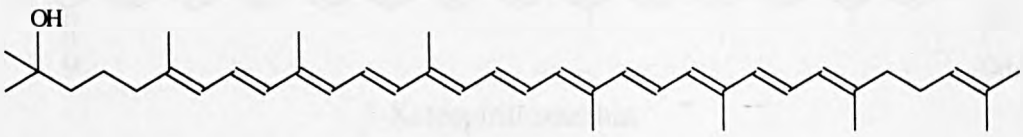
Okenone



Echinenone

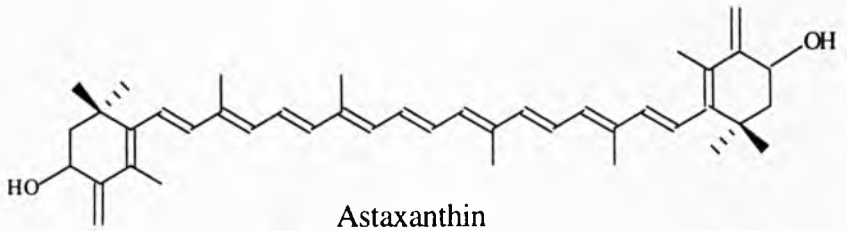


Myxoxanthophyll

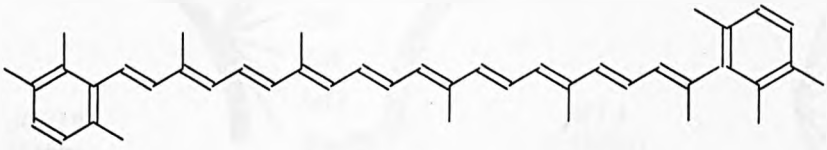


Rhodopin

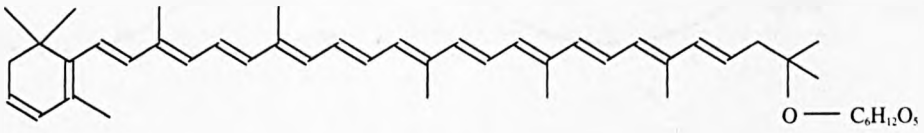
Appendix I Structures of carotenoids referred to in the thesis.



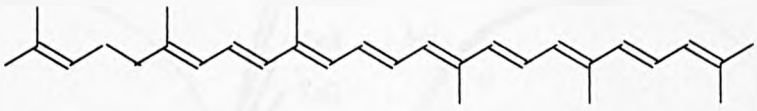
Astaxanthin



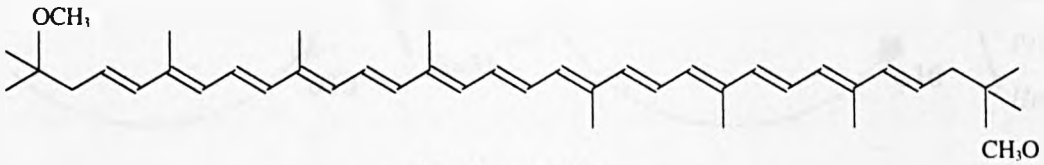
Leprotene



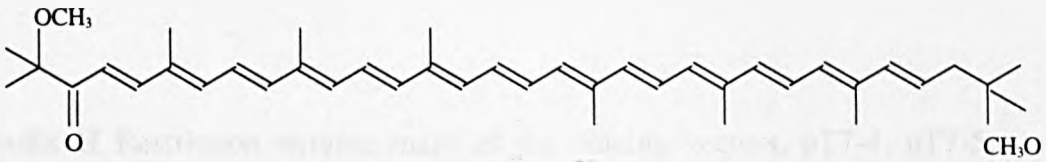
Myxobactin



4,4'-Diaponeurosporene

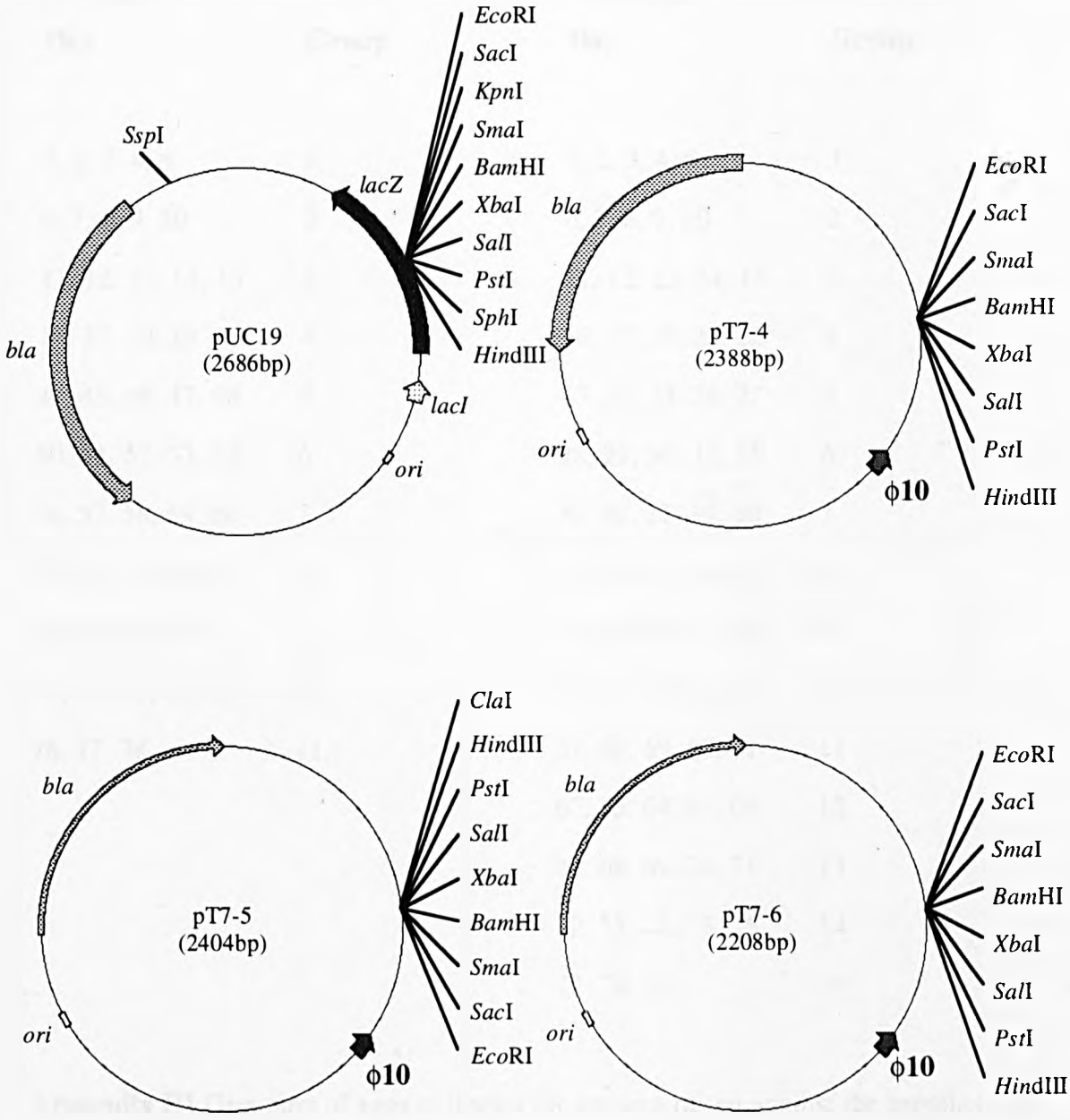


Spirilloxanthin



2-Ketospirilloxanthin

Appendix I (continued)



Appendix II Restriction enzyme maps of the cloning vectors, pT7-4, pT7-5, pT7-6 and pUC19

CrtI eggs		CrtY eggs	
Day	**Group**	**Day**	**Group**
1, 2, 3, 4, 5	1	1, 2, 3, 4, 5	1
6, 7, 8, 9, 10	2	6, 7, 8, 9, 10	2
11, 12, 13, 14, 15	3	11, 12, 13, 14, 15	3
16, 17, 19, 29	4	16, 17, 18, 20, 22	4
44, 45, 46, 47, 48	5	23, 24, 25, 26, 27	5
50, 51, 52, 53, 55	6	28, 29, 30, 31, 33	6
56, 57, 58, 59, 60	7	34, 36, 37, 39, 40	7
61, 62, 63, 64, 65	8	41, 42, 43, 44, 45	8
66, 67, 68, 69	9	46, 47, 48, 49, 50	9
71, 72, 73, 74, 75	10	51, 52, 53, 54, 55	10
76, 77, 78	11	57, 58, 59, 60, 61	11
		62, 63, 64, 65, 66	12
		67, 68, 69, 70, 71	13
		72, 73, 74, 75, 76	14
		77, 78, 79	15

Appendix III Grouping of eggs collected for antisera raised against the proteins CrtI and CrtY. Injection of antigen was performed at 5 and 27 days. Missing numbers represent days when no egg was produced.