THE GENE ENCODING THE GLYPHOSATE-TOLERANT EPSP SYNTHASE FROM *Anabaena variabilis*

Gillian Morag Muir

This thesis is submitted to the University of Glasgow for the degree of Doctor of Philosophy.

Division of Biochemistry and Molecular Biology,

Institute of Biomedical and Life Sciences,

July, 1996.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

To my Mother and Father, my sister Carole and

David for their encouragement and support.

Acknowledgements

There are a great many people to whom I am deeply indepted.

I must thank my supervisors, Professor John Coggins and Dr. Nigel Kerby, for their guidance and support. I am extremely grateful to Dr. Ian Anton, an excellent tutor, who introduced me to the practicalities of molecular biology. My thanks must also be extended to Ranjit Deka and Malcolm Horsburgh for their help and counsel. I am very thankful to Shriharsa Pradhan for his excellent words of advice on matters concerning science, life, the universe and everything. I would like to extend my gratitude to Professor Roger Adams for taking the time to proof read my thesis and to Dr Ian Hamilton for his advise.

I would like to thank the AFRC for their financial support.

Many thanks to all those in the Biochemistry Department for their support and good humour especially: Fiona Gillies, Graeme Thomson, Will Goodwin, Tim Sawbridge, Nigel Urwin, Colin Johnson (Chromatin Man), Sally Martin, Saiqa Khan, Tino Krell, Sureka Chackrewarthy, David Gourley, David Gilloly and Andrew Elwell (for his computer skills). A big thank you goes to John Greene for his patience, excellent technical support and his great sense of humour.

I am extremely grateful to my Mother and Father and my sister Carole for their unending support and financial contributions. Special thanks go to David who has been a constant source of encouragement. I am so grateful to him for everything.

Abbreviations

ATP	adenosine triphosphate
bp	base pair
CIAP	calf intestinal alkaline phosphatase
cpm	counts per minute
(k)Da	(kilo) Daltons
DAHP	3-deoxy-D-arabino-heptulosonate-7-phosphate
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
ds	double stranded
Ds-Co	cobalt requiring DAHP synthase
DTNB	5, 5'-dithio-bis (2-nitrobenzoic acid)
E. Coli®	*Epicurian Coli*®
EDTA	ethylene diamine tetra acetate
EPSP	5-enolpyruvylshikimate-3-phosphate
%GC	percentage of guanine and cytosine
GCG	genetics computer group
genembl	GenBank and European Molecular Biology Laboratory DNA data bases
IPTG	isopropyl-β-D-thiogalatoside
kb	kilobase pair
LB	Luria-Bertani medium
l.m.p.	low melting point
MOPS	3-[N-morpholino] propane sulphonic acid
ODx	optical density at x nm measured with a 1 cm path
PEG	polyethylene glycol
PEP	phosphoenolpyruvate
PFU	plaque forming units

phe	phenylalanine
Pi	inorganic phosphate
PMG	*N*-phosphonomethyl glycine (or glyphosate)
®	registered trade mark
RF	replicative form
RNA	ribonucleic acid
rpm	revolutions per minute
SDS	sodium dodecyl sulphate
S3P	shikimate-3-phosphate
ss	single stranded
Taq	*Thermus aquaticus*
TEMED	N,N,N'N'-tetramethylethylene diamine
Tris	tris (hydroxymethyl) aminomethane
trp	tryptophan
tyr	tyrosine
U.V.	ultra violet
X-gal	5-bromo-4-chloro-3-indolyl-β-galactoside

Contents

Dedication i

Acknowledgments ii

Abbreviations iii

Contents v

List of Figures xiv

List of Tables xvi

Summary xvii

Chapter 1 Introduction 1 - 38

1.1 Cyanobacteria 1

1.2 The Shikimate Pathway 4

1.2.1 Organisation in Prokaryotes 7

1.2.2 Organisation in Fungi, Yeast and Algae 7

1.2.3 Organisation in Higher Plants 9

1.3 The Pathways to Phenylalanine and Tyrosine 10

1.4 Tryptophan Biosynthesis 12

1.5 Aromatic Amino Acid Biosynthesis in Cyanobacteria 12

1.6 The Herbicide Glyphosate 17

1.6.1 General Properties 17

1.6.2 The Mode of Action of Glyphosate 17

1.7 5-Enolpyruvylshikimate-3-phosphate (EPSP) Synthase 18

1.7.1 The Reaction Mechanism of EPSP Synthase 18

1.7.2 Amino Acids Involved at the Active Site of EPSP Synthase 22

1.7.3 Purification of EPSP Synthases from a Diverse Range of Sources 26

1.7.4 Isolation of the EPSP Synthase Gene (*aro*A) from a Diverse Range of Sources 27

1.8 Genetic Engineering of Glyphosate Tolerance 30

1.9 Cyanobacteria and Herbicides 33

1.9.1 General Aspects 33

1.9.2 Glyphosate Tolerance in Cyanobacteria 34

1.10 Aim of this Project 37

Chapter 2 Materials and Methods 39 - 75

2.1 Materials 39

2.1.1 Chemicals and Biochemicals 39

2.1.2 Proteins and Enzymes 39

2.1.3 Chromatography Matrices 40

2.1.4 Bacterial Strains and Cloning Vectors 40

2.2 Growth Media 40

2.2.1 Rich Media 40

2.2.2 Minimal Medium 40

(a) 2 x M9 Salts 44

(b) Minimal Medium Agar 44

(c) Concentrated Mix 44

(d) Amino Acid Supplements For Minimal Medium (2 x concentrated) 44

2.2.3 Media Supplements 44

(a) Antibiotic Supplements: (i) Ampicillin, (ii) Tetracycline 44

(b) *lac* Operon Inducer: Isopropyl-β-D-thiogalactopyranoside (IPTG) 45

(c) Chromogenic Substrate : X-gal
(5-bromo-4-chloro-3-indolyl-β-galactoside) 45

2.3 Growth of Axenic *Anabaena variabilis* ATCC 29413 Cultures 45

2.3.1 Culturing *Anabaena variabilis* 45

2.3.2 Examining *Anabaena variabilis* Cultures for Contamination 46

2.4 General Methods 46

2.4.1 pH Measurement 46

2.4.2 Quantitation of DNA 47

2.4.3 Storage of Bacterial Strains 47

2.5 Small Scale Plasmid Preparation 47

2.5.1 Modified Holmes and Quigley Method 47

2.5.2 Promega's Wizard Plasmid Purification Kit 48

2.6 Large Scale Plasmid/ds M13 Phage DNA Preparation 48

2.6.1 Alkaline Lysis (modified Birnboim and Doly) Method 48

2.6.2 Qiagen Plasmid Preparations 49

2.7 Purification of High Molecular Weight *Anabaena variabilis* Genomic DNA 49

2.8 Caesium Chloride (CsCl)/Ethidium Bromide Density Gradient Centrifugation 50

2.9 Digestion of DNA with Restriction Enzymes 51

2.9.1 Restriction Digestion of Genomic DNA 51

2.9.2 Partial Digestion of Genomic DNA with *Sau*3AI 51

2.9.3 Restriction Digestion of PCR products, Plasmid or Lambda DNA 52

2.9.4 Double Digestions 52

2.10 Phosphatasing DNA with Calf Intestinal Alkaline Phosphatase (CIAP) 52

2.11 Agarose Gel Electrophoresis of DNA 53

2.12 Extraction and Purification of DNA 53

2.12.1 Preparation of Organic Reagents 53

(a) Phenol 53

(b) Chloroform 53

(c) Phenol / Chloroform (1:1) 54

2.12.2 Purification of DNA by Extraction with Phenol and Chloroform 54

2.12.3 Purification of DNA from Low Melting Point Agarose Gels 54

2.12.4 Purification of DNA Using Commercial Kits 55

2.13 Synthesis of Oligonucleotides 55

2.14 Polymerase Chain Reaction 55

2.14.1 PCR Amplification of DNA 55

2.14.2 Purification and Analysis of PCR Products 56

2.14.3 Sub Cloning PCR Products 56

2.15 Creation of an *Anabaena variabilis* Genomic DNA Library in pBluescript SK- 57

2.16 Complementation 58

2.17 Transformation of *E. coli* with Plasmid or M13 Bacteriophage DNA 58

2.17.1 Preparation of Competent Cells 58

(a) $CaCl_2$ Method 59

(b) TFB Method (Modified Hanahan Method) 59

2.17.2 Transformation of Competent Cells 60

2.17.3 Plating Transformed Cells 60

(a) Plating Cells Transformed with M13 60

(b) Plating Cells Transformed with Plasmid DNA 60

2.17.4 Selection of Recombinant Plasmid and M13 Vectors 61

2.18 Creation of an *Anabaena variabilis* Genomic Library in Lambda FIX II 61

2.18.1 Preparation of *Anabaena variabilis* Genomic DNA 61

2.18.2 Quantitation of *A. variabilis* Genomic DNA 62

2.18.3 Ligation of *A. variabilis* Genomic DNA to Lambda FIX II 63

2.18.4 Packaging Reactions 63

2.19 Titre Determination of Phage Genomic Library 63

2.19.1 Preparation of Host Cells for Infection with Phage Library 63

2.19.2 Infection of *E. coli* XL1-Blue MRA (P2) with Phage Library 64

2.19.3 Selection of Lambda FIX II Recombinants 64

2.20 Amplification of Phage Genomic Library 64

2.21 Radiolabelling ds DNA Probes with [α-^{32}P]dCTP 65

2.22 Purification of Radiolabelled DNA Probes and Measuring their Specific Activity 66

2.23 Southern Blotting 66

2.24 Southern Hybridisation of Radiolabelled DNA Probes to Membrane Bound DNA 67

2.25 Removal of Bound Radiolabelled Probe from Filters 68

2.26 Colony Hybridisation 68

2.26.1 Colony Lysis and Binding DNA to Nylon Membranes 68

2.26.2 Hybridising Radiolabelled DNA Probes to Plasmid Library 69

2.26.3 Washing Procedure 69

2.26.4 Purification of Putative Positive Clones from Plasmid Library 70

2.27 Plaque Hybridisation 70

2.27.1 Transfer of DNA to Nylon Membranes 70

2.27.2 Hybridisation of Radiolabelled DNA Probe to Phage Library 71

2.27.3 Purification of Putative Positive Clones from the Phage Library 71

2.27.4 Purification of Lambda Phage DNA 72

2.28 Preparation of Recombinant Vectors for DNA Sequencing 72

2.28.1 Preparation of Single Stranded M13 Template DNA 72

2.28.2 Preparation of ds Template DNA for Sequencing 73

2.29 Dideoxy DNA Sequencing 73

2.29.1 Annealing Reaction 73

(a) Annealing Primer to ss M13 Template DNA 73

(b) Annealing Primer to Denatured ds Template DNA 74

2.29.2 Labelling Reaction 74

2.29.3 Termination Reaction 74

2.29.4 DNA Sequencing Gel Electrophoresis 74

2.30 Computer Programmes for the Analysis of DNA Sequences 75

Chapter 3 Amplification of *Anabaena variabilis* Genomic DNA by PCR 76 - 101

3.1 Introduction 76

3.2 The Polymerase Chain Reaction 76

3.3 Design of Degenerate PCR Primers 77

3.4. Strategy for PCR amplification of *A. variabilis* Genomic DNA 78

3.4.1 Design of Degenerate *aro*A Primers 78

3.4.2 PCR Amplification of *A. variabilis* Genomic DNA Using Primers 1 and 2 78

3.4.3 Cloning PCR Amplified Products 83

3.4.4 Attempts to Clone the Full Length, PCR Amplified *aro*A Fragment 84

3.5 Southern Blots of *Anabaena variabilis* Genomic DNA 86

3.6 New Strategy for PCR Amplification of an *Anabaena variabilis aro*A Gene Fragment. 90

3.6.1 PCR Amplification of Uncontaminated *Anabaena variabilis* Genomic DNA with Degenerate Primers 90

3.6.2. Cloning and DNA Sequencing of the 900-bp PCR Product Amplified with Primers 3 and 4 93

3.7 Southern Blot of *Anabaena variabilis* Genomic DNA Using 900-bp PCR Product as Probe 95

3.8 Discussion 96

3.8.1 PCR and Contamination 96

3.8.2 Cleavage of the 800-bp PCR Amplified Product 99

3.8.3 Primer Design 100

Chapter 4 Hybridisaton of the *Synechocystis aro*A Probe to an *Anabaena variabilis* Genomic DNA Library 102 - 123

4.0 Common Introduction to Chapters 4, 5 and 6 102

4.1 Heterologous Hybridisation 103

4.2 Strategy for Isolating the *Anabaena variabilis aro*A Gene By Heterologous Hybridisation 104

4.3 Southern Hybridisation of the *Synechocystis* sp. PCC 6803 *aro*A Probe to *Anabaena variabilis* Genomic DNA 104

4.3.1 Preparation of the *Synechocystis* sp. PCC 6803 *aro*A Probe 104

4.3.2 Optimal Conditions for Hybridisation of *Synechocystis* sp. PCC 6803 *aro*A Probe to *Anabaena variabilis* Genomic DNA 105

4.3.3 Binding Specificity of the *Synechocystis* sp. PCC 6803 *aro*A Probe to *Anabaena variabilis* Genomic DNA 108

4.4 Construction of the *Anabaena variabilis* Genomic DNA Library for Heterologous Screening 111

4.4.1 Preparation of *Anabaena variabilis* Genomic DNA for Library Construction 111

4.4.2 Ligations of *Anabaena variabilis* genomic DNA into pBluescript SK- 112

4.5 Evaluation of the *Anabaena variabilis* Genomic DNA Library in pBluescript SK- 116

4.5.1 Preparation of the *Anabaena variabilis vnf*DGK Probe 116

4.5.2 Colony Hybridisation of the Homologous *vnf*DGK Probe to the *Anabaena variabilis* Genomic DNA Library 117

4.6 Screening the *Anabaena variabilis* Genomic DNA Library in pBluescript with the *Synechocystis aro*A Probe 117

4.7 Discussion 122

Chapter 5 Screening an *Anabaena variabilis* Genomic DNA Library with *aro*A Probes from Different Sources 124 - 147

5.1 Introduction 124

5.2 Strategy for the Construction of the *Anabaena variabilis* Genomic DNA Library in Lambda FIX® II 124

5.3 Reasons for Using the Replacement Vector, Lambda FIX® II 126

5.4 Southern Hybridisation of the *Synechocystis aro*A Probe to *Sau*3AI Digested *Anabaena variabilis* Genomic DNA 127

5.5 Construction of the *Anabaena variabilis* Genomic DNA Library in Lambda FIX® II 129

5.5.1 Preparation of *Anabaena variabilis* Genomic DNA for Ligation into Lambda FIX® II 129

5.5.2 Ligation of *Sau*3AI Cut, Partially Filled-In *Anabaena variabilis* DNA to Lambda FIX® II 132

5.5.3 Amplification of the *Anabaena variabilis* Genomic DNA Library in Lambda FIX® II 133

5.6 The Quality of the Amplified *Anabaena variabilis* Genomic DNA Library in Lambda FIX® II 134

5.7 The Heterologous Probes Used to Screen the Amplified *Anabaena variabilis* Genomic DNA Library in Lambda FIX® II 137

5.7.1 Preparation of the *Synechocystis*, *E. coli* and Pea *aro*A Probes 139

5.7.2 Screening the Amplified *Anabaena variabilis* Genomic DNA Library in Lambda FIX® II with the *Synechocystis*, *E. coli* and Pea *aro*A Probes 140

5.7.3 Analysis of the Purified Recombinant Clones Isolated from the *Anabaena variabilis* Library 141

5.8 Discussion 144

5.8.1 Overview of Phage Library Construction and Screening Procedures 144

5.8.2 Justification for Using the *Synechocystis*, *E. coli* and Pea *aro*A Genes as Heterologous Probes 145

5.8.3 Reasons Why the *Anabaena variabilis aro*A Gene was not Isolated from the Phage Library 147

Chapter 6 Screening an *Anabaena variabilis* Genomic DNA Library by Phenotypic Complementation 148 - 158

6.1 Complementation 148

6.2 Strategy for Isolating the *A. variabilis aro*A Gene by Complementation 148

6.3 Construction of an *Anabaena variabilis* Genomic DNA Library in pBluescript SK- for Complementation 149

6.3.1 Preparation of *A. variabilis* Genomic DNA for Library Construction 149

6.3.2 Ligation of Size Fractionated *Eco*RI Digested *Anabaena variabilis* DNA into pBluescript SK- 151

6.3.3 Quality of the *A. variabilis Eco*RI Genomic DNA Library 151

6.4 Complementation of the Auxotrophic Mutant, *E. coli* AB2829, with the Cloned EPSP Synthase Gene of *E. coli* K12 152

6.5 Complementation of *E. coli* AB2829 with the *Anabaena variabilis Eco*RI Genomic DNA Library 153

6.6 Discussion 154

Chapter 7 General Discussion 159 - 164

7.1 The Objectives of this Work 159

7.2 Alternative Strategies for Cloning the *Anabaena variabilis aro*A Gene 160

7.3 Reasons for Cloning the *Anabaena variabilis aro*A Gene 162

Chapter 8 References 165 - 181

List of Figures

1.1 The shikimate pathway. 5

1.2 Some of the aromatic compounds derived from chorismate. 6

1.3 Organisaton of the five central enzymes of the shikimate pathway. 8

1.4 The pathways leading to biosynthesis of phenylalanine and tyrosine from chorismate. 11

1.5 Pathway to tryptophan from chorismate. 13

1.6 Regulation of aromatic amino acid biosynthesis in cyanobacteria. 15

1.7 The structural formula for glyphosate. 19

1.8 The EPSP synthase reaction. 19

1.9 The proposed kinetic mechanism of the *Neurospora crassa* arom EPSP synthase (Boocock and Coggins, 1983). 21

3.1 Strategy for PCR amplification of genomic DNA. 79

3.2 Design of degenerate primers 1 and 2 for the amplification of a segment of the *Anabaena variabilis aro*A gene. 80

3.3 PCR of *Anabaena variabilis* genomic DNA with degenerate primers 1 and 2. 82

3.4 EPSP synthase sequence alignment. 85

3.5 Southern hybridisation of the 320-bp putative *aro*A gene fragment to 7 µg *A. variabilis Hin*dIII digested genomic DNA. 88

3.6 Southern hybridisation of *Rhodobacter sphaeroides* LCHII gene to *Rb. sphaeroides* genomic DNA. 89

3.7 The degenerate primers 3 and 4 used for the amplification of an *Anabaena variabilis aro*A segment. 91

3.8 PCR of uncontaminated *Anabaena variabilis* genomic DNA with degenerate primers 3 and 4. 94

3.9 Southern hybridisation of the (A) *Synechocystis aro*A probe and (B) 900-bp PCR product to 2 µg *Anabaena variabilis* genomic DNA. 97

4.1 The positions of the *Synechocystis aro*A probes relative to the full length EPSP synthase gene. 109

4.2 Hybridisation of different sized *Synechocystis aro*A gene fragments to *Anabaena variabilis* genomic DNA. 110

4.3 Autoradiogram illustrating hybridisation of the *Anabaena variabilis vnf*DGK probe to replicate membranes carrying approximately 20,000 clones of the *A. variabilis* genomic DNA library in pBluescript (grown in *E. coli*® XL2-Blue MRF'). 118

4.4 Autoradiogram showing hybridisation of the *Synechocystis* 618-bp *aro*A probe to duplicate membranes containing 500 clones of the *Anabaena variabilis* genomic DNA library in pBluescript (grown in *E. coli*® XL2-Blue MRF'). 121

5.1 DNA library construction in lambda FIX® II. 125

5.2 Hybridisation of the *Synechocystis aro*A probe to *Sau*3AI digested *Anabaena variabilis* genomic DNA. 128

5.3 Preparation of *Anabaena variabilis* genomic DNA for phage library construction. 130

5.4 Restriction analysis of randomly selected clones from the amplified *Anabaena variabilis* library in lambda FIX II. 136

5.5 Autoradiogram resulting from hybridisation of the *Anabaena variabilis vnf*DGK probe to replicate membranes carrying approximately 50,000 clones of the *A. variabilis* genomic DNA library in lambda FIX II (grown in *E. coli*® XL1-Blue MRA (P2)). 138

5.6 Restriction analysis of purified clones isolated from the *Anabaena variabilis* library using the *E. coli aro*A probe. 142

6.1 Southern hybridisation of the *Synechocystis* 618-bp *aro*A probe to *Anabaena variabilis* genomic DNA. 150

List of Tables

1.1 Sequence comparison of the highly conserved region of EPSP synthases from various sources. 25

1.2 Kinetic parameters of three purified EPSP synthase enzymes. 27

1.3 Comparison of some kinetic and physical properties of wild type and mutant EPSP synthases isolated from various sources. 36

2.1 Bacterial strains used for molecular cloning experiments. 41

2.2 Phagemid and phage vectors used for molecular cloning experiments. 42

2.3 Rich Growth Media. 43

2.4 Composition of BG11o medium. 46

2.5 Reaction conditions used for PCR. 56

3.1 PCR Primers for amplification of *Anabaena variabilis aro*A gene fragment. 92

4.1 Number of recombinant clones in each *A. variabilis* test library made in pBluescript SK-. 113

4.2 The number of recombinant clones to be screened to have a 99% probability of isolating the *A. variabilis aro*A gene. 115

Summary

Glyphosate is a broad-spectrum, non-selective, post-emergence herbicide, active against a variety of weed and crop species. The primary site for herbicidal action is 5-enolpyruvylshikimate-3-phosphate (EPSP) synthase, the penultimate enzyme of the shikimate pathway. In higher plants and microorganisms, this pathway provides a metabolic route to the synthesis of the aromatic amino acids - phenylalanine, tyrosine and tryptophan - plus other aromatic compounds. EPSP synthase has been purified from various sources to investigate its kinetic characteristics and the inhibition properties of glyphosate.

The cyanobacteria are the largest, most diverse and widely distributed group of photosynthetic prokaryotes. The physiological and biochemical effects of glyphosate on certain cyanobacterial strains have been examined. The filamentous, nitrogen fixing strain *Anabaena variabilis* ATCC 29413, like other cyanobacteria, is tolerant to glyphosate. Tolerance is due to an EPSP synthase that is uninhibited by the herbicide. A detailed kinetic and molecular study of this enzyme has been hindered by the consistently low yield of EPSP synthase protein purified from *A. variabilis* cells. As such, the purpose of the work described in this thesis was to isolate the *A. variabilis* EPSP synthase gene (termed *aro*A) and overexpress the encoded enzyme. A number of cloning methods were undertaken in an attempt to reach this goal.

The polymerase chain reaction was employed to amplify a defined segment of the *aro*A gene from *A. variabilis* genomic DNA using different sets of degenerate oligonucleotide primers. These primers were designed from conserved regions of EPSP synthase sequence from various plants and microorganisms. A truncated fragment of an *aro*A gene was synthesised, however, this piece of DNA did not originate from *A. variabilis* DNA. The source of this contaminating PCR product has yet to be identified. The design of the PCR primers may have affected the specificity of the amplification reaction and could explain why other primer sets failed to amplify the sequence of interest.

Subsequently, a library of *A. variabilis* genomic DNA was constructed in the phagemid vector, pBluescript SK-. After evaluating its size and quality, the library was screened with an *aro*A probe from the unicellular, non-nitrogen fixing cyanobacterium *Synechocystis* sp. PCC 6803 to isolate the clone of interest. Prior to screening, Southern blot experiments had demonstrated that the *Synechocystis* probe hybridised to specific fragments of restriction digested *A. variabilis* DNA. This heterologous probe was, therefore, considered suitable for screening purposes. Persistent problems with non-specific hybridisation between the probe and the genetic material of the host cell harbouring the library frustrated the attempts made to locate an *aro*A clone. Reducing the level of background hybridisation required a slightly different approach. It was established that plaque hybridisation was more sensitive than colony hybridisation. Consequently, another *A. variabilis* library was made, but on this occasion the phage, lambda FIX® II, was used as the vector system. Heterologous *aro*A probes from *Synechocystis*, *E. coli* and pea were employed to maximise the possibility of finding the *A. variabilis* EPSP synthase clone. Despite control experiments signifying the good quality of the newly constructed library, not one of the heterologous probes pulled out the clone of interest.

The final strategy involved isolating the *A. variabilis* EPSP synthase gene by phenotypic complementation of an *aro*A$^-$ auxotrophic mutation of the *E. coli* strain AB2829. A control experiment showed that expression of the *E. coli* EPSP synthase from cloned DNA complemented the deficiency of the host and enabled the mutant to grow on selective medium. The *E. coli* mutant was then transformed with a library of *A. variabilis* genomic DNA made from *Eco*RI restriction fragments thought to contain the entire *aro*A coding sequence. Succeeding experiments showed that such a library could not complement the *aro*A mutation of *E. coli* AB2829.

The advantages and disadvantages of each of the above techniques are discussed in detail with specific regard to cloning the *A. variabilis* EPSP synthase gene. Other gene cloning strategies not used in this work are described and the practical reasons for not employing these techniques are debated. The possible exploitation of

the EPSP synthase gene from *A. variabilis* by genetic engineers for the construction of a glyphosate-tolerant crop plant is discussed, as is the contribution the gene sequence could have made to the debate regarding the cyanobacterial origin of higher plant and algal plastids.

Chapter 1

Introduction

1.1 Cyanobacteria

Cyanobacteria are an ancient group of eubacteria and are probably the largest, most diverse and widely distributed group of prokaryotes (see Stanier and Cohen-Bazire, 1977; Carr and Whitton, 1982; Packer and Glazer, 1988 for reviews). The earliest forms appeared in the fossil record approximately 2,250 million years ago (the Precambrian era) and are believed to have introduced molecular oxygen into the Precambrian environment (Doolittle, 1982). Unicellular and filamentous species exist and live in a range of different habitats. Terrestrial, marine and freshwater species have been found and strains also occur in salt marshes, hot and sulphurous springs, soda lakes and polar ice caps. The cyanobacteria have evolved in order to inhabit such diverse ecological niches. Although predominantly photoautotrophic, some strains can grow photoheterotrophically and chemoheterotrophically using glucose, fructose, ribose or glycerol as a carbon source (Rippka *et al.*, 1979). In the absence of an available combined nitrogen source, some species can reduce atmospheric nitrogen in specialised cells of filamentous strains called heterocysts or in vegetative cells of non-heterocystous strains (Stanier and Cohen-Bazire, 1977). These adaptive processes are thought to involve changes in gene expression at the transcriptional level. The information regarding transcription in cyanobacteria is limited but is reviewed by Curtis and Martin (1994). The genetic heterogeneity of the cyanobacteria is illustrated by the %GC content of their DNA. This ranges from 35 to 71%, almost as much as the whole bacterial kingdom (Herdman *et al.*, 1979).

The cyanobacteria were originally categorised as blue-green algae primarily due to their plant-like photosynthetic properties. However, the prokaryotic nature of these organisms was recognised with the development of electron microscopy which lead to the classification of living things as prokaryotes or eukaryotes (Stanier and Cohen-Bazire, 1977). Cyanobacteria have characteristic prokaryotic features: instead of a nucleus they contain a single circular chromosome; they have no sub-cellular organelles; their ribosomes are small (70s) and their genetic material is almost free of introns (Barinaga, 1990). The cyanobacteria themselves have also been categorised and

are sub-divided into five taxonomic groups on the basis of their morphology (Rippka *et al.*, 1979). Various biochemical and molecular methods have since been used in an attempt to improve this classification scheme (reviewed in Wilmotte, 1994).

Unlike other photosynthetic bacteria that contain bacteriochlorophyll, the cyanobacteria contain the light harvesting pigment chlorophyll *a* and perform plant-like oxygenic photosynthesis. Chlorophyll *a* and other components of the photosynthetic apparatus (including the carotenoids, the photochemical reaction centre and the photosynthetic electron transport chain) are located in membranous vesicles called the thylakoids (Stanier and Cohen-Bazire, 1977). Cyanobacteria also contain another major group of pigments called the phycobiliproteins. These pigments are divided into three major classes: phycoerythrin, allophycocyanin and phycocyanin, and are located in structures called phycobillisomes adjacent to the thylakoid membranes. During photosynthesis, light energy is initially absorbed by the phycobiliproteins then efficiently transferred to chlorophyll *a* in the thylakoid membrane where chemical energy is generated in the form of ATP (Grossman *et al.*, 1993).

Studies of photosynthesis in cyanobacteria as well as analyses of various protein and nucleic acid sequences have led to many suggestions that these organisms share a common ancestry with the chloroplasts of photosynthetic eukaryotes (reviewed in Douglas, 1994). Extensive analysis of 16S rRNA sequences revealed that higher plant and algal plastids arose from a cyanobacterium or a cyanobacterium-like ancestor (Giovannoni *et al.*, 1988; Douglas and Turner, 1991). These and other studies also proposed that the cyanobacteria had diversified into different species before the plastid progenitor established an endosymbiotic relationship with the precursor plant or algal cell (Nelissen *et al.*, 1995).

One particularly important property of many cyanobacteria is their ability to fix atmospheric nitrogen. This process is performed by the enzyme, nitrogenase, which can only function in anaerobic conditions. Since cyanobacteria are oxygen evolving photosynthesisers, some filamentous species form specialised cells called heterocysts where nitrogen fixation takes place (Stewart *et al.*, 1969). Heterocysts have a thickened

cell envelope and lack photosystem II, the oxygen generating component of the photosynthetic apparatus. Collectively these features help to exclude oxygen from the interior of the heterocyst and create an environment in which the nitrogenase can function. The heterocyst can still generate ATP via photosystem I, but cannot manufacture fixed carbon compounds and must depend on neighbouring vegetative cells for their supply (Wolk, 1982). During nitrogen fixation, nitrogen is reduced to ammonia by the nitrogenase enzyme and assimilated via the amination of glutamate to glutamine by glutamine synthetase. Glutamine is then transported to vegetative cells where glutamate is regenerated by glutamate: 2-oxoglutarate aminotransferase (Meeks *et al.*, 1978). Cyanobacteria form heterocysts only when the supply of combined nitrogen is limiting. The regulation of heterocyst differentiation has, therefore, been studied closely . The *nif* genes, which encode the nitrogenase and other components involved in nitrogen fixation, are rearranged during cellular differentiation (Haselkorn *et al.*, 1986). This is one of only a few known examples of environmentally regulated genome rearrangement. Other unicellular and filamentous cyanobacteria that cannot form heterocysts can still synthesise nitrogenase and, hence, fix nitrogen but only in anaerobic conditions (Rippka and Waterbury, 1977).

A wide range of organisms have formed symbiotic relationships with cyanobacteria to exploit their nitrogen fixing abilities (Whitton and Carr, 1982). Both symbiotic and free-living nitrogen fixing strains are of great practical and economic importance in rice cultivation where they act as biofertilizers in flooded fields (Kerby *et al.*, 1989).

As well as heterocysts, filamentous cyanobacteria can form hormogonia. These are short, undifferentiated, motile filaments formed by fragmentation of the heterocystous filament and are a major means of reproduction in these organisms. Additionally, akinetes frequently occur adjacent to heterocysts in filamentous cyanobacteria. Their role is analogous to bacterial endospores since these specialised cells can survive for many years, resist adverse conditions and germinate to produce

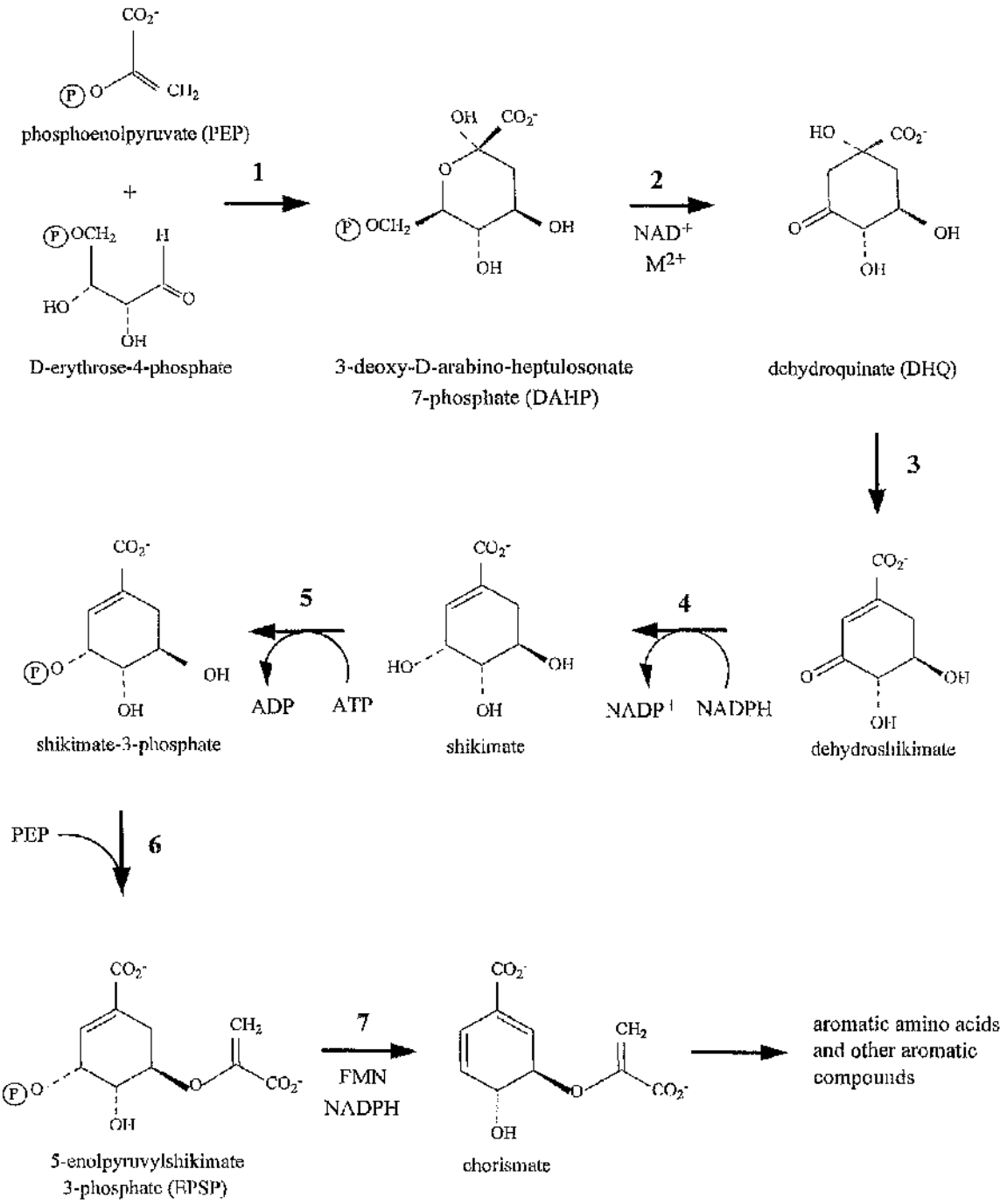
vegetative cells in response to improved environmental conditions (Nichols and Adams, 1982).

Knowledge of photosynthesis, nitrogen metabolism and heterocyst differentiation has improved with the ability to clone genes and analyse their structure and expression. The most widely used method for gene cloning is screening a DNA library with an appropriate heterologous probe. However, antibodies and oligonucleotides can also be used to isolate the gene of interest. Other gene cloning techniques involve complementation of isolated and characterised mutations (Bryant and Tandeau de Marsac, 1988). Gene expression can be measured using reporter systems, particularly *lacZ* and *lux*AB, which can identify environmentally regulated genes. These techniques have been reviewed by Thiel (1994).

Cyanobacteria are potentially of great importance to the biotechnology industry and could be used for the production of food-stuffs, pigments and pharmaceuticals (see Elhai, 1994 for review). The development of strains able to liberate ammonia or other nitrogenous compounds (such as amino acids) are being developed and may increase the value of cyanobacteria in agriculture (Kerby *et al.*, 1989).

1.2 The Shikimate Pathway

In plants and microorganisms the biosynthesis of all aromatic compounds proceeds by way of the shikimate pathway (see Bentley, 1990 and Haslam, 1993 for reviews). Erythrose-4-phosphate and phosphoenolpyruvate (PEP) are converted to chorismate via the intermediate shikimate by a series of seven enzymatic reactions (Figure 1.1). Chorismate is a major metabolic branch point since it is the common precursor for the synthesis of an array of different aromatic compounds (Figure 1.2). The aromatic amino acids, phenylalanine, tyrosine and tryptophan, are the major end products of the shikimate pathway. Animals lack the capacity to synthesise aromatic compounds via this route and must, therefore, obtain aromatic amino acids, folate coenzymes and vitamins from their diet.



(1) DAHP synthase, **(2)** dehydroquinate synthase, **(3)** dehydroquinase, **(4)** shikimate dehydrogenase, **(5)** shikimate kinase, **(6)** EPSP synthase, **(7)** chorismate synthase

Figure 1.1 The shikimate pathway.

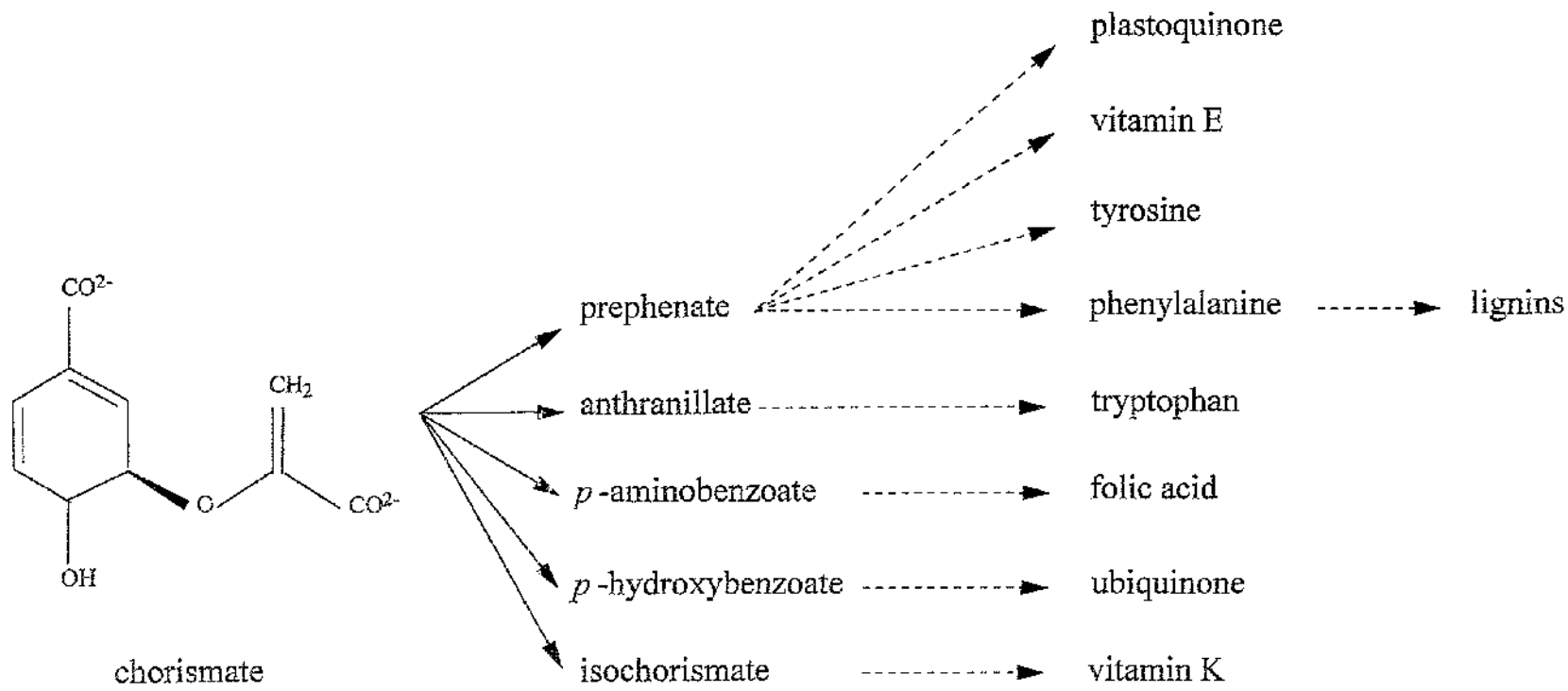


Figure 1.2 Some of the aromatic compounds derived from chorismate.

——► = one step biocemical process, - - -► = multi-step biochemical process.

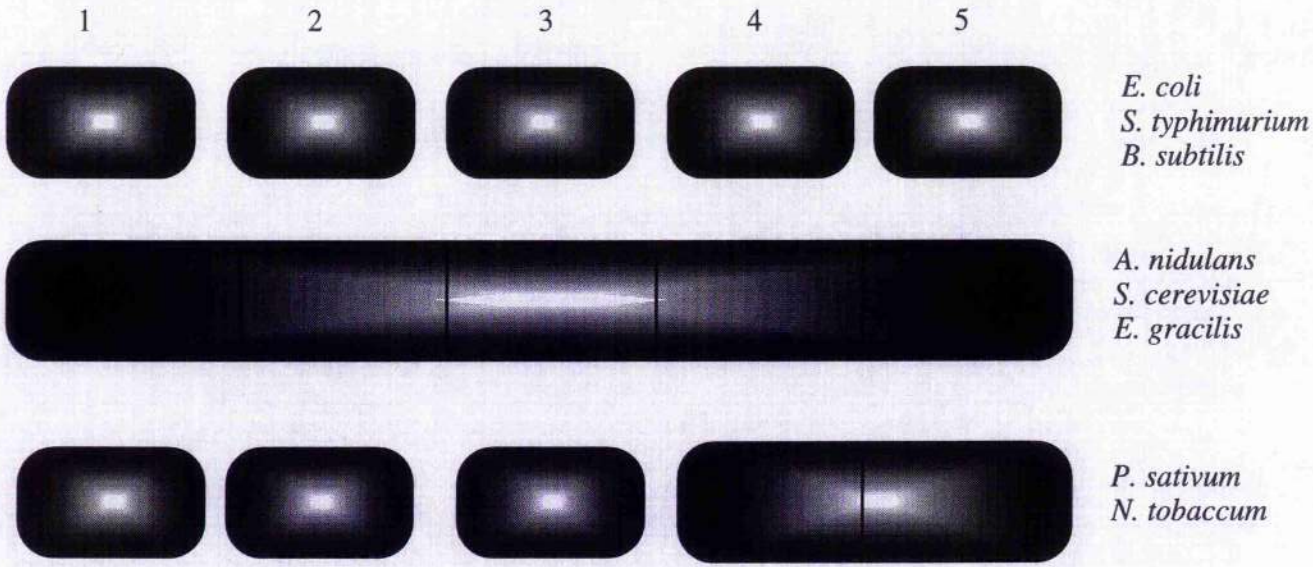
The chemical intermediates between erythrose-4-phosphate and chorismate were first isolated and characterised from microorganisms by Davis, Weiss, Sprinson and Gibson (Davis, 1955; Levin and Sprinson, 1964; Gibson and Pittard, 1968). The route to chorismate is believed to be identical in all plants and microorganisms. There are, however, remarkable differences in the organisation of the shikimate pathway genes and enzymes between different species (illustrated in Figure 1.3).

1.2.1 Organisation in Prokaryotes

In most bacteria including the cyanobacterium, *Anabaena variabilis*, the shikimate pathway enzymes are separable and monofunctional (Berlyn and Giles, 1969; Berlyn *et al.*, 1970; Coggins *et al.*, 1985 and Figure 1.3). The *Bacillus subtilis* strain 168 is the exception to this rule since its DAHP synthase and chorismate mutase activities occur on a single polypeptide (Nakatsukasa and Nester, 1972). The genes of the shikimate pathway (termed *aro*) in *E. coli* (Pittard and Wallace, 1966) and *Salmonella typhimurium* (Gollub *et al.*, 1967) are located at unlinked regions of the chromosome. In contrast, most of the *Bacillus subtilis* shikimate pathway genes are organised as part of a supra-operon with other genes involved in aromatic amino acid biosynthesis (Henner *et al.*, 1986; Henner and Yanofsky, 1993). The genes encoding 5-enolpyruvylshikimate-3-phosphate (EPSP) synthase (*aro*A) and chorismate synthase (*aro*C) have been isolated from the unicellular cyanobacterium, *Synechocystis* sp. PCC 6803 and are located at different regions of the chromosome (Mayes *et al.*, 1993; Schmidt *et al.*, 1993). The *aro*A gene appears to be transcribed singly (dalla Chiesa *et al.*, 1994) whilst *aro*C forms part of an operon with genes encoding ribosomal proteins (Schmidt *et al.*, 1993).

1.2.2 Organisation in Fungi, Yeast and Algae

In 1969, Ahmed and Giles found that the five central enzymes of the shikimate pathway in fungi co-sedimented on a sucrose density gradient. It has since



1 = dehydroquinate synthase, 2 = EPSP synthase,
3 = shikimate kinase, 4 = dehydroquinase,
5 = shikimate dehydrogenase.

Figure 1.3 Organisation of the five central enzymes of the shikimate pathway.

been discovered that these enzyme activities occur on two identical pentafunctional polypeptide chains, termed *arom*, in the fungi *Neurospora crassa* (Lumsden and Coggins, 1977), *Aspergillus nidulans* (Hawkins, 1985) and *Pneumocystis carinii* (Banerji *et al.*, 1993), the yeast *Saccharomyces cerevisiae* (Duncan *et al.*, 1987) and the alga *Euglena gracilis* (Patel and Giles, 1979). The *arom* complex is encoded by a single gene which has been cloned from *A. nidulans* (Charles *et al.*, 1986), *P. carinii* (Banerji *et al.*, 1993) and *S. cerevisiae* (Duncan *et al.*, 1987).

The amino acid sequences of the five central shikimate pathway enzymes from *E. coli* have been compared to the equivalent sequences of the *arom* polypeptide from *P. carinii* (Banerji *et al.* 1993) and *S. cerevisiae* (Duncan *et al.*, 1987). Homologies between the bacterial, fungal and yeast sequences confirmed the order of enzyme activities along the *arom* polypeptide (as shown in Figure 1.3). These analyses also lead to the conclusion that the *arom* complex is the product of gene fusion of ancestral *E. coli*-like genes.

There are a number of hypotheses attempting to explain the advantages of multifunctional proteins. These include: catalytic efficiency, substrate channelling and coordinate expression of enzyme activities. From the available evidence only the last explanation can be applied to the *arom* polypeptide (Coggins *et al.*, 1985; Duncan *et al.*, 1987; Charles *et al.*, 1986; Banerji *et al.* 1993).

The remaining two enzyme activities that do not form part of the *arom* complex, namely DAHP synthase and chorismate synthase, are monofunctional and separable in *N. crassa* (Nimmo and Coggins, 1981; White *et al.*, 1988).

1.2.3 Organisation in Higher Plants

In higher plants, five enzymes of the shikimate pathway can be separated by various chromatographic techniques and are, therefore, monofunctional (Mousdale and Coggins 1984, 1985, 1986 and Figure 1.3). The 3-dehydroquinase and shikimate dehydrogenase activities (responsible for catalysing steps 3 and 4 of the pathway) have been shown to occur on a bi-functional polypeptide in the moss *Physcomitrella patens*

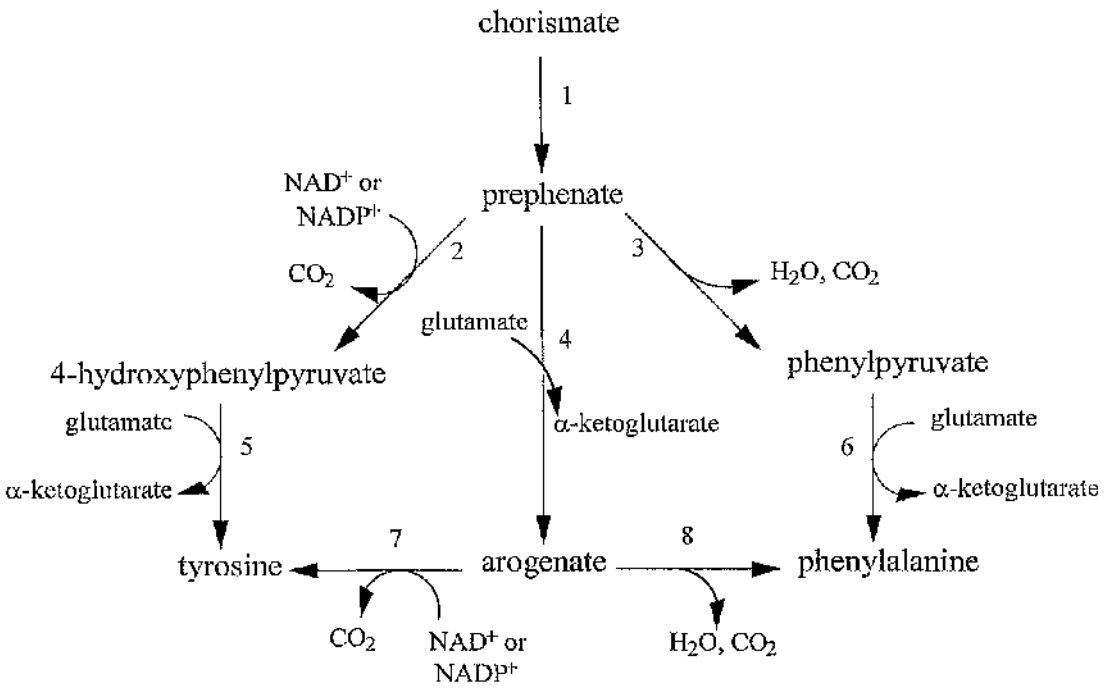
(Polley, 1978), *Phaseolus mungo* seedlings (Koshiba, 1978) and *Pisum sativum* seedlings (Deka *et al.*, 1994). Various cDNAs and genes coding for shikimate pathway enzymes in plants have been isolated and characterised. Schmid and Amrhein (1995) have recently reviewed the molecular organisation of the shikimate pathway in higher plants. This review also confirms the work performed by Mousdale and Coggins (1985 and 1986) that the plastids are the major, if not the only, site of aromatic amino acid biosynthesis in plant cells.

1.3 The Pathways to Phenylalanine and Tyrosine

In all organisms so far studied, the first step leading to the synthesis of phenylalanine (phe) and tyrosine (tyr) is the transformation of chorismate to prephenate catalysed by chorismate mutase (Haslam, 1993). From prephenate the pathways diverge (as illustrated in Figure 1.4) and the biosynthetic routes to phenylalanine and tyrosine vary from species to species.

Both *E. coli* and the cyanobacteria *Anacystis nidulans* and *Agmenellum quadruplicatum* synthesise phenylalanine via phenylpyruvate (Davis, 1953; Byng *et al.*, 1982; Jensen, 1986). However, the former synthesises tyr through 4-hydroxyphenylpyruvate whereas the latter transaminate prephenate to form arogenate, a precursor of tyrosine. Alternatively in higher plants, green algae and *Euglena gracilis*, arogenate is the precursor for both phenylalanine and tyrosine (Bonner *et al.*, 1995). In some pseudomonad species such as *Pseudomonas aeroginosa* both alternative pathways co-exist (Haslam, 1993).

Some of the enzymes involved in the synthesis of these two aromatic amino acids show a lack of substrate specificity (Jensen, 1986; Bentley, 1990). For example, some microbial aminotransferases can transaminate prephenate as well as phenylpyruvate and 4-hydroxyphenylpyruvate. Also, a single dehydratase from *Pseudomonas aeroginosa* can use both prephenate and arogenate as a substrate. The apparent existence of two different routes leading to phenylalanine and tyrosine biosynthesis may, therefore, not reflect the existence of two separate sets of enzymes



(1) chorismate mutase, (2) prephenate dehydrogenase, (3) prephenate dehydratase, (4) prephenate aminotransferase, (5) tyrosine aminotransferase, (6) phenylalanine aminotransferase, (7) arogenate dehydrogenase, (8) arogenate dehydratase

Figure 1.4 The pathways leading to biosynthesis of phenylalanine and tyrosine from chorismate.

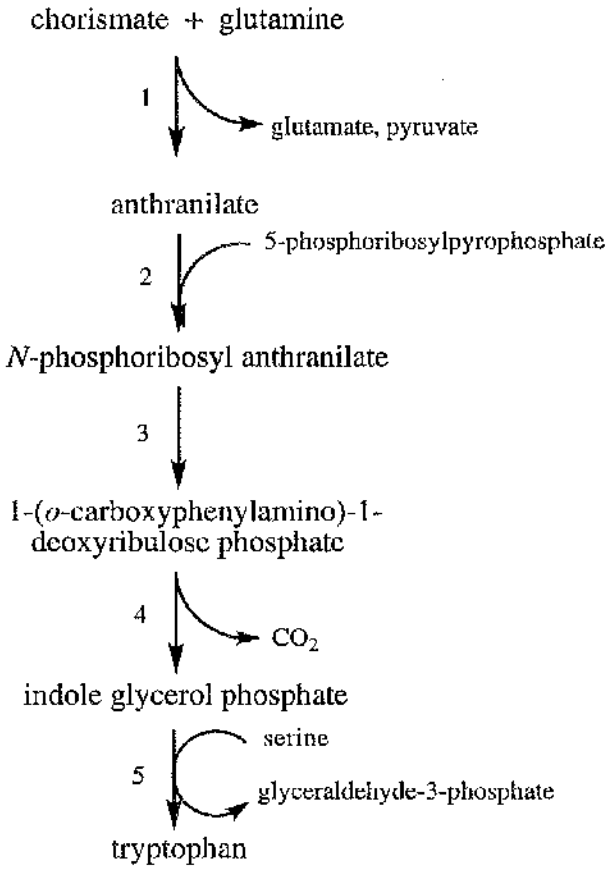
encoded by two sets of genes. However, further investigations in this area are required before definitive answers can be found.

1.4 Tryptophan Biosynthesis

The biochemical steps leading to tryptophan (*trp*) biosynthesis from chorismate are identical in all organisms studied to date, but organisation of the enzymes and genes of the pathway varies from species to species. The five step process requires five enzymatic activities encoded by seven genes (Figure 1.5). In *E. coli*, the *trp* genes are organised into a single operon where *trp*G.D and *trp*C.F are fused. Thus, the tryptophan synthesising enzymes are organised into three complex protein molecules: the anthranilate synthase / phosphoribosyl anthranilate transferase complex (steps 1 and 2); a bi-functional polypeptide carrying the *N*-phosphoribosyl anthranilate isomerase and indole glycerol phosphate synthase activities (steps 3 and 4) and the tryptophan synthase complex (step 5). In contrast to *E. coli*, the *trp* genes of *Saccharomyces cerevisiae* are found at different locations in the chromosome. However, the *trp*G.C. and *trp*A.B genes of this yeast are, like *E. coli*, fused to produce polypeptides with several functional domains. Thus in *S. cerevisiae*, tryptophan biosynthesis is catalysed by four separable enzymatic components: steps 2, 3 and 5 are catalysed by the appropriate, separate and independent enzymes but anthranilate synthase (step 1) and indole glycerol phosphate synthase (step 4) are found as a complex of two polypeptides. The information regarding tryptophan biosynthesis in cyanobacteria is limited but is discussed in the next section (Haslam, 1993).

1.5 Aromatic Amino Acid Biosynthesis in Cyanobacteria

Cyanobacteria tend to direct their metabolism towards the synthesis of essential metabolites rather than scavenging compounds from the environment. The metabolism of these organisms have, therefore, been classified as endo-oriented by Jensen and Hall (1982). In such systems the primary site for enzyme regulation is at the first committed step of the biosynthetic pathway. With regard to the shikimate

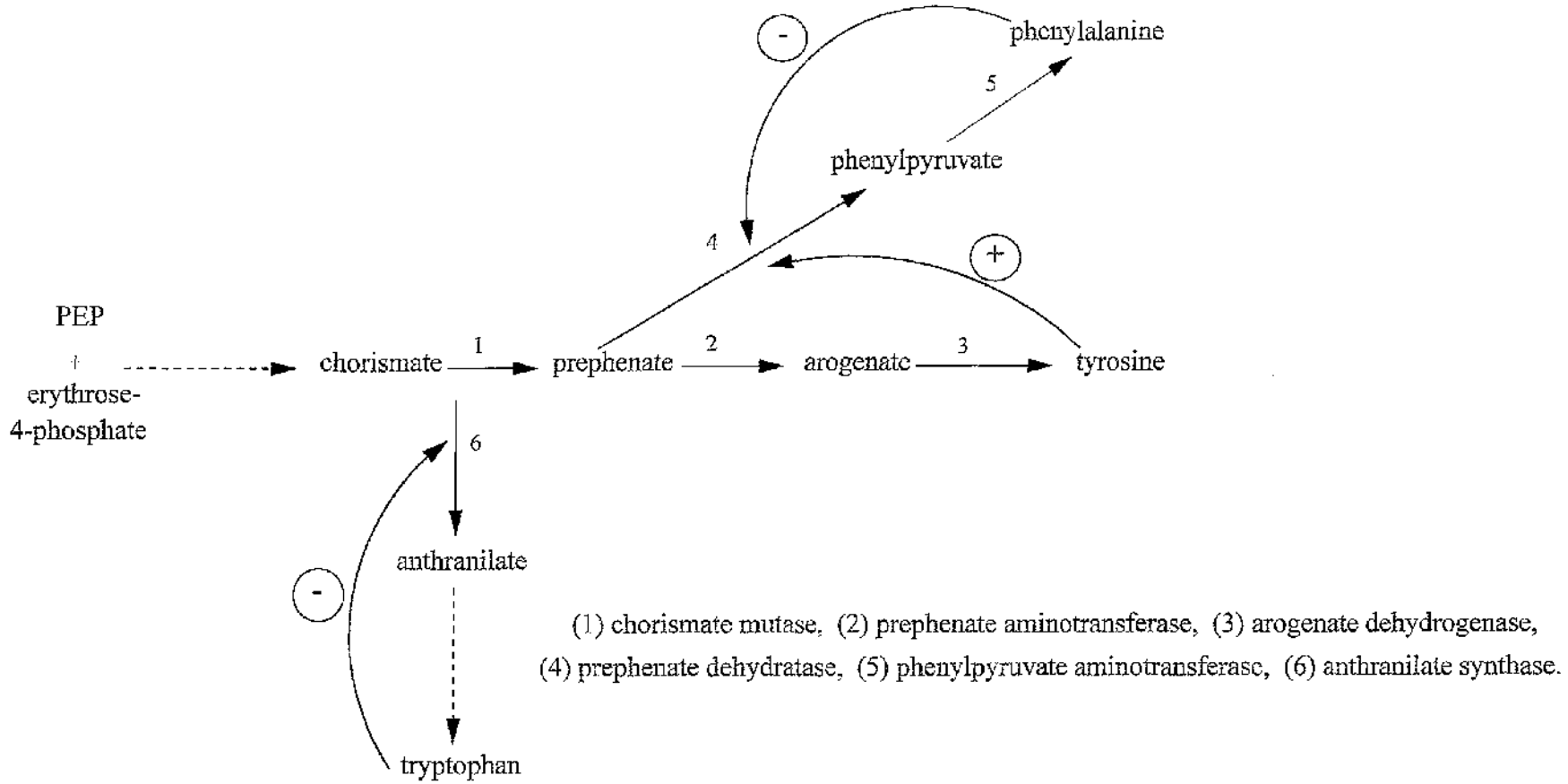


(1) anthranilate synthase (*trp*E and G),
(2) phosphoribosyl anthranilate transferase (*trp*D),
(3) *N*-phosphoribosyl anthranilate isomerase (*trp*F),
(4) indole glycerol phosphate synthase (*trp*C),
(5) tryptophan synthase (*trp*A and B).

Figure 1.5 Pathway to tryptophan from chorismate.

pathway in cyanobacteria, DAHP synthase is the principle site of regulation (Figure 1.1). When examining 48 cyanobacterial strains representative of the five major taxonomic groupings (Rippka *et al.*, 1979), Hall *et al.* (1982) discovered a number of different regulatory patterns controlling DAHP synthase activity. Enzyme activity was controlled by the concentration of phenylalanine in approximately half of the strains studied while unimetabolite inhibition by tyrosine was found in a much smaller number. These results were consistent with the existence of a single form of DAHP synthase. Indeed, this enzyme has been purified from *Anacystis nidulans* (Weber and Bock, 1968; Holtzclaw *et al.*, 1972) and *Synechocystis* sp. strain ATCC 29108 (Hall and Jensen, 1980). In some of the more complex filamentous cyanobacteria, DAHP synthase exhibits a more sophisticated level of control involving cumulative, concerted and additive inhibitions (Hall *et al.*, 1982). The additive inhibition patterns suggest the presence of regulatory isozymes. Indeed, isozymic forms of DAHP synthase have been purified from *Anabaena* sp. strain ATCC 29151: one inhibited by phenylalanine and the other by tyrosine (Hall and Jensen, 1981a). Niven *et al.*, (1988) have also isolated two forms of DAHP synthase regulated by phenylalanine and tyrosine in *Anabaena variabilis* ATCC 29413, but the existence of two separate isozymes was not proven conclusively.

Cyanobacteria synthesise phenylalanine via phenylpyruvate and tyrosine via arogenate as illustrated in Figure 1.4 (Stenmark *et al.*, 1974; Jensen and Stenmark, 1975). Chorismate mutase, which converts chorismate to prephenate, and arogenate dehydrogenase, which forms tyrosine from arogenate, are unregulated. The control of phenylalanine and tyrosine biosynthesis occurs by regulation of the prephenate dehydratase activity, which catalyses the formation of phenylpyruvate from prephenate (Figure 1.6). This enzyme is inhibited by phenylalanine and stimulated by tyrosine (Jensen and Hall, 1982). The activating effects of tyrosine on this enzyme are usually as potent as the inhibitory effects of phenylalanine (Hall *et al.*, 1982). A single transaminase is capable of catalysing prephenate transamination to arogenate and phenylpyruvate transamination to phenylalanine in *A. quadruplicatum* (Stenmark *et al.*,



(1) chorismate mutase, (2) prephenate aminotransferase, (3) arogenate dehydrogenase, (4) prephenate dehydratase, (5) phenylpyruvate aminotransferase, (6) anthranilate synthase.

Figure 1.6 Regulation of aromatic amino acid biosynthesis in cyanobacteria.

→ = one step biochemical process, ----→ = multi-step biochemical process,

(-) = inhibition of enzyme, (+) = activation of enzyme.

1974; Jensen and Stenmark, 1975). The level of phenylalanine and tyrosine influences the behaviour of the enzyme. When tyrosine is in excess, the enzyme functions as a phenylpyruvate transaminase and metabolic flow is directed towards phenylalanine biosynthesis. Alternatively, when phenylalanine is in excess, phenylpyruvate dehydratase is inhibited and so prephenate is guided towards tyrosine biosynthesis (Figure 1.6). This pattern of control allows the appropriate intracellular levels of phenylalanine and tyrosine to be maintained, but it is unclear if transaminase regulation occurs in cyanobacteria other than *A. quadruplicatum.*

Information regarding the enzymes involved in tryptophan biosynthesis in cyanobacteria is limited. However, anthranilate synthase, which converts chorismate to anthranilate (Figure 1.5), has been characterised in *Agmenellum quadruplicatum* (Friedman and Jensen, 1978). This enzyme has been found to be sensitive to feedback regulation by tryptophan (Hall and Jensen, 1981b).

At present the only genes involved in aromatic amino acid biosynthesis that have been cloned encode the enzymes EPSP synthase (Mayes *et al.*, 1993) and chorismate synthase (Schmidt *et al.*, 1993) of the unicellular cyanobacterium, *Synechocystis* PCC 6803 (Section 1.2.1). These cyanobacterial enzymes have yet to be purified and characterised. EPSP synthases have been purified from a large number of different sources because this enzyme is the principle target site for the commercially successful herbicide, glyphosate (Section 1.7.3). Since *Synechocystis* PCC 6803 is naturally tolerant to glyphosate (Powell *et al.*, 1991), characterisation of this enzyme would be of particular interest. The filamentous cyanobacterium, *Anabaena variabilis*, also exhibits natural tolerance to glyphosate (Powell *et al.*, 1991). The EPSP synthase from this organism has been studied in some detail and is discussed in Section 1.9.2.

1.6 The Herbicide Glyphosate

1.6.1 General Properties

Glyphosate (*N*-phosphonomethyl glycine) is a foliar applied, broad-spectrum, non-selective herbicide highly toxic toward the majority of annual and perennial plants (Figure 1.7). Glyphosate is rapidly translocated from the foliar tissue to the metabolically active root and shoot tips to exert its herbicidal effects. Glyphosate is widely used because it has favourable environmental features such as rapid inactivation and degradation by soil microorganisms, low toxicity to animals and minimum soil mobility. The monoisopropylamine salt of glyphosate is the active ingredient in a number of herbicidal formulations, for example: Roundup®, Rodeo®, Tumbleweed® and Vision®. Franz (1985) provides a more detailed review of glyphosate.

1.6.2 The Mode of Action of Glyphosate

Early studies on the mechanism of action of glyphosate indicated that growth inhibition of plants and bacteria by this herbicide was partially alleviated by either phenylalanine or a combination of phenylalanine and tyrosine (Jaworski, 1972). These results suggested that glyphosate was interfering with aromatic amino acid biosynthesis. Subsequently, glyphosate was shown to inhibit the incorporation of [^{14}C]-shikimate into aromatic amino acids and other aromatic compounds, such as lignin, in Buckwheat seedlings (Hollander and Amrhein, 1980). This work was supported by Amrhein and co-workers (1980) who showed that the conversion of [^{14}C]-shikimate to [^{14}C]-chorismate was inhibited by glyphosate in the bacterium, *Klebsiella pneumoniae*. The herbicide also caused significant amounts of shikimate to accumulate (up to 10% of the dry weight) in cultured plant cells (Amrhein *et al.*, 1980). These experiments indicated that the target site for glyphosate inhibition was one of the enzymes catalysing the transformation of shikimate to chorismate. Steinrucken and Amrhein (1980) provided the first evidence that glyphosate specifically inhibits the penultimate enzyme of the shikimate pathway, 5-enolpyruvylshikimate-3-phosphate

(EPSP) synthase, in cell free extracts of *K. pneumoniae*. Similar properties of EPSP synthase inhibition have been found in other bacteria (Duncan *et al.*, 1984a), fungi (Boocock and Coggins, 1983) and plants (Mousdale and Coggins, 1984; Ream *et al.*, 1988; Forlani *et al.*, 1994). As mentioned in Section 1.2., the EPSP synthase reaction of the shikimate pathway occurs exclusively in plants and microorganisms which explains the low toxicity of glyphosate to animals.

Although the principle metabolic target site for glyphosate is EPSP synthase, the herbicide can also inhibit DAHP synthase. In plants, DAHP synthase (which catalyses the first step in the shikimate pathway, Figure 1.1) has a number of isozymic forms each requiring different metal ions for activation. Only the cobalt requiring enzyme (DS-Co) is inhibited by glyphosate. Inhibition of DS-Co by glyphosate in *Nicotiana silvestris* can be relieved by the addition of excess Co^{2+} or by replacing Co^{2+} with magnesium or manganese as the divalent metal ion activator. Enzyme inhibition appears to be due to the formation of cobalt(II) : glyphosate complexes which reduce the amount of Co^{2+} available to the enzyme (Ganson and Jensen, 1988). In addition to the inhibitory effects of glyphosate on DAHP synthase, sub-lethal concentrations of the herbicide increase the level of this enzyme in potato (*Solanum tuberculosum* L.) cells grown in suspension culture. The molecular mechanism for this glyphosate induced increase in DAHP synthase activity is unknown (Pinto *et al.*, 1988).

1.7 5-Enolpyruvylshikimate-3-phosphate (EPSP) Synthase

1.7.1 The Reaction Mechanism of EPSP Synthase

EPSP synthase catalyses the penultimate reaction of the shikimate pathway in which 5-enolpyruvylshikimate-3-phosphate is formed from shikimate-3-phosphate (S3P) and phosphoenolpyruvate (PEP). This reversible reaction involves the catalytic transfer of the enolpyruvyl moiety of PEP to S3P to yield EPSP and inorganic

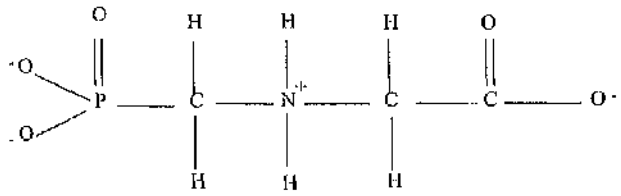


Figure 1.7 The structural formula for glyphosate.

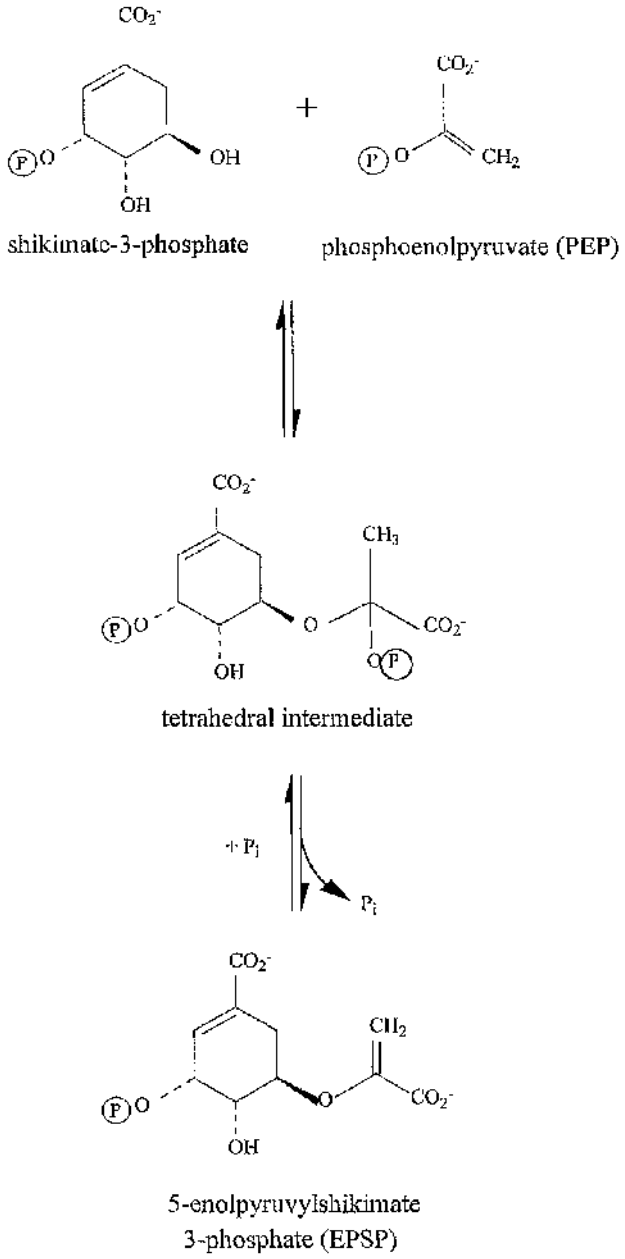


Figure 1.8 The EPSP synthase reaction.

phosphate (Figure 1.8). Hence, the formal name for this enzyme is phosphoenolpyruvate:3-phosphoshikimate 5-*O*(1-carboxyvinyl) transferase.

Levin and Sprinson (1964) originally proposed that during the EPSP synthase reaction a tetrahedral intermediate is formed when the C-5 hydroxyl group of S3P bonds with C-2 of PEP (as illustrated in Figure 1.8). EPSP is produced after inorganic phosphate is eliminated. Consistent with this mechanism, Bondinell *et al.* (1971) showed that an exchange of protons occurred with the solvent at the C-3 position of PEP also that the C-O bond of PEP was cleaved releasing inorganic phosphate during the reaction. The isolation and characterisation of the tetrahedral intermediate tightly bound to the EPSP synthase of *E. coli* by Anderson and co-workers (1988a, 1988b, 1990) provides strong evidence for the initial proposals made by Levin and Sprinson (1964) and Bondinell *et al.* (1971). The definition of a kinetically competent, tetrahedral intermediate in the EPSP synthase reaction has prompted the design of mechanism-based inhibitors (Alberg and Bartlett, 1989 and 1992). Close structural analogues of the tetrahedral intermediate are potent inhibitors of the EPSP synthase reaction.

Glyphosate has been characterised as a reversible competitive inhibitor with respect to PEP and uncompetitive with respect to S3P in most organisms studied so far (Haslam, 1993). The inhibition patterns of EPSP synthase by glyphosate have been used to provide an insight into the kinetic mechanism of the enzyme. The effect of glyphosate on the forward reaction of the *Neurospora crassa* enzyme indicates that substrate and inhibitor binding follows an ordered sequential mechanism. PEP or glyphosate compete with each other to bind to an enzyme:S3P binary complex (Boocock and Coggins, 1983 and Figure 1.9).

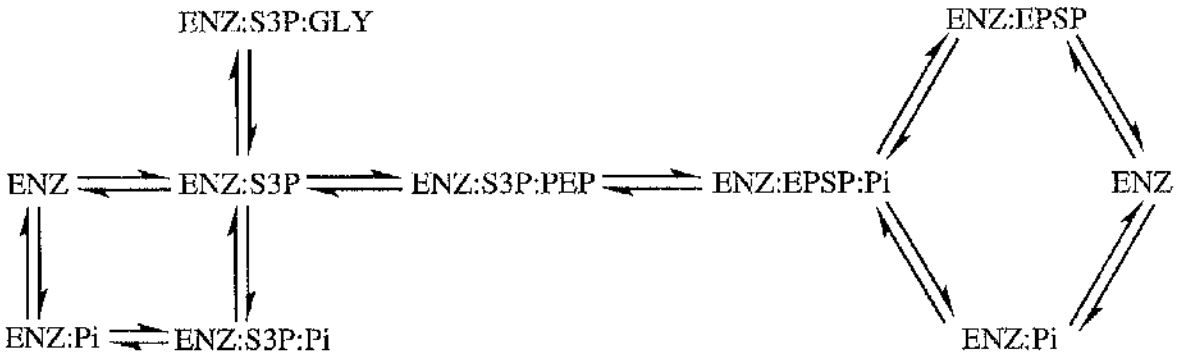


Figure 1.9 The proposed kinetic mechanism of the *Neurospora crassa arom* EPSP synthase (Boocock and Coggins, 1983).

Recently however, detailed kinetic and ligand binding studies of the *E. coli* EPSP synthase forward and reverse reactions have shown that PEP, EPSP and to a lesser extent glyphosate can bind to the free enzyme (Ream *et al.*, 1992; Gruys *et al.*, 1992 and 1993). This suggests a more random character to the enzyme. Nevertheless, this work also showed that both the substrates and glyphosate have greater binding affinities for the binary EPSP synthase:S3P complex than for the free enzyme.

Glyphosate is not a ground state analogue of PEP since it does not bind to the free enzyme as tightly as PEP (as described above) and because it does not inhibit other PEP utilising enzymes (Padgette *et al.*, 1989). Therefore, it has been proposed that glyphosate is a transition state analogue of PEP. Numerous studies have been carried out to characterise the ternary complex formed when glyphosate interacts with the EPSP synthase:S3P complex. The conformation of S3P and glyphosate in the ternary complex is predicted to resemble the tetrahedral intermediate formed when the substrates, S3P and PEP, are bound to the *E. coli* enzyme (Castellino *et al.*, 1989; Christensen and Schaefer, 1993). Sammons *et al.* (1995) hypothesised that if glyphosate was indeed a transition state analogue of PEP, the inhibitor would occupy the same space in the enzyme active site as PEP and, through substrate turnover, EPSP. Therefore, in the reverse EPSP synthase reaction, where EPSP and inorganic

phosphate (P_i) are the substrates, there would be little if any interaction of glyphosate with the enzyme:EPSP or enzyme:EPSP:P_i complexes. However, the results obtained from steady-state kinetic experiments contradicted this hypothesis. These studies revealed that EPSP or a combination of EPSP and inorganic phosphate did not block the interaction of glyphosate with the *E. coli* enzyme. Additionally, various ligand binding experiments showed that glyphosate could form a ternary complex with EPSP synthase in the presence of EPSP. These results do not support the theory that glyphosate acts as a transition state analogue of PEP, but instead suggest that glyphosate binds to amino acid residues near to but outside the enzyme active site.

1.7.2 Amino Acids Involved at the Active Site of EPSP Synthase

To study the mechanism of catalysis and glyphosate inhibition of EPSP synthase, it is useful to identify the amino acids located at the enzyme active site. Identification of active site residues has been largely based on chemical modification experiments and site directed mutagenesis.

Huynh *et al.* (1988a) predicted that lysine and arginine residues would be present at the active site of EPSP synthase due to the anionic nature of the substrates and the inhibitor, glyphosate. As such, lysine residues at positions 340 (Huynh, 1990) and 411 (Huynh, 1991) of the *E. coli* EPSP synthase sequence and arginine residues at positions 28 and 131 of the petunia enzyme (Padgette *et al.*, 1988a) have been shown to be essential for enzyme function. However, the exact roles played by these residues remain unclear.

The importance of lysine-22 in substrate binding has been explored by chemical modification of the *E. coli* EPSP synthase with pyridoxal 5'-phosphate (Huynh *et al.*, 1988a), *o*-phthalaldehyde (Huynh, 1990) and pyruvate in the presence of sodium cyanoborohydride (Huynh, 1992). To verify its function, the lysine residue at position 23 of the petunia enzyme (equivalent to position 22 of the *E. coli* enzyme) was replaced with an arginine, alanine or glutamate residue by site directed

mutagenesis. The arginine mutant retained some activity but the alanine and glutamate variants were inactive (Huynh *et al.*, 1988b). Fluorescence measurements, used to study ligand binding, showed that S3P and glyphosate could interact with the wild type enzyme and arginine mutant but not with the alanine variant. This indicates that the cationic group of lysine-23 plays an important role in substrate binding.

Site directed mutagenesis of the conserved arginine at position 104 of the *Bacillus subtilis* EPSP synthase was performed to analyse the role of this residue in enzyme catalysis (Selvapandiyan *et al.*, 1995). This arginine residue is known to occur in a highly conserved region crucial for binding the phosphate moiety of PEP (discussed below and in Padgette *et al.*, 1991). Mutant enzymes were constructed with lysine and glutamine residues at position 104. Kinetic analyses of both the wild type and variant enzymes indicated that the arginine at this position is involved in both PEP and glyphosate binding.

Chemical modification and photooxidation experiments highlighted the potential importance of the conserved histidine residue at position 385 of the *E. coli* EPSP synthase (Huynh, 1987 and 1993). Since S3P and glyphosate protected the enzyme from inactivation by histidine specific reagents, histidine-385 was thought to constitute part of the enzyme active site. Site directed mutagenesis was performed to investigate the role of this amino acid (Shuttleworth and Evans, 1994). Two mutant enzymes were constructed: one containing a glutamine residue, the other containing an alanine at position 385. Evidence obtained from the wild type and mutant enzymes by kinetic analyses and binding studies (using fluorescence and NMR spectroscopy) revealed that histidine-385 is located at the PEP binding site but is not involved in enzyme catalysis.

Glutamate-418 is thought to be located at or near the EPSP synthase active site since inhibition of the *E. coli* enzyme by carboxyl specific group reagents was prevented by S3P and glyphosate. The negatively charged carboxyl group of this residue may interact with the positively charged imino group of glyphosate but this has yet to be confirmed (Huynh, 1988).

Cysteine-408 is proximal to the active site of the EPSP synthase from *E. coli*, but is not essential for catalytic activity or substrate binding. The observed inactivation of the enzyme by the bulky thiol reagent, DTNB, was merely due to steric hindrance at the active centre (Padgette *et al.*, 1988b).

Padgette and co-workers (1991) were interested in characterising the active site of EPSP synthases by identifying amino acid residues which, when substituted by site directed mutagenesis, influenced ligand binding and catalysis . Comai *et al.* (1983) had previously isolated a glyphosate resistant mutant of *Salmonella typhimurium*. In this strain, tolerance is due to an altered EPSP synthase which has a serine substituted for a proline residue at position 101 in the sequence (Stalker *et al.*, 1985). In 1986, Kishore *et al.* isolated a glyphosate tolerant strain of *E. coli* B. Resistance in this mutant strain is associated with the substitution of a glycine for an alanine at position 96 of the EPSP synthase sequence (Kishore *et al.*, unpublished results, cited by Padgette *et al.*, 1991). Alignment of EPSP synthase sequences allowed Padgette *et al.*, (1991) to identify a highly conserved region from residues 90 to 102 from which a consensus sequence was derived (old consensus in Table 1.1). The importance of this region in various plant EPSP synthases was confirmed by site directed mutagenesis of the corresponding amino acids at positions 96 and 101 and subsequent kinetic analyses of the mutant enzymes. This region was found to be crucial for the interaction of the phosphate moiety of PEP and inorganic phosphate as well as the phosphonate moiety of glyphosate. The consensus sequence defined by Padgette *et al.* (1991) did not consider the corresponding regions in the EPSP synthases of Gram positive bacteria and so was refined by O'Connell *et al.* (1993) (new consensus in Table 1.1).

Table 1.1 Sequence comparison of the highly conserved region of EPSP synthases from various sources.

Source of EPSP synthase	Sequence equivalent to residues 90 to 102 in the *E. coli* EPSP synthase sequence	References
Escherichia coli	L F L G N A G T A M R P L	1
Salmonella typhimurium	L F L G N A G T A M R P L	2
Petunia	L F L G N A G T A M R P L	3
Tomato	L F L G N A G T A M R P L	3
Aspergillus nidulans [a]	L Y L G N A G T A S R F L	4
Saccharomyces cerevisiae [a]	L Y L G N A G T A S R F L	5
Old consensus [b]	L X L G N A G T A X R X L	6
Mycobacterium tuberculosis	- - - G L A G T V L R F V	7
Bacillus subtilis	L D V G N S G T T I R L M	8
Staphylococcus aureus	L Y T G N S G T T T R L L	9
New consensus [b]	X X X G X X G T X X R X X	9

[a] Part of the *arom* multifunctional complex. [b] The symbol X of the consensus sequence represents a non-conserved amino acid. References (1) Duncan *et al.*, 1984b; (2) Stalker *et al.*, 1985; (3) Gasser *et al.*, 1988; (4) Charles *et al.*, 1986; (5) Duncan *et al.*, 1987; (6) Padgette *et al.*, 1991; (7) Garbe *et al.*, 1990; (8) Henner *et al.*, 1986; (9) O'Connell *et al.*, 1993.

For a more detailed account of the reaction mechanism of EPSP synthase and its inhibition by glyphosate, the results obtained from the various chemical and molecular techniques described above need to be correlated to the 3-dimensional structure of the enzyme. Stallings *et al.* (1991) have crystallised the native EPSP synthase from *E. coli* and determined its structure at 3×10^{-10} m resolution, although only the back-bone structure of the enzyme has been published. The *E. coli* EPSP synthase consists of two globular domains with the active site apparently located at the inter-domain region. Previous results obtained from fluorescence measurements have shown that EPSP synthase undergoes significant conformational changes upon ligand

binding (Anderson and Johnson, 1990). So to accurately determine which residues are located within the active site, the structure of EPSP synthase complexed with S3P and PEP or S3P and glyphosate is required, but has not yet been elucidated. In spite of this and although the X-ray crystal structure of EPSP synthase is not available in the public domain, it is thought that histidine-385 (Shuttleworth and Evans, 1994), cysteine-408 and lysine-411 (Huynh, 1991) are in the vicinity of the enzyme active site.

1.7.3 Purification of EPSP Synthases from a Diverse Range of Sources

Prior to 1983, EPSP synthase had only been purified to homogeneity as part of the *arom* multifunctional complex of *Neurospora crassa* (Lumsden and Coggins 1977) and *Euglena gracilis* (Patel and Giles, 1979). The interest in EPSP synthase has greatly increased since the discovery that it is the target site for glyphosate inhibition. As a result, the enzyme has been purified to near homogeneity from an array of different sources, including: *E. coli* (Lewendon and Coggins, 1983) and *Klebsiella pneumoniae* (Steinrucken and Amrhein, 1984a), and partially purified from several other bacterial species (Schulz *et al.*, 1985). Higher plant EPSP synthases have also been isolated from *Pisum sativum* (Mousdale and Coggins, 1984), *Petunia hybrida* (Steinrucken *et al.*, 1986), *Sorgum bicolor* (Ream *et al.*, 1988) and *Zea mays* (Forlani *et al.*, 1994). The EPSP synthase of higher plants is predominantly located in the chloroplasts and root plastids of plant cells as demonstrated by Mousdale and Coggins, (1985), Smart and Amrhein (1987), della Cioppa *et al.* (1986) and Schmid and Amrhein (1995). Mousdale and Coggins (1986) utilised the compartmentation of this enzyme to develop a rapid, small scale purification protocol for EPSP synthases from a range of higher plant species.

EPSP synthases from bacteria and plants are generally monomeric and monofunctional proteins of similar molecular mass (43 - 52 kDa), while the EPSP synthase activity of fungi forms part of the *arom* pentafunctional complex. Irrespective of the source of the enzyme, EPSP synthases have broadly similar kinetic

characteristics (see Table 1.2 below and Table 1.3 in Section 1.9.2). The EPSP synthase from the filamentous cyanobacterium, *Anabaena variabilis*, has been purified. The physical and kinetic properties of this enzyme are discussed in Section 1.9.2.

Table 1.2 Kinetic parameters of three purified EPSP synthase enzymes

	Source of EPSP synthase		
Kinetic Parameter	*Escherichia coli*	*Neurospora crassa*	*Pisum sativum*
K_m for PEP	16 μM	2.7 μM	5.2 μM
K_m for S3P	2.5 μM	0.36 μM	7.7 μM
K_m for EPSP	3.0 μM	0.25 μM	5.2 μM
K_m for phosphate	2.5 μM	1.8 μM	4.0 μM
K_i for glyphosate	0.9 μM	1.1 μM	0.08 μM

Reproduced from Lewendon and Coggins (1987).

Isozymic forms of EPSP synthase have been isolated from a range of higher plant species including: *Pisum sativum* (Mousdale and Coggins, 1985), *Nicotiana silvestris* (Rubin *et al.*, 1984), *Sorgum bicolor* (Ream *et al.*, 1988) and *Zea mays* (Forlani *et al.*, 1994). The exact role of these isozymes has yet to be elucidated.

The *Bacillus subtilis* EPSP synthase has recently been overexpressed and purified to near homogeneity (Majumder *et al.*, 1995). The kinetic characteristics of this enzyme do not conform to EPSP synthases isolated from other sources. The experimental data show that the *B. subtilis* EPSP synthase exhibits allosteric behaviour and may harbour an additional non-catalytic PEP binding site.

1.7.4 Isolation of the EPSP Synthase Gene (*aro*A) from a Diverse Range of Sources

The EPSP synthase or *aro*A gene has been cloned from a variety of sources. The *aro*A gene of *E. coli* has been cloned (Duncan and Coggins, 1986) and is transcribed with the *ser*C gene which encodes an enzyme involved in serine

biosynthesis. This mixed function operon is conserved in other Gram negative bacteria such as *S. typhimurium* (Hoiseth and Stocker, 1985) and *K. pneumoniae* (Sost and Amrhein, 1990). The *aro*A gene of *Bacillus subtilis* (termed *aro*E) forms part of a supra-operon with other genes involved with aromatic amino acid biosynthesis (Henner *et al.*, 1986). In *Synechocystis* sp. PCC 6803, the *aro*A gene does not appear to constitute part of an operon, however, analysis of the DNA sequence 5' to the open reading frame has not revealed any convincing promoter sequences (dalla Chiesa *et al.*, 1994). The *aro*A gene has been cloned from a large variety of medically important bacteria since Hoiseth and Stocker (1981) found that inactivation of the *S. typhimurium* gene caused a loss of virulence in this organism. Mutant *aro*A strains of *S. typhimurium* are, therefore, very effective vaccines. As such the EPSP synthase gene has been cloned from *Bordetella pertussis* (Maskell *et al.*, 1988), *Yersinia enterocolitica* (O'Gaora *et al.*, 1989), *Aeromonas salmonicida* (Vaughan *et al.*, 1993), *Dichelobacter nodosus* (Alm *et al.*, 1994) and *Pasteurella haemolytica* (Tatum *et al.*, 1994).

In *Aspergillus nidulans, Saccharomyces cerevisiae* and *Pneumocystis carinii* the nucleotide sequence for the EPSP synthase activity is part of a single gene encoding the *arom* pentafunctional complex (Charles *et al.*, 1986; Duncan *et al.*, 1987; Banerji *et al.*, 1993).

In higher plants, the *aro*A gene is located in the nuclear DNA and is interrupted by introns: seven in *Arabidopsis* (Klee *et al.*, 1987) and petunia (Gasser *et al.*, 1988) and eight in *Brassica napus* (Gasser and Klee, 1990). cDNA sequences for EPSP synthases have been cloned from a number of plant species, for example: petunia (Shah *et al.*, 1986) tomato (Gasser *et al.*, 1988) and pea (Granger, 1989). Nucleotide sequence analysis of the petunia EPSP synthase cDNA suggests that the enzyme is synthesised as a precursor polypeptide (pre-EPSP synthase) containing a transit peptide sequence which targets the enzyme to the chloroplast (Shah *et al.*, 1986). *In vitro* uptake studies have confirmed the pre-EPSP synthase is rapidly translocated into the chloroplast stroma where it undergoes proteolytic maturation. Interestingly, the precursor enzyme has catalytic activity and is as sensitive to inhibition by glyphosate as

the mature enzyme (della Cioppa *et al.*, 1986). Import of the pre-EPSP synthase into chloroplasts is also significantly reduced in the presence of glyphosate (della Cioppa and Kishore, 1988). This adds a new dimension to glyphosate inhibition of EPSP synthase if these *in vitro* studies reflect the *in vivo* situation in plant cells.

The amino acid sequences of EPSP synthases from various prokaryotes and eukaryotes have been analysed and compared. Evolutionary trees depicting relationships between EPSP synthase sequences isolated from different sources have been constructed (Griffin and Griffin, 1991) and have divided the enzymes into four separate groups: Gram negative bacteria, Gram positive bacteria, fungi and plants. This work revealed that Gram negative bacterial and higher plant enzyme sequences are more closely related to each other than to the fungal sequences. Thus, the trees presented in this study show that divergence between the Gram negative bacterial and plant EPSP synthases occurred some time later than the divergence of the fungal synthase from an ancestral enzyme. This same conclusion was reached by Gasser *et al.* (1988). The findings of Griffin and Griffin (1991) and Gasser *et al.* (1988) support the hypothesis that the plastids of higher plants are descendants of endosymbiotic prokaryotes that invaded ancestral plant cells (Margulis, 1970). If this is indeed the case, then it is possible that the plant *aro*A sequence may represent a gene that migrated from the plastids to the nucleus (Gasser *et al.*, 1988). As such, these genes would be expected to retain homology to the equivalent sequences in prokaryotes (Weedon, 1981). Since a cyanobacterium-like ancestor is thought to have given rise to plastids from higher plants and algae (Giovannoni *et al.*, 1988; Douglas and Turner, 1991), it is interesting to note that the EPSP synthase sequence of the cyanobacterium, *Synechocystis* sp. PCC 6803, shares more homology with the Gram positive bacterium, *Bacillus subtilis*, than with the *Euglena gracilis* or the available higher plant sequences (Mayes *et al.*, 1993). Giovannoni *et al.* (1988) and Nelissen *et al.* (1995) have provided evidence to suggest that the cyanobacteria had diversified into different species prior to endosymbiosis with precursor eukaryotic cells. Thus, *Synechocystis* may not be the

present-day representative of the ancestral cyanobacterial sub-line that became the progenitor of the plastid organelle.

The cloning of *aro*A genes from different organisms has allowed EPSP synthases to be overexpressed for detailed structural and mechanistic work (previously described in Sections 1.7.1/2). Genetic engineering techniques have also allowed the *aro*A gene from different sources to be introduced into plant cells for the construction of transgenic plants resistant to glyphosate. This is discussed in more detail in the next section.

1.8 Genetic Engineering of Glyphosate Tolerance

As stated in Section 1.6, glyphosate is a non-selective herbicide and so cannot distinguish between crop plants and weeds. This can cause crop injury and results in a significant loss in crop yield (Padgette *et al.*, 1989). Consequently, the use of this herbicide has been restricted. Since not one crop plant studied to date is naturally resistant to glyphosate, it is highly desirable to engineer selective tolerance to this herbicide into plants. Transgenic plants resistant to glyphosate would permit the application of this environmentally safe herbicide for weed management during active growth of the crop.

The development of glyphosate resistant transgenic plants has focused on two approaches, both of which involve the target site of the inhibitor, EPSP synthase. The first approach requires the overproduction of wild type EPSP synthase where resistance is conferred by the residual EPSP synthase activity in the presence of glyphosate. The second approach is based on the expression of a mutated, glyphosate tolerant EPSP synthase. Tolerance is achieved because the enzyme is resistant to inhibition by glyphosate.

The first evidence that glyphosate tolerance could be accomplished by overproduction of EPSP synthase was reported by Rogers *et al.* (1983). In this work, *E. coli* cells harbouring the *E. coli* gene on a multicopy number plasmid overproduced the enzyme approximately 17-fold. This caused an 8-fold increase in tolerance to

glyphosate. Subsequently, plant cell cultures were selected that exhibited tolerance to glyphosate. Resistance due to overproduction of EPSP synthase occurred in *Corydalis sempervirens* (Amrhein *et al.*, 1983), carrot (Nafziger *et al.*, 1984), petunia (Steinrucken *et al.*, 1986) and tobacco cells (Dyer *et al.*, 1988). The molecular basis for the increased EPSP synthase activity was due to gene amplification in petunia (Shah *et al.*, 1986), carrot (Hauptmann *et al.*, 1988; Shyr *et al.*, 1993) and tobacco cells (Goldsbrough *et al.*, 1990; Wang *et al.*, 1991). In *Corydalis sempervirens*, glyphosate tolerance was acquired by an increased rate of transcription and a reduced turnover of the enzyme (Hollander-Czytko *et al.*, 1992).

Overexpression of EPSP synthase has also been achieved in petunia and *Arabidopsis* cells by cloning their respective *aro*A genes behind the strong cauliflower mosaic virus 35s (CaMV35s) promoter (Shah *et al.*, 1986; Klee *et al.*, 1987). This promoter causes high level, constitutive gene expression in most plant cells and tissues. The resulting transgenic plants were more resistant to glyphosate but not at the levels required for commercial use.

The main site of glyphosate inhibition in plants is in the leaves where the herbicide is applied. Phloem transport of glyphosate to the root and apical meristems also blocks EPSP synthase function at these locations (Arnaud *et al.*, 1994). To confer complete glyphosate tolerance at the whole plant level, significant overproduction of EPSP synthase in these target tissues is required. To circumvent this problem, mutant EPSP synthases with a reduced affinity for glyphosate have been isolated. These enzymes decrease the level of expression required in plant cells to achieve tolerance.

Mutant forms of *S. typhimurium* (Comai *et al.*, 1983) and *E. coli* (Kishore *et al.*, 1986) that are insensitive to glyphosate have been isolated. Both mutant strains contain an altered EPSP synthase with reduced sensitivity to glyphosate encoded by a mutated *aro*A gene. In the variant *S. typhimurium* EPSP synthase, a proline to serine substitution at position 101 of the enzyme sequence confers resistance (Stalker *et al.*, 1985), whereas in *E. coli* (Kishore, unpublished results in Padgette *et al.*, 1991) the mutant enzyme has a glycine to alanine substitution at position 96. Both mutations

occur in a highly conserved region thought to form part of the enzyme active site (Section 1.7.2 and Table 1.1).

The mutant *S. typhimurium aro*A gene has been expressed in the cytosol of tobacco (Comai *et al.*, 1985), tomato (Fillatti *et al.*, 1987) and populus cells (Fillatti *et al.*, 1988). In each case, the resulting transgenic plants showed an increased resistance to glyphosate, but the levels of tolerance were not commercially useful. As discussed in Section 1.2.3, EPSP synthase and other enzymes of the shikimate pathway in plants are predominantly located in the chloroplasts and root plastids. The *S. typhimurium aro*A gene lacks the transit peptide sequence crucial for targeting the synthesised enzyme to the plastids. *In vitro* studies using the mutated *E. coli aro*A gene fused to the chloroplast transit peptide (CTP) of the petunia EPSP synthase have shown that the resulting fusion protein is efficiently imported into isolated chloroplasts (della Cioppa *et al.*, 1987). As such, transgenic tobacco plants more resistant to glyphosate have been constructed by expressing the resistant *E. coli* EPSP synthase in the chloroplasts (Kishore *et al.*, 1992). These experiments highlight the importance of targeting EPSP synthase to the plastids and confirmed that a significant proportion of aromatic amino acid biosynthesis must occur in these plant organelles.

Mutations corresponding to the glycine to alanine substitution at position 96 in the *E. coli* EPSP synthase have been made in various plant enzymes (Padgette *et al.*, 1991). These altered enzymes have a decreased binding affinity (or increased K_i) for glyphosate but are unsuitable for making commercially viable tolerant plants. The problem being that the mutational change also results in plant enzymes with a reduced affinity for their substrates, especially PEP. This compromises the ability of the enzyme to perform its catalytic function under physiological conditions. Therefore, new glyphosate tolerant plant variant enzymes have been isolated that are more catalytically active due to lower K_m values (or higher binding affinities) for PEP. Work is ongoing to establish if these new altered enzymes are suitable for the construction of glyphosate tolerant plants (Kishore *et al.*, 1992).

Very recently, Padgette *et al.* (1995) constructed a transgenic plant line by the expression of a glyphosate tolerant EPSP synthase from the wild type *Agrobacterium* strain CP4 in soybean. The bacterial enzyme was expressed from the CaMV35s promoter and targeted to the chloroplasts using the petunia CTP. The yield obtained from this transgenic line did not vary significantly from the parent line (Delannay *et al.*, 1995). Therefore, glyphosate tolerant soybean should soon be available to farmers as an alternative to traditional crops.

The possible use of herbicide resistant crops has attracted many criticisms as it may lead to the creation of super weeds by the transfer of resistant genes from crops to related weed species. However, the results obtained from crosses between a variety of crops and their weed relatives do not support this claim (Kishore *et al.*, 1992). Plant genetic engineering has also courted controversy for fostering the use of chemicals in agriculture. Nevertheless, glyphosate is an environmentally safe herbicide and provides a cost-effective means of weed management. This technology may even promote the use of more desirable herbicides like glyphosate.

1.9 Cyanobacteria and Herbicides

1.9.1 General Aspects

Most herbicides target plant functions performed in the chloroplast. Cyanobacteria have been used as models to determine the toxic effects of herbicides at the physiological and biochemical level since these organisms are considered to share a common ancestry with plant chloroplasts (Giovannoni *et al.*, 1988).

A variety of mutant cyanobacterial strains resistant to a selection of herbicides have been isolated. Some of these strains contain altered genes whose products confer resistance to these chemicals. For example, the D1 protein is an essential sub-unit for photosystem II and is a target for several herbicides. A number of herbicide resistant mutations map within the gene encoding this protein in *Synechocystis* sp. PCC 6714 (Ajlani *et al.*, 1989; Boujoub *et al.*, 1993). Additionally, phytoene desaturase (*pds*) is a

key enzyme in the carotenoid biosynthetic pathway and is a target for several bleaching herbicides. Mutant strains of *Synechococcus* resistant to these herbicides contain specific point mutations in the *pds* gene (Chamovitz *et al.*, 1991 and 1993). These studies have increased the understanding of the role of D1 in photosystem II and phytoene desaturase in carotenoid biosynthesis.

Much research has been carried out to elucidate the toxic effects of herbicides on cyanobacteria and other aquatic and soil microorganisms to assess the potential hazards of these chemicals to non-plant organisms during weed control. A study performed by Peterson *et al.* (1994) illustrated that a range of cyanobacteria from the five taxonomic groupings (Rippka *et al.*, 1979) are naturally tolerant to the expected environmental concentrations of a variety of herbicides. Cyanobacteria have also been shown to exhibit variable levels of tolerance to phosphinothricin, a glutamine synthetase inhibitor (Lea *et al.*, 1984). Additionally, a wild type strain of *Synechocystis* is naturally tolerant to a sulphonylurea herbicide which inhibits acetolactate synthase, an enzyme involved in branched chain amino acid biosynthesis (Friedberg and Seijffers, 1988).

Both mutant and wild type cyanobacteria could, therefore, act as a source of genes encoding herbicide resistant proteins for the development of transgenic plants as discussed in Section 1.8.

1.9.2 Glyphosate Tolerance in Cyanobacteria

The physiological and biochemical effects of glyphosate on the filamentous, nitrogen fixing cyanobacterium, *Anabaena variabilis* ATCC 29413, and the unicellular, non-nitrogen fixing cyanobacterium *Synechocystis* sp. PCC 6803 have been examined (Powell, 1991; Powell *et al.*, 1991; Powell *et al.*, 1992). Both cyanobacteria showed a high degree of tolerance to glyphosate when applied as the free acid, the monoisopropylamine salt and the commercial formulation (Roundup). However, each chemical form of the herbicide exhibited a different level of toxicity: Roundup was the most toxic and the free acid, the least.

There are several mechanisms by which glyphosate resistance may arise in microorganisms and plants. These are: reduced uptake or degradation of the herbicide; overproduction of EPSP synthase or expression of a glyphosate resistant EPSP synthase. Each of these possibilities were investigated to determine the mechanism(s) for glyphosate resistance in *A. variabilis* and *Synechocystis* and the reasons for the differential toxic effects of the herbicide formulations.

HPLC analysis was performed to measure the intracellular levels of glyphosate in *A. variabilis* and *Synechocystis* cells. Each chemical form of glyphosate had a different rate of uptake which accounted for the varied levels of toxicity. It was concluded that cyanobacterial tolerance to glyphosate was not due to inefficient uptake as the internal concentrations of the herbicide would have been high enough to inhibit EPSP synthases from other wild type organisms. For example, when *A. variabilis* cells were grown in the presence of 20 mM of the free acid of glyphosate, the intracellular concentration of the herbicide was approximately 9 mM. This concentration of glyphosate would inhibit the EPSP synthase from *E. coli* and other sources (K_i values, Table 1.3). Further HPLC analysis also revealed that the cyanobacteria could not degrade glyphosate, so tolerance was not caused by detoxification of the herbicide. (Powell *et al.*, 1991).

Other investigations indicated that both the monoisopropylamine salt and Roundup inhibited the shikimate pathway in the cyanobacteria. This was because the intracellular concentration of the central metabolite, shikimate, increased in the presence of these chemicals. Also, the addition of exogenous aromatic amino acids partially alleviated their toxic effects. Therefore, the target enzyme for glyphosate, EPSP synthase, was purified to near homogeneity from *A. variabilis* cells and characterised (Powell *et al.*, 1992). It was established that glyphosate resistance in *A. variabilis* was not due to the overproduction of EPSP synthase. Indeed, the enzyme was found to have broadly similar kinetic characteristics as EPSP synthases from other sources (Table 1.3). Notably however, the cyanobacterial enzyme had a raised K_i for glyphosate (approximately ten times higher than EPSP synthases from other sources).

Table 1.3 Comparison of some kinetic and physical properties of wild type and mutant EPSP synthases isolated from various sources.

Organism	K_m PEP (μM)	K_m S3P (μM)	K_m EPSP (μM)	K_i PMG [a] (μM)	MW [b]	Ref. [f]
Escherichia coli	16	2.5	3.0	0.90	49	1
Klebsiella pneumoniae	17	10 [c]	2.0 [c]	1.00	43	2
Neurospora crassa	3	0.4	0.3	1.10	ND [d]	3
Corydalis sempervirens	86	21	ND	0.32	46	4
Pisum sativum	5	7.7	5.2	0.08	50	5
Sorghum bicolor	8	7.0	ND	0.16	51	6
Petunia hybrida	14	8.5	ND	0.17	52	7
Fagopyrum esculentum	22	ND	ND	1.75	49	8
Pseudomonas sp.	220	ND	ND	200	ND	9
E. coli G96A [e]	220	ND	ND	4100	46	10
Petunia G96A [e]	200	ND	ND	2000	52	10
Anabaena variabilis	34	6.0	1.0	350	49	11

[a] PMG = glyphosate. [b] MW = molecular weight (kDa). [c] values calculated from data in reference. [d] ND = not determined. [e] G96A = mutant EPSP synthases with an alanine residue substituted for a glycine residue at position 96 in the enzyme sequence. [f] Ref = References. (1) Lewendon and Coggins, 1987; (2) Steinrucken and Amrhein, 1984b; (3) Boocock and Coggins, 1983; (4) Smart *et al.*, 1985; (5) Mousdale and Coggins, 1984; (6) Ream *et al.*, 1988; (7) Steinrucken *et al.*, 1986; (8) Amrhein *et al.*, 1987; (9) Kishore and Shah, 1988; (10) Kishore *et al.*, 1992; (11) Powell *et al.*, 1992.

The *A. variabilis* EPSP synthase is, therefore, tolerant to high concentrations of glyphosate. The lower affinity of this enzyme for the herbicide may be due to the slightly elevated K_m value (reduced binding efficiency) for PEP as illustrated in Table 1.3. The specific activity of the enzyme has not yet been accurately determined because of the low yield of EPSP synthase protein extracted from *A. variabilis* cells.

Several EPSP synthases resistant to inhibition by glyphosate have been isolated from cells selected following growth in the presence of the herbicide (Comai *et al.*, 1983; Kishore *et al.*, 1992; Section 1.8). But the glyphosate tolerant EPSP synthase from *A. variabilis* had been purified from wild type cells previously unexposed to the herbicide. Several *Pseudomonas* species (Shultz *et al.*, 1985) and more recently an *Agrobacterium tumefaciens* strain (Padgette *et al.*, 1995) have been isolated that are also naturally tolerant to glyphosate due to resistant EPSP synthases. As well as having a reduced affinity for glyphosate, the pseudomonad EPSP synthases plus the mutated bacterial and plant enzymes do not bind PEP efficiently. This can interfere with the ability of the enzyme to perform its function *in vivo*. The wild type *Agrobacterium* EPSP synthase, however, is more catalytically active since it has a high affinity for PEP (Section 1.8). The K_m for PEP of the wild type *A. variabilis* EPSP synthase is only slightly higher than those values obtained from glyphosate sensitive EPSP synthases, so the cyanobacterial enzyme can bind this substrate efficiently (Table 1.3). For this reason and because cyanobacteria are thought to share a common ancestry with higher plant chloroplasts (Giovannoni *et al.*, 1988; Douglas and Turner, 1991), the *aro*A gene encoding the *A. variabilis* EPSP synthase may be suitable for the construction of transgenic plants resistant to glyphosate.

1.10 Aim of this Project

The aim of this project was to isolate the *A. variabilis* EPSP synthase gene. Cloning the *aro*A gene into an appropriate plasmid vector would allow the overexpression of the EPSP synthase in a suitable host and subsequent purification of large quantities of the enzyme. This would permit a more detailed study of the protein

which was previously hindered by the very low yield of EPSP synthase obtained from *A. variabilis* cells (only 60 ng of protein was isolated from 25 g wet weight of cells (Powell, 1991). If greater quantities of the enzyme were available, the kinetic data collated by Powell *et al.* (1992) could be verified, the specific activity calculated and the structure of the enzyme active site could be investigated. This would determine how PEP and glyphosate interact with the *A. variabilis* EPSP synthase.

Chapter 2

Materials and Methods

2.1 Materials

2.1.1 Chemicals and Biochemicals

Ampicillin, tetracycline, bromophenol blue, rubidium chloride, MOPS, Ficoll 400, polyvinylpyrrolidone and ethidium bromide were obtained from the Sigma Chemical Company, Dorset, England.

Bacto-tryptone, bacto-yeast extract, and bacto-agar were obtained from Difco Laboratories, Detroit, U.S.A.

Bacteriological agar no. 1 was obtained from Oxoid, Unipath Limited, Hampshire, England.

Phenylalanine, tyrosine and tryptophan as well as *para*-aminobenzoic acid and *para*-hydroxybenzoic acid were obtained from BDH Laboratory Supplies, Dorset, England.

Agarose, low melting point agarose, and X-gal were obtained from GIBCO BRL Life Technologies, Paisley, Scotland.

IPTG was obtained from Nova Biochem (UK) Limited, Nottingham, England.

Radioactive [α-^{32}P]dCTP and [α-^{35}S]dATP were obtained from Amersham International, Slough, England.

2.1.2 Proteins and Enzymes

Lysozyme and bovine serum albumin were obtained from the Sigma Chemical Company, Dorset, England.

Restriction enzymes and their buffers were obtained from GIBCO BRL Life Technologies, Paisley, Scotland; New England Biolabs. Incorporated, Hertfordshire, England and Promega Corporation, Southampton, England.

Bacteriophage T4 DNA ligase, calf intestinal alkaline phosphatase, *Taq* polymerase, Klenow fragment of *E. coli* DNA polymerase I and appropriate buffers were obtained from Promega Corporation, Southampton, England.

Sequenase Version 2.0 plus related reagents were obtained from United States Biochemical, Amersham International plc., Cleveland, Ohio, U.S.A.

Klenow fragment of *E. coli* DNA polymerase I and related buffers for the megaprime labelling kit were obtained from Amersham International plc., Berkshire, England.

2.1.3 Chromatography Matrices

Sephadex G50 was obtained from Pharmacia Biotechnologies, Milton Keynes, England.

2.1.4 Bacterial Strains and Cloning Vectors

Bacterial strains, phagemid and phage vectors used for molecular cloning experiments are listed in Tables 2.1 and 2.2.

2.2 Growth Media

2.2.1 Rich Media

The media listed in Table 2.3 were used to support the growth of most *E. coli* strains in this project. Sterilised media by autoclaving for 20 minutes at 15 lb/sq. inch.

2.2.2 Minimal medium

Minimal medium was used for the growth of *E. coli* TG1 and AB2829. Only top grade reagents were used to make this medium to avoid contamination with amino acids.

To make solid minimal medium, added 200 mls 2 x (2 times concentrated) M9 salts to 200 mls bacto-agar (final concentration 1.25 % agar). When the medium cooled to 50 - 55 ^{o}C, added 4.8 mls concentrated mix (glucose, $CaCl_2$, and vitamin B_1 to a final concentration of 2 g/litre, 0.1 mM and 2 μg/ml respectively) and mixed

Table 2.1 Bacterial strains used for molecular cloning experiments.

Strain	Genotype	Origin / Reference
E. coli TG1	*sup*E *hsd*Δ5 *thi*Δ(*lac-pro*AB) F'[*tra*D36 *pro*AB$^+$ *lac* I^q *lac*ZΔM15]	Gibson, 1984
Epicurian Coli ® SURE™	e14(*mcr*A)Δ(*mcr*CB-*hsd*SMR-*mrr*)171 *end*A1 *sup*E44 *thi*-1 *gyr*A96 *rel*A1 *lac* *rec*B *rec*J *sbc*C *umu*CTn5(kan^r) *uvr*C F'[*pro*AB *lac*I^{q}ΔM15Tn10(tet)]	Greener, 1990
E. coli AB2829	*aro*A$^-$	Pittard and Wallace, 1966
E. coli DH5α	*sup*E44 Δ*lac*U169 (ø80 *lac*ZΔM15) *hsd*R17 *rec*A1 *end*A1 *gyr*A96 *thi*-1 *rel*A1	Hanahan, 1985
Epicurian Coli ® XL2-Blue MRF'	Δ(*mrc*A)183 Δ(*mrc*CB-*hsd*SMR-*mrr*)173 *end*A1 *sup*E44 *thi*-1 *rec*A1 *gyr*A96 *rel*A1 *lac*[F'*pro*Ab *lac*I^{q}ZΔM15Tn10 (Tet^r Amy Cam^r)]	Greener and Jerpseth, 1993
Epicurian Coli ® XL1-Blue MRA (P2)	Δ(*mcr*A)183 Δ(*mcr*CB-*hsd*SMR-*mrr*)173 *end*A1 *sup*E44 *thi*-1 *gyr*A96 *rel*A1 *lac* (P2 lysogen)	Elgin *et al.*, 1991

Table 2.2 Phagemid and phage vectors used for molecular cloning experiments.

Vector	Host	Origin / Reference
M13mp19RF	*E. coli* TG1 *Epicurian coli* SURE™	Messing, 1983
pBluescript SK-	*E. coli* DH5α, *E. coli* AB2829, *Epicurian coli* XL2-Blue MRF'	Short *et al.*, 1988
lambda FIX® II	*E. coli* XL1-Blue MRA (P2)	Elgin *et al.*, 1991

Table 2.3 Rich Growth Media.

Media	Composition per Litre
LB (Luria-Bertani) Broth	10 g bactotryptone 5 g yeast extract 10 g NaCl
LB agar	LB broth plus 15 g bactoagar
2 x TY	16 g bactotryptone 10 g bactoyeast extract 5 g NaCl
H broth	10 g bactotryptone 8 g NaCl
H agar	H broth plus 12 g bactoagar
H top agar	H broth plus 8 g bactoagar
Ψ broth	20 g bactotryptone 5 g bactoyeast extract 4 g $MgSO_4$. $7H_2O$ 0.74 g KCl
Ψ agar	Ψ broth plus 15 g bactoagar
NZY	5 g NaCl 2 g $MgSO_4.7H_2O$ 5 g bactoyeast extract 10 g NZ amine (casein hydrolysate) Adjusted pH 7.5 with NaOH

thoroughly. Two x M9 salt solution and bacto-agar were autoclaved separately to avoid precipitation problems. The components of minimal medium were as follows:

(a) 2 x M9 Salts

For 1 litre: 2 g NH_4Cl, 0.26 g $MgSO_4.7H_2O$, 6 g KH_2PO_4, 12 g Na_2HPO_4. Autoclaved to sterilise.

(b) Minimal Medium Agar

Added 5 g of Oxoid no. 1 agar to 200 mls distilled water (2.5% agar) and autoclaved to sterilise.

(c) Concentrated mix

For 75 mls: 15 g D-glucose, 1.5 mls 0.5 M $CaCl_2$. Autoclaved to sterilise. Once cooled added 15 mls of filter sterilised 1 mg/ml vitamin B_1 (thiamine hydrochloride). Stored in dark at 4 °C.

(d) Amino Acid Supplements For Minimal Medium (2 x concentrated)

To support the growth of the auxotrophic *aro*A mutant *E. coli* AB2829 (Table 2.1), the following amino acids were added to 2 x M9 salts prior to autoclaving. For 1 litre: 160 mg phenylalanine, 160 mg tyrosine, 80 mg tryptophan, 0.64 mg *p*-aminobenzoic acid, 0.64 mg *p*-hydroxybenzoic acid.

2.2.3 Media Supplements

The following supplements were added to growth media to select and identify bacteria transformed with vector DNA.

(a) Antibiotic Supplements

(i) Ampicillin

A 25 mg/ml stock solution of ampicillin dissolved in distilled water was sterilised by filtration through a 0.22 μm filter. Aliquots of 1 ml were stored at -20 °C in the dark. Once growth media cooled to 50 - 55 °C, added ampicillin to a final concentration of 100 μg/ml. Agar plates supplemented with the antibiotic could be stored at 4 °C for up to 4 weeks.

(ii) Tetracycline

A stock solution 5 mg/ml of tetracycline was stored as 1 ml aliquots at -20 °C in the dark. The appropriate amount of tetracycline was dissolved in 50% ethanol. Filter sterilisation was not required. Media was cooled to 50 - 55 °C before addition of antibiotic to 20 µg/ml final concentration. Tetracycline plates were stored at 4 °C for up to 4 weeks.

(b) *lac* Operon Inducer

Isopropyl-β-D-thiogalactopyranoside (IPTG)

Made a 1 M stock solution by adding 2 g of IPTG to 8 mls distilled water. Increased volume to 10 mls once dissolved. Filter sterilised solution through a 0.22 µM filter and stored as 1 ml aliquots at -20 °C.

(c) Chromogenic Substrate

X-gal (5-bromo-4-chloro-3-indolyl-β-galactoside)

Made a 2% stock solution by dissolving 20 mg X-gal in dimethylformamide. The solution was stored in a glass container in the dark at -20 °C. To select for cells harbouring recombinants by α-complementation, spread 2 µl of 1 M IPTG and 40 µl of 2 % X-gal over each agar plate.

2.3 Growth of Axenic *Anabaena variabilis* ATCC 29413 Cultures

2.3.1 Culturing *Anabaena variabilis*

The filamentous, nitrogen fixing cyanobacterium, *A. variabilis*, was grown in a medium based on BG11o, which lacks nitrates (Rippka *et al.*, 1979). The recipe for BG11o is shown in Table 2.4. Batch cultures of 100 mls were grown in 250 ml conical flasks, with constant shaking, at ambient temperature and a photon fluence rate of 30 - 50 µmol m^{-2} s^{-1} at the surface of the vessel. Cells were grown for 3 - 4 weeks or until the late the log phase of growth was reached. The concentration of chlorophyll *a* was measured to determine the stage of growth the cells had reached. To do this, 1 ml of culture was centrifuged at 12 krpm for 10 minutes. The cell pellet was resuspended in 1 ml methanol in the dark and kept at 4 °C for 12 hours. Centrifuged

suspension as before to sediment cellular debris and measured absorbance of supernatant at 663 nm on a Philips PU 8720 UV/VIS scanning spectrophotometer. The chlorophyll *a* concentration was calculated according to MacKinney (1941).

Table 2.4 Composition of BG11o medium.

Ingredient	Amount in BG11o medium [1] (g/L)
K_2PO_4	0.04
$MgSO_4.7H_2O$	0.075
$CaCl_2.2H_2O$	0.036
$FeSO_4.7H_2O$	0.006
EDTA	0.001
Na_2CO_3	0.02
Trace elements [2]	1 ml/L

[1] Adjusted pH of medium to 7.4 with 1 M HCl prior to autoclaving. [2] Trace elements (g/L): H_3BO_4, 2.86; $MnCl_2.4H_2O$, 1.81; $ZnSO_4.7H_2O$, 0.22; $NaMoO_4.2H_2O$, 0.39; $CuSO_4.5H_2O$, 0.079; $Co(NO_3)_2.6H_2O$, 0.049.

2.3.2 Examining *Anabaena variabilis* Cultures for Contamination

Samples of *A. variabilis* cultures were examined under a light microscope at all magnifications (x 100, x 500, x 1000). In a pure axenic culture, only heterocystous filaments were observed.

2.4 General Methods

2.4.1 pH Measurement

pH measurements were made using a Radiometer Model 26 pH meter (Copenhagen, Denmark) with a Russel combination electrode (Auchtermuchty, Scotland) at room temperature.

2.4.2 Quantitation of DNA

Nucleic acid concentrations were determined spectrophotometrically at 260 nm (Sambrook *et al.*, 1989) with a Philips PU 8720 UV/VIS scanning spectrophotometer using quartz cuvettes. An absorbance of 1 in a 1 cm path length cuvette corresponded to 50 µg/ml of double stranded DNA and 20 µg/ml of single stranded oligonucleotides. The ratio of OD_{260nm}/OD_{280nm} provided an estimate of DNA purity. Pure preparations of DNA had OD_{260nm}/OD_{280nm} values between 1.8 - 2. A value below 1.8 indicated phenol or protein contamination.

2.4.3 Storage of Bacterial Strains

Stored bacterial strains as glycerol stocks. These were made by adding glycerol to growing bacterial cultures to a final concentration of 20%. This culture stock could be stored at -80 °C for 5 - 10 years. Small liquid cultures were also stored for 24 - 48 hours at 4 °C. Alternatively, cultures were streaked onto suitable agar plates and stored at 4 °C for 4 - 6 weeks.

2.5 Small Scale Plasmid Preparation

2.5.1 Modified Holmes and Quigley Method

This protocol is based on the Holmes and Quigley method (1981). Inoculated 1 ml LB (plus antibiotic) with appropriate transformed *E. coli* strain. Collected cell pellet by centrifugation. Completely resuspended pellet in 200 µl STET buffer (0.1 M NaCl, 10 mM Tris/HCl pH 8, 1 mM EDTA pH 8, 5% Triton X 100). Added 16 µl 10 mg/ml lysozyme to cell suspension and incubated at 37 °C for 2 minutes. Boiled tubes for 1 minute then centrifuged for 10 minutes at 12,000 rpm. Removed pelleted cellular debris with a sterile tooth pick. Added 100 µl isopropanol to the supernatant and left at room temperature for 5 minutes. Harvested precipitate by spinning at 12,000 rpm, 10 minutes, 4 °C. Washed pellet with ice cold 70% ethanol then dissolved DNA in 30 µl

TE. Added RNase to solution immediately before loading sample into an agarose / ethidium bromide gel. Yields of up to 3 μg DNA were obtained using this method. Plasmid DNA was suitable for restriction analysis but was of poor quality for other manipulations.

2.5.2 Promega's Wizard Plasmid Purification Kit

Purified 3 - 5 μg good quality DNA using protocol recommended by manufacturer. DNA was suitable for restriction digestions, ligations and transformations.

2.6 Large Scale Plasmid/ds M13 Phage DNA Preparation

2.6.1 Alkaline Lysis (modified Birnboim and Doly) Method

This method is a scaled up version of the Birnboim and Doly alkaline lysis method (1979) modified by Sambrook (1989). The following protocol omits the chloramphenicol amplification step designed to increase the yield of low copy number plasmids. This is because only high copy number plasmid and ds M13 phage DNA were purified using this method.

Inoculated 20 mls LB medium with a single colony or plaque of *E. coli* transformed with the appropriate plasmid or phage. The culture was grown to late log phase with antibiotic (if necessary). Inoculated 500 mls LB medium (in a 2 L flask) with the 20 ml culture and grew for 12 - 16 hours, shaking vigorously at 37 ^{o}C. Cooled cells on ice then centrifuged at 5 krpm, 4 ^{o}C for 15 minutes in the Mistral 6L centrifuge. Removed supernatant and resuspended cell pellet in 10 mls lysis solution (50 mM glucose, 25 mM Tris/HCl pH 8, 10 mM EDTA pH 8). Dissolved lysozyme in an additional 10 mls lysis solution, mixed enzyme gently into the cell suspension and left at room temperature for 5 minutes. Added 40 mls of freshly prepared 0.2 M NaOH/1% SDS to the mixture, mixed gently but thoroughly and left on ice for no more than 10 minutes. Next, added 30 mls 5 M potassium acetate pH 4.8 to the suspension ,

mixed several times then left on ice for 10 minutes (during which time a white precipitate formed). Removed precipitate by centrifugation at 5 krpm, 4 ^{o}C for 15 minutes. Filtered supernatant through double layered muslin. Added 0.6 volumes isopropanol to the clear supernatant and left at room temperature for 15 minutes. Recovered precipitate by spinning suspension at 5 krpm at room temperature for 15 minutes. Re-dissolved pellet in 6 mls TE (10 mM Tris/HCl pH 8, 1 mM EDTA) buffer. To purify plasmid or ds phage from any residual chromosomal DNA, RNA or protein, set up a CsCl/ethidium bromide density gradient (Section 2.8). High quality plasmid or phage DNA was isolated using this method and was suitable for dideoxy DNA sequencing as well as other manipulations.

2.6.2 Qiagen Plasmid Preparations

High quality DNA was purified using protocol recommended by manufacturers. Isolated DNA was suitable for all manipulations including dideoxy DNA sequencing.

2.7 Purification of High Molecular Weight *Anabaena variabilis* Genomic DNA

Ten 100 ml cultures of *A. variabilis* were grown to the late log phase of growth in BG11o medium as in Section 2 3 (Rippka *et al.*, 1979). Harvested cells by centrifugation at 6 krpm for 20 minutes. Cellular pellet tended to be very loose so care was taken when removing the supernatant. Sometimes a second centrifugation step was required to recover the majority of cells. Cell pellet was washed in 40 mls of TE buffer (10 mM Tris/HCl pH 8, 1 mM EDTA pH 8) and centrifuged as before. Resuspended pellet in 20 mls of sucrose buffer (50 mM Tris/HCl pH 8, 1 mM EDTA pH 8, 25% w/v sucrose). Cell suspension was treated with 5 mg/ml lysozyme in the presence of 100 mM EDTA pH 8 and shaken gently for 1 hour at 37 ^{o}C. Added 500 mM EDTA pH 8 to a final concentration of 250 mM to the cell mixture and kept on ice for 5 minutes then added 10 mls of lytic mix (60 mM EDTA pH 8, 1% (v/v) Triton X

100, 0.4% (w/v) sodium deoxycholate) and kept on ice for 20 minutes. Added 20% sodium dodecyl sulphate (SDS) to a final concentration of 2.5% to the mixture and incubated for 15 minutes at 45 °C. The lysate was then extracted with an equal volume of phenol by shaking gently at 4 °C for 10 minutes and then centrifuging at 3 krpm, 4 °C for 15 minutes. Retained the top, aqueous layer and back extracted the organic layer with TE buffer. The pooled aqueous layers were then extracted twice with a 1:1 ratio of phenol-chloroform and once with chloroform. To the final aqueous phase, 3 M sodium acetate pH 7.4 was added to a final concentration of 0.3 M along with 0.54 volumes isopropanol to precipitate nucleic acids. Recovered precipitate by centrifugation at 3 krpm, 4 °C for 30 minutes. The pellet was dried and resuspended in 6 mls TE buffer. Separated genomic DNA from plasmid DNA, RNA and protein by CsCl/ethidium bromide density gradient centrifugation (Section 2.8). 50 - 100 µg Genomic DNA was extracted from *A. variabilis* cells using this method.

2.8 Caesium Chloride (CsCl)/Ethidium Bromide Density Gradient Centrifugation

This protocol was used to separate genomic and plasmid DNA from RNA and protein according to their respective densities in the CsCl gradient.

Added 4 mg CsCl per 4.1 mls DNA solution, and 0.8 mls ethidium bromide for every 10 mls DNA/CsCl mixture. The density gradient was created by centrifugation at 45,000 rpm, 25 °C for 36 - 48 hours in a Beckman L8-55M ultracentrifuge. DNA bands could then be recovered using a sterile wide-bore needle and syringe. Extracted DNA solution 5 - 6 times with an equal volume of isopropanol (saturated with CsCl and TE) to remove ethidium bromide. Removed CsCl from DNA solution by dialysing against 2 L of TE at room temperature for 1 hour and again for 24 - 48 hours at 4 °C against a fresh batch of buffer.

2.9 Digestion of DNA with Restriction Enzymes

Restriction digests of DNA were performed with enzymes and buffers from GIBCO BRL, New England Biolabs or Promega and were carried out as in Sambrook *et al.*, 1989. A series of buffers with varying salt concentrations were available for a range of restriction enzymes. Each buffer was supplied as a 10 x stock solution and the appropriate volume added to give the correct final concentration.

2.9.1 Restriction Digestion of Genomic DNA

Preparative digests of genomic DNA were performed in large volumes (up to 100 µl) with 7 - 10 units of enzyme per µg DNA. Mixed genomic DNA gently but thoroughly with a wide bore tip before incubation at 37 °C for 12 - 16 hours.

2.9.2 Partial Digestion of Genomic DNA with *Sau*3AI

Pilot reactions were performed in order to determine the correct ratio of *Sau*3AI and DNA required to result in a favourable size distribution of restriction fragments. Ten tubes were set up each containing 0.5 µg genomic DNA except tube one which had 1 µg DNA. 2 units *Sau*3AI per µg DNA was added to tube one and mixed gently for 2 minutes on ice with a wide bore tip. A half volume of the reaction mixture from tube one was transferred to tube 2 which diluted the enzyme concentration 2 fold. Continued to serially dilute enzyme until the penultimate tube where half the reaction mixture was discarded. So the final tube was used as a no enzyme control. Test reactions were performed in a total volume of 100 µl. Incubated reactions for one hour at 37 °C to digest DNA then at 65 °C for 20 minutes to heat inactivate *Sau*3AI. Electrophoresed DNA through a 0.3% agarose / ethidium bromide gel at 10 mA for 12 - 16 hours (Section 2.11) to ascertain the size distribution of the restriction fragments.

Once the optimum *Sau*3AI / DNA ratio was determined, large scale digestions of genomic DNA were performed. 10 µg Genomic DNA was digested with the optimum dilution of *Sau*3AI previously determined. In case DNA became over or under digested, incubated DNA with *Sau*3AI at dilutions 2 fold higher and 2 fold lower

than optimum enzyme concentration. Aliquots of the reactions were electrophoresed through a 0.3% agarose / ethidium bromide gel as before to check size distribution of fragments. Digested DNA was purified as in Section 2.12.2 for further manipulation.

2.9.3 Restriction Digestion of PCR products, Plasmid or Lambda DNA

Digests were performed in a 20 µl volume with 2 units of enzyme per µg DNA at 37 °C. Length of incubation varied. Generally, plasmid and lambda DNA digested within 2 hours. PCR products were incubated for 2 - 16 hours depending on the efficiency of digestion at their 5' end.

2.9.4 Double Digestions

DNA digestions with 2 restriction enzymes were performed simultaneously if both enzymes digested well in the same buffer. Alternatively, digestions were performed simultaneously in the buffer that suited the enzyme with the more stringent requirements. If this was not suitable, DNA was digested with one enzyme, cleaned up as in Section 2.12.2 then digested with the second enzyme.

2.10 Phosphatasing DNA with Calf Intestinal Alkaline Phosphatase (CIAP)

Terminal 5' phosphates were removed from vector DNA cut with a single restriction enzyme to prevent self ligation. Digested plasmid DNA with appropriate restriction enzyme (Section 2.9.3). Added 1 x CIAP buffer (10 x: 0.5 M Tris/HCl pH 9.3, 10 mM $MgCl_2$, 1 mM $ZnCl_2$, 10 mM Spermidine) and CIAP directly to the digested DNA and increased volume of reaction to 100 µl with water. Incubated DNA at 37 °C for 1 hour then added 2.5 µl 0.5 M EDTA pH 8 to stop the reaction.

The nature of the digested termini determines the amount of CIAP required for each reaction. For protruding 5' termini, 0.02 units CIAP/pmol of ends is required while 1 unit CIAP/pmol of ends is necessary for blunt or recessed 5' termini.

Dephosphorylated DNA was purified using Promega's Magic DNA Clean Up System (Section 2.12.4).

2.11 Agarose Gel Electrophoresis of DNA

Agarose gel electrophoresis of DNA was performed in horizontal, submerged, slab gels in tris-borate buffer at room temperature. Suitable gel concentrations (0.3 - 2.5%) were made by adding the correct amount of agarose to 1 x TBE buffer (89 mM Tris-borate, 2 mM EDTA). Melted agarose by heating the solution to nearly boiling point in a microwave. Once the gel had cooled, added ethidium bromide to a final concentration of 0.5 µg/ml and left to set for 30 minutes at room temperature. A 1/5 to 1/10 volume of agarose gel sample buffer (0.2% bromophenol blue, 50% sucrose in 1 x TBE) was added to DNA samples prior to loading into gel. Electrophoresis was carried out in 1 x TBE with a current of 10 - 100 mA until the dye front had migrated a suitable length of the gel. DNA fragments were visualised using a UV transilluminator. The UV light absorbed by the ethidium bromide through the DNA emitted a red-orange fluorescence. Photographs of the gel were taken as a permanent record.

2.12 Extraction and Purification of DNA

2.12.1 Preparation of Organic Reagents

(a) Phenol

Liquefied and re-distilled phenol was aliquoted and stored at -20 ^{o}C. Equilibrated phenol with 1/2 volume 1 M Tris/HCl pH 8.1. Discarded most of the upper aqueous layer. Phenol was again equilibrated with 1/2 volume distilled water then stored in the dark at 4 ^{o}C for up to 6 weeks.

(b) Chloroform

Little pre-treatment was required. Chloroform was merely saturated with distilled water before use and stored at room temperature.

(c) Phenol / Chloroform (1:1)

The appropriate volume of phenol (prepared as in Section (c) above) and chloroform were mixed in a 1:1 ratio and stored in the dark at 4 ^{o}C for up to 6 weeks.

2.12.2 Purification of DNA by Extraction with Phenol and Chloroform

Extracted DNA with phenol and chloroform to remove enzymes, salts or other reagents.

Increased volume of the DNA solution to 200 µl with TE buffer. Added an equal volume of phenol to DNA solution, mixed thoroughly and centrifuged at 12,000 rpm for 3 minutes to separate the organic and aqueous layers. Retained the upper phase and back extracted the organic layer with 80 µl TE as before. Pooled the two aqueous layers then extracted DNA twice more with phenol and three times with chloroform. To the final aqueous phase, added 3 M sodium acetate pH 7.4 to a final concentration of 0.3 M and 2.15 volumes ethanol. Precipitated DNA at -20 ^{o}C for at least 1 hour. Recovered DNA by centrifuging at 12,000 rpm for 15 minutes. Washed pellet in 70% ice - cold ethanol. Dried pellet and re-dissolved DNA in the appropriate volume of TE or distilled water.

2.12.3 Purification of DNA from Low Melting Point Agarose Gels

This technique was used for the recovery of DNA from low melting point agarose gels and was based on a method found in Sambrook *et al.*, 1989.

Low melting point agarose gels were prepared in the same way as in Section 2.11. However, to ensure full polymerisation, the gel was left to solidify at room temperature for 30 minutes then at 4 ^{o}C for a further 30 minutes. A maximum current of 70 mA was applied to the gel during electrophoresis. The desired DNA segment(s) was cut out of the gel using a sterile needle and syringe. 0.3 mls TE buffer was added to the gel slice which was then melted thoroughly at 65 ^{o}C for 10 minutes then cooled

to 37 °C. Extracted DNA with phenol (heated to 37 °C) and chloroform as in Section 2.12.2. Precipitated DNA with ethanol (Section 2.12.2), washed pellet in 70% ice - cold ethanol and re-dissolved DNA in the buffer of choice for subsequent manipulations.

2.12.4 Purification of DNA Using Commercial Kits

These kits were a quick and easy way to remove enzymes, salts or deoxynucleotides from DNA solutions. The following kits were used: Wizard DNA Clean Up Kit (Promega Corporation) and Wizard PCR Clean Up Kit (Promega Corporation). The protocols recommended by the manufacturer were followed at all times.

2.13 Synthesis of Oligonucleotides

Oligonucleotides were used as PCR primers for the amplification of defined fragments of genomic and plasmid DNA. Dr. V.B. Math (University of Glasgow, Scotland) synthesised the oligonucleotides using a solid-phase phosphotriester method on an Applied Biosystems model 280A DNA synthesiser.

2.14 Polymerase Chain Reaction

2.14.1 PCR Amplification of DNA

Sterile filter tips and fresh, sterile reagents were used at all times to avoid contamination. The following reagents were added to a sterile, 0.5 ml PCR tube:
0.1 - 1 µg DNA, 1 x *Taq* DNA polymerase reaction buffer (10 mM Tris/HCl pH 9, 50 mM KCl, 1% Triton X 100), 1.5 - 4 mM $MgCl_2$, 100 - 200 µM dNTP, 40 - 100 pmol primers, 2.5 units *Taq* DNA polymerase.

Increased volume of reaction to 100 µl with sterile H_2O, mixed thoroughly and overlaid with 100 µl paraffin oil. PCR was performed using a DNA Thermal

Cycler (Perkin Elmer Cetus Corporation). The reaction conditions are shown in Table 2.5.

Table 2.5 Reaction conditions used for PCR.

Step	Temperature (^{o}C)	Time (minutes)
Denaturation	94	3 - 5
Cycle *	94	1
	42 - 55	1 - 2
	72	1 - 2
Extension	72	5
Soak	4	indefinitely

* Repeated cycle step 20 - 40 times.

2.14.2 Purification and Analysis of PCR Products

Paraffin oil was removed from PCR products by extraction with an equal volume of chloroform saturated with H_2O. Analysed PCR products by electrophoresis through an agarose gel (Section 2.11). 10% of the PCR reaction volume was used for gel analysis.

2.14.3 Sub Cloning PCR Products

The PCR products of interest were cloned into the vectors M13mp19RF or pBluescript SK- for further analysis. Purified PCR products from low melting point agarose gels as described in Section 2.12.3. Vector and insert DNA were then digested using the relevant restriction enzymes (Sections 2.9.3/4). Cleaned vector DNA by organic extraction and ethanol precipitation (Section 2.12.2) or by using one of the commercial kits described in Section 2.12.4.

Ligation mixtures contained 50 ng vector and 50 - 100 ng of insert DNA. Added the appropriate amount of water to the reaction to give a final volume of 10 µl.

Incubated reactions at 56 °C for 5 minutes to melt cohesive termini, then added 1 µl 10 x ligase buffer (1 x: 30 mM Tris/HCl pH 7.8, 10 mM $MgCl_2$, 10 mM DTT, 0.5 mM ATP) and 1 unit (Weiss *et al.*, 1968) of bacteriophage T4 DNA ligase. The reaction was incubated for 16 hours at 8 - 10 °C.

Ligation mixtures were transformed into the appropriate *E. coli* host cells (Section 2.17).

2.15 Creation of an *Anabaena variabilis* Genomic DNA Library in pBluescript SK-

A collection of recombinant clones representing the genome of *A. variabilis* was constructed by digesting purified genomic DNA and plasmid vector with the appropriate restriction enzyme then ligating the genomic fragments into pBluescript SK.

Digested 20 µg genomic DNA with 7 - 10 units of the appropriate restriction enzyme per µg DNA (Section 2.9.1). Inactivated restriction enzymes by incubating at 65 °C for 20 minutes. The restriction pattern was checked by running a sample of the digested DNA through an agarose / ethidium bromide gel (Section 2.11). Digested DNA was size fractionated by separating fragments through a l.m.p. agarose / ethidium bromide gel and cutting out a gel slice containing DNA in the desired size range. To prevent DNA damage, the electrophoresed DNA was not exposed to UV light. Instead lambda size markers, electrophoresed alongside the DNA sample, were visualised and photographed beside a ruler. Using the markers and ruler as a reference, genomic DNA in the appropriate size range was purified for the l.m.p. agarose gel (Section 2.12.3). Cut 5 µg pBluescript SK- with the appropriate restriction enzymes as indicated in Section 2.9.3. Phosphatased the plasmid with calf intestinal alkaline phosphatase (Section 2.10) to minimise self ligation. After incubation, purified vector DNA with Promega's Magic DNA Clean Up Kit (Section 2.12.4).

Plasmid and genomic fragments were quantitated (Section 2.4.2) then mixed in a 1:1 - 4:1 ratio (plasmid: genomic DNA) with 1 x ligation buffer (1 x: 30 mM Tris/HCl pH 7.8, 10 mM $MgCl_2$, 10 mM DTT, 0.5 mM ATP) and 3 units T4 DNA

ligase. The volume of the reactions were increased to 10 µl with distilled water and were incubated at 8 - 10 °C for at least 20 hours.

The ligation products were then transformed into competent *E. coli* DH5α or *Epicurian Coli* XL2-Blue MRF' cells (Section 2.17). Plated up to 500 colonies onto 9 cm agar plates or 5,000 colonies onto 15 cm agar plates containing ampicillin, IPTG and X-gal. Grew plates overnight at 37 °C and examined for cells containing recombinants clones (Section 2.17.4).

2.16 Complementation

Digested *A. variabilis* genomic DNA with the appropriate restriction enzyme and cloned the resulting fragments into pBluescript SK- as described in Section 2.15. Transformed 50 ng of the plasmid library into competent *E. coli* AB2829 (*aro*A$^-$) and plated up to 100 µl cell mixture onto minimal medium agar plates (Section 2.2.2) containing ampicillin as described in Section 2.17. Grew plates at 37 °C for 2 - 3 days then examined for bacterial colonies.

The resulting colonies were picked and grown overnight on LB/ampicillin plates. Plasmid DNA was purified (Section 2.5.2) and analysed by agarose gel electrophoresis (Section 2.11). Recombinant plasmid DNA was re-transformed into *E. coli* AB2829 and grown on minimal medium (as above) to double check that the host cell mutation had been complemented by protein expressed by the plasmid DNA.

2.17 Transformation of *E. coli* with Plasmid or M13 Bacteriophage DNA

2.17.1 Preparation of Competent Cells

Epicurian Coli strains SURE^{TM} and XL2-Blue MRF' were supplied as competent cells by Stratagene Ltd., Cambridgeshire, England (Table 2.1). However, other strains used for plasmid and phage transformations had to be made competent using the methods outlined below.

(a) $CaCl_2$ Method

This method was routinely used to make competent *E. coli* TG1, the host strain for bacteriophage M13. Single colonies of TG1 were picked from minimal medium plates (Section 2.2.2) to select for the F' episome. 10 mls 2 x TY was inoculated with a single colony picked from a stock plate and incubated overnight (12 - 16 hours) at 37 ^{o}C. Inoculated 20 mls growth medium with 1 ml overnight culture and incubated at 37 ^{o}C until OD_{600nm} of the mixture reached 0.3. Stored cells temporarily on ice then centrifuged at 5000 rpm, 5 minutes, 4 ^{o}C. Cells were resuspended in 1/2 volume of ice cold 0.1 M $CaCl_2$ and left on ice for 30 minutes. Harvested competent cells as before and dissolved pellet in 0.9 mls ice cold 0.1 M $CaCl_2$. Stored cells at 4 ^{o}C for up to 24 hours or until ready for use. At best, a transformation efficiency of 10^6 colonies per μg DNA was achieved.

(b) TFB Method (Modified Hanahan Method)

Required a high transformation efficiency when constructing plasmid libraries of genomic DNA. Routinely used this protocol to make competent cells of *E. coli* DH5α and AB2829. Obtained up to 10^7 transformants per μg supercoiled plasmid DNA.

Single colonies of *E. coli* DH5α or AB2829 were picked from freshly grown LB plates and used to inoculate 10 mls Ψ broth. Grew culture overnight at 37 ^{o}C then stored on ice. Inoculated 1 ml of this stationary phase culture into 20 mls of Ψ broth and incubated at 37 ^{o}C, shaking vigorously until the $O.D_{600nm}$ reached 0.3. Cooled cells on ice for 10 minutes then harvested at 5000 rpm, 4 ^{o}C for 5 minutes. Removed supernatant and drained off any traces of Ψ broth. Carefully resuspended cells in 5 mls ice cold TFB 1 solution (100 mM RbCl, 50 mM $MnCl_2$, 30 mM potassium acetate, 10 mM $CaCl_2$, 15% glycerol, pH adjusted to 5.8 with glacial acetic acid) and left on ice for 1 hour. Centrifuged cells as before and gently resuspended cell pellet in 0.8 mls TFB 2 solution (10 mM MOPS, 10 mM RbCl, 75 mM $CaCl_2$, 15% glycerol, pH adjusted to 7 with glacial acetic acid). Competent cells were then kept on ice ready for use.

2.17.2 Transformation of Competent Cells

Epicurian Coli strains SURE^{TM} and XL2-Blue MRF' were transformed with vector DNA according to the manufacturer's instructions. The volume of the ligation reactions were increased to 50 µl with TNE buffer (50 mM Tris/HCl pH 7.5, 50 mM NaCl, 1 mM EDTA pH 8). 200 µl Aliquots of competent cells were mixed thoroughly but gently with no more than 100 ng of the ligation products in sterile 1.5 ml eppendorf tubes. Incubated mixture on ice for 30 minutes, heat shocked at 37 ^{o}C for 90 seconds then transferred cells back to ice for a further 15 - 30 minutes.

2.17.3 Plating Transformed Cells

(a) Plating Cells Transformed with M13

This method was applied for both *E. coli* TG1 and SURE^{TM}. Transformed cells were added to a mixture of 40 µl 100 mM IPTG; 40 µl 2% X-gal and 200 µl fresh *E. coli* (made fresh cells by inoculating 40 mls 2 x TY medium with 2 mls of an overnight culture of the appropriate strain and incubating at 37 ^{o}C for 3 hours). Added 3 mls of molten H top agar to the cell mixture and poured onto pre-warmed H agar plates. Left agar to set at room temperature then incubated plates overnight at 37 ^{o}C.

(b) Plating Cells Transformed with Plasmid DNA

This method applies to *E. coli* DH5α and XL2-Blue MRF' which grew on Ψ and LB media respectively. 0.8 mls of Ψ or LB medium was added to transformed cells which was then incubated at 37 ^{o}C, shaking gently, for 1 hour 45 minutes. This allowed expression of the vector encoded antibiotic resistance gene. Up to 200 µl of the cell mixture was plated onto Ψ or LB plates containing ampicillin and/or IPTG and X-gal for selection of recombinants by α complementation (Section 2.17.4). Plates were dried at room temperature then incubated at 37 ^{o}C. For plating out transformed *E. coli* AB2829 refer to Section 2.16.

2.17.4 Selection of Recombinant Plasmid and M13 Vectors

M13 and pBluescript SK- have the *lac* Z' gene which encodes the first 146 amino acids of the N terminal region of *E. coli* β-galactosidase. Host cells for these vectors, e.g.. *E. coli* TG1, SURE^{TM}, DH5α and XL2-Blue encode the carboxy region of the enzyme. When the cell has been transformed with non-recombinant vector the N terminal fragment of β-galactosidase associates with the C terminal portion to produce active enzyme. This process is called α-complementation. The polycloning region of the vector is tagged onto the 3' end of the *lac* Z' gene but this only adds a few extra amino acids onto the N terminus of the enzyme so the reading frame is not disrupted. When transformed cells are grown on medium containing the *lac* operon inducer, IPTG, functional β-galactosidase cleaves the chromogenic substrate (X-Gal) so forming bromochloroindole, which produces the blue coloured plaques or colonies. α Complementation cannot occur in cells transformed with recombinant vector since the insert DNA interrupts the reading frame. Therefore, white plaques or colonies are formed instead. False positive whites are rare and may be caused by incorrect self ligation of the vector (Yanisch-Perron *et al.*, 1985).

2.18 Creation of an *Anabaena variabilis* Genomic Library in Lambda FIX II

2.18.1 Preparation of *Anabaena variabilis* Genomic DNA

Constructed a phage library of *A. variabilis* genomic DNA in lambda FIX II (Stratagene Ltd., Cambridgeshire, England). The vector was supplied pre-cut with *Xho*I and the restriction sites were partially filled in with dCTP and dTTP. The vector arms were purified from the stuffer fragment to minimise self ligation of the vector.

10 μg *A. variabilis* DNA was partially digested with 0.031 units *Sau*3AI/μg DNA (Section 2.9.2) resulting in a range of fragments 23 - 5-kb in size. Purified digested DNA as in Section 2.12.2 except ethanol precipitated at -20 ^{o}C for 12 - 16

hours and harvested DNA by centrifugation at 12 krpm at 4 °C for 30 minutes. Re-dissolved DNA in 10 µl TE.

Partially filled in *Sau*3AI restriction sites with dATP and dCTP to complement the ends of lambda FIX II using the Klenow fragment of *E. coli* DNA polymerase I. Mixed 10 µg *Sau*3AI digested DNA (in 10 µl TE) with 3 µl Klenow buffer (10 x: 500 mM Tris/HCl pH 7.5, 70 mM $MgCl_2$), 5 µl 10 mM dATP, 5 µl 10 mM dGTP, 10 units (1u/µg DNA) Klenow fragment. Increased volume of reaction to 30 µl with distilled water and incubated for 30 minutes at 37 °C. Purified DNA as in Section 2.12.2 except increased reaction volume from 30 µl to 300 µl with distilled water before addition of sodium acetate and ethanol. This reduced the amount of salt precipitation. Dissolved DNA in 10 µl TE buffer.

The principle of the partial fill-in technique is illustrated in Chapter 5, Figure 5.1.

2.18.2 Quantitation of *A. variabilis* Genomic DNA

A DNA fluorimeter type TKO 100 was used to quantitate the *Sau*3AI digested, partially filled-in genomic DNA because it can accurately measure small amounts of DNA. A dye, bis benzimidizole (commonly known as Hoechst 33258), binds to the DNA and fluoresces at a wavelength of 460 nm when initially exposed to light at a wavelength of 365 nm.

The fluorimeter was standardised using 1 ml of the fluorescent dye solution (For 100 mls: 90 mls double distilled water, 10 mls 10 x TNE (1 x: 50 mM Tris/HCl pH 7.5, 50 mM NaCl, 1 mM EDTA), 10 µl Hoechst 33258). 1 µl of the DNA sample (diluted 5 fold) was then added to 1 ml dye solution. The fluorimeter reading obtained was compared to those given by calf thymus DNA standards and the concentration of the DNA sample was determined.

2.18.3 Ligation of *A. variabilis* Genomic DNA to Lambda FIX II

Ligation reactions contained 1 μg lambda FIX II (1 μl) and 0.4 μg *Sau*3AI digested, partially filled-in *A. variabilis* genomic DNA. Control reactions were performed in parallel with 1 μg lambda FIX II and 0.3 μg pMF (test insert supplied by Stratagene Ltd, Cambridgeshire, England). To each reaction tube added 1 μl 10 x ligase buffer (1 x: 30 mM Tris/HCl pH 7.8, 10 mM $MgCl_2$, 10 mM DTT, 0.5 mM ATP), 1 μl (3 Weiss units) T4 DNA ligase and increased volume to 10 μl with distilled water. Incubated reactions at 8° C for 24 hours.

2.18.4 Packaging Reactions

Packaged phage DNA into their coat proteins using the protocol and reagents supplied by Stratagene, Cambridgeshire, England. Added 0.25 - 0.35 μg ligation products to a partially thawed, freeze - thaw extract. Kept mixture on ice then added 15 μl freshly thawed sonic extract. Mixed gently by stirring but did not pipette. Incubated phage mixture at 22 °C for 1.5 - 2 hours. Added 500 μl SM buffer (1 litre: 5.8 g NaCl, 2 g $MgSO_4.7H_2O$, 50 mls 1 M Tris/HCl pH 7.5, 5 mls 2% gelatin solution) and 20 μl chloroform. Mixed then centrifuged briefly to sediment any debris. Stored packaged phage at 4 °C. To optimise efficiency of packaging reaction, melted freeze - thaw and sonic extracts just before use.

2.19 Titre Determination of Phage Genomic Library

2.19.1 Preparation of Host Cells for Infection with Phage Library

Picked a single colony of *E. coli* XL1-Blue MRA (P2) from an LB plate, and inoculated into 20 mls NZY medium. Grew culture at 37 °C for 4 - 6 hours ensuring OD_{600nm} did not exceed a value of 1. Harvested cells at 3000 rpm for 10 minutes then

resuspended in 10 mM $MgSO_4$ until OD_{600nm} = 0.5. Stored cells at 4 °C for up to 48 hours.

2.19.2 Infection of *E. coli* XL1-Blue MRA (P2) with Phage Library

Made serial dilutions of packaged phage and mixed 10 µl of each dilution with 200 µl host cells prepared as in Section 2.19.1. Incubated at 37 °C for 15 minutes shaking gently. Added 3 mls molten NZY top medium (equilibrated to 45 °C) to mixture and plated onto 9 cm, dry, pre-warmed NZY plates. Allowed top agar to harden for 30 minutes then incubated plates at 37 °C for 8 hours. Calculated the phage titre using the following equation: (number of plaques x dilution factor x total packaging volume) ÷ (number of µg DNA x number of µl mixture plated).

2.19.3 Selection of Lambda FIX II Recombinants

Lambda FIX II system is sensitive to P2 inhibition (spi). Lambda phage with *red* and *gam* genes cannot grow in host cells containing a P2 lysogen (e.g., *E. coli* XL1-Blue MRA (P2)) whereas lambda phage lacking these genes can grow in such hosts. The *red* and *gam* genes of lambda FIX II are situated in the stuffer fragment, so wild type phage cannot multiply in *E. coli* XL1-Blue MRA (P2). However, when the stuffer fragment is replaced by foreign DNA, lambda FIX II becomes red^-, gam^- and can grow in the P2 lysogenic strain. In principle, plating the phage library in *E. coli* XL1-Blue MRA (P2) would only allow recombinant phage to grow.

2.20 Amplification of Phage Genomic Library

The phage library was amplified to create a high titre stock. Prepared *E. coli* XL1-Blue MRA (P2) as in Section 2.19.1. Mixed 3 x 125 µl packaged phage with 600 µl of the host cells and incubated mixture at 37 °C for 15 minutes. Added 8 mls molten NZY top medium (cooled to 45 °C) to mixture and plated onto 15 cm, dry, pre-warmed NZY plates. Allowed top agar to harden then incubated plates at 37 °C for 6 - 8 hours

until individual plaques could be resolved. Confluent growth was avoided to decrease the possibility that a single cell would be infected by more than one phage particle. Overlaid plates with 8 mls SM buffer (1 litre: 5.8 g NaCl, 2 g $MgSO_4.7H_2O$, 50 mls 1 M Tris/HCl pH 7.5, 5 mls 2% gelatin solution) and left overnight at 4 ^oC shaking gently to allow phage to diffuse into solution. Decanted phage solution and washed plates with a further 2 mls SM buffer. Stored pooled phage stock solution in polypropylene tubes. Liberated remaining phage with chloroform added to a final concentration of 5%. Clarified supernatant by centrifugation at 3000 rpm for 10 minutes at 4 ^oC. Added chloroform to amplified library to a final concentration of 0.3% and stored at 4 ^oC. Titred phage as in Section 2.19.

2.21 Radiolabelling ds DNA Probes with [α-^{32}P]dCTP

Megaprime labelling kit was obtained from Amersham International, Slough, England. For standard reactions, labelled 25 ng DNA. However, 100 ng DNA was labelled for screening plasmid and phage genomic libraries. The technique used is based on a method developed by Feinberg and Vogelstein 1983, 1984.

The following reagents were added to a sterile, 1.5 ml, screw top eppendorf tube: 25 ng or 100 ng ds DNA dissolved in TE buffer or H_2O; 5 µl nonamer primers; H_2O to give a total volume of 50 µl. Boiled solution for 5 minutes to denature DNA. Centrifuged tube briefly and kept DNA at room temperature when adding next set of reagents: 10 µl labelling buffer (dATP, dGTP, dTTP in Tris/HCl pH 7.5, 2-mercaptoethanol, $MgCl_2$); 50 µCi [α -^{32}P]dCTP; 2 µl Klenow fragment. Mixed contents of tube gently but thoroughly, centrifuged briefly and incubated for 10 minutes at 37 ^oC. Added 5 µl 0.2 M EDTA to mixture to stop the reaction. Stored labelled DNA on ice until required.

2.22 Purification of Radiolabelled DNA Probes and Measuring their Specific Activity

On completion of the radiolabelling reaction, the probe was passed through a packed Sephadex G50 column. The chromatography matrix retained any free nucleotides and allowed the radiolabelled probe to pass straight through the column.

Equilibrated Sephadex G50 with TE buffer by autoclaving at 15 p.s.i. for 20 minutes. To prepare column, plugged a 1 ml sterile syringe with siliconised glass wool. Filled the syringe with equilibrated Sephadex G50, placed syringe in a 15 ml corex tube and centrifuged at 3000 rpm for 4 minutes. Repeated this procedure until the column was packed to a 1 ml volume. Added 100 µl of TE to the matrix and centrifuged as before to ensure the correct sample volume was eluted. Placed a de-capped, 1.5 ml microfuge tube below the column and applied the radiolabelled DNA sample (made up to 150 µl with TE) to the matrix. Centrifuged column as before. Eluted DNA was collected in the microfuge tube and stored at 4 ^{o}C until required.

The specific activity of the probe was then measured. Spotted a 1µl aliquot of radiolabelled DNA onto Whatmann DE81 paper, dried paper then placed into a scintillation vial with 5 mls scintillation fluid (Ecoscint from National Diagnostics). The number of counts of radioactivity emitted per minute per µg DNA was measured using a scintillation counter.

2.23 Southern Blotting

Digested genomic DNA with restriction enzymes as described in Sections 2.9.1 or 2. Separated DNA fragments by electrophoresis through a 0.8% agarose gel (Section 2.11). Soaked gel in several volumes of denaturing solution (1.5 M NaCl, 0.5 M NaOH) shaking gently for 1 hour. Rinsed gel with distilled water and incubated in neutralising solution (1 M Tris/HCl pH 8, 1.5 M NaCl) for 1 hour with 2 buffer changes. If the transfer of high molecular weight DNA was required, DNA was depurinated in 0.25 M HCl for 15 minutes prior to denaturation. Once neutralised, blotted DNA onto nylon membrane (Hybond N, Amersham International, Slough,

England) using 10 x SSC (20 x SSC: 0.15 M NaCl, 0.015 M tri-sodium citrate) as described by Sambrook *et al.*, 1989. Dried membrane at room temperature then baked between 2 sheets of 3MM paper at 80 °C for 2 hours to fix DNA. If not for immediate use, the filters could be stored between two sheets of 3MM paper in a cool, dry place until required.

2.24 Southern Hybridisation of Radiolabelled DNA Probes to Membrane Bound DNA

Floated the baked filters in 4 x SET (20 x SET: 0.4 mM Tris/HCl pH 8, 3 M NaCl, 20 mM EDTA) until wet then soaked them for 5 minutes at room temperature. Placed filters in a heat sealable bag and added 0.2 mls of prehybridisation solution (4 x SET buffer, 5 x Denhardt's solution, 0.1% SDS, 0.1% tetra-sodium pyrophosphate, 100 µg/ml denatured salmon sperm DNA) for each square centimetre of filter. Sealed bag and incubated filters at the required temperature (50 - 65 °C) for 12 - 16 hours, shaking gently. Removed prehybridisation solution then added 50 µl/cm^2 of hybridisation solution (prehybridisation solution containing radiolabelled probe heated to the appropriate temperature) to the filters. Denatured radiolabelled probe (prepared as in Section 2.21) by boiling for 5 minutes and added to hybridisation solution. Sealed bag and incubated for 12 - 24 hours shaking gently at the appropriate temperature.

Filters were washed once for 20 minutes with each of the following buffers described below at the temperature used for hybridisation. However, the filters were only washed at a high stringency if necessary.

500 mls:	2 x SET	0.1% SDS	0.1% NaPPi
500 mls:	1 x SET	0.1% SDS	0.1% NaPPi
250 mls:	0.5 x SET	0.1% SDS	0.1% NaPPi
250 mls:	0.1 x SET	0.1% SDS	0.1% NaPPi

2.25 Removal of Bound Radiolabelled Probe from Filters

Poured boiling 0.1% SDS over filter. Allowed solution to cool to room temperature and repeated procedure 2 - 3 times more. Autoradiographed filter to ensure all of the membrane bound probe had been removed.

2.26 Colony Hybridisation

The technique described below was used to replicate bacterial colonies transformed with recombinant plasmid onto nylon membrane. DNA was then liberated from bacterial cells, denatured and irreversibly bound to the filter. A radiolabelled DNA probe was hybridised to the immobilised DNA to identify recombinants carrying related sequences.

2.26.1 Colony Lysis and Binding DNA to Nylon Membranes

Labelled a nylon membrane disc (15 cm in diameter) with a ball point pen and placed carefully onto the surface of the agar plate, avoiding air bubbles. Left for 1 minute and, with a sterile syringe needle, keyed the filter to the master plate in an asymmetric pattern. Lifted the filter from the plate and transferred to a fresh Ψ or LB agar plate (containing the appropriate antibiotic) colony side up. Made a duplicate filter by placing a fresh membrane disc on top of the original filter, taking care not to trap air bubbles. Keyed duplicate to original filter. Incubated the filter plates and master plates at 37 ^{o}C for approximately 3 hours until the colonies were 1 - 2 mm in size.

Placed each set of duplicate filers on Whatmann 3MM paper saturated with 10% SDS for 3 minutes. Ensured upper side of filters did not get wet. Transferred filters, keeping them horizontal, to 3MM paper soaked with denaturing solution (1.5 M NaCl, 0.5 M NaOH) for 7 minutes then neutralising solution (1 M Tris/HCl pH 8, 1.5 M NaCl), twice, for 3 minutes. Lastly, peeled apart filter sandwich and saturated membranes with 2 x SSC (20 x SSC: 0.15 M NaCl, 0.015 M tri-sodium citrate) for 5 minutes after which they were dried at room temperature for at least 30 minutes on fresh

3MM paper. Between each transfer, blotted membranes on dry 3MM paper to remove excess fluid. Fixed DNA to nylon membrane by exposing filters to 240 nm U.V. light using a U.V. transilluminator (Ultra-violet Products Inc.) for 2 - 3 minutes each.

2.26.2 Hybridising Radiolabelled DNA Probes to Plasmid Library

Floated the dried membranes in 2 x SSC until wet then submerged them for 5 minutes. Transferred filters to 250 mls of pre-washing solution (5 x SSC, 0.5% SDS, 1 mM EDTA pH 8) and soaked for 30 minutes at 50 °C. Repeated pre-washing step. Reduced background hybridisation by gently wiping filters with mediwipe tissues to remove bacterial debris. Incubated filters in 50 mls of prehybridisation solution (4 x SET, 5 x Denhardt's solution, 0.1% SDS, 0.1% tetra-sodium pyrophosphate, 50 µg/ml denatured salmon sperm DNA) for 12 - 16 hours at the appropriate temperature. Added radiolabelled denatured probe (prepared as in Section 2.21) to 25 mls freshly made prehybridisation solution and incubated for 24 - 48 hours at the required temperature. Filters were incubated in sealed hybridisation tubes that rotated slowly in a Techne hybridisation oven (model HB - 1D).

2.26.3 Washing Procedure

Washed filters once in each of the following buffers for 1 hour at room temperature or the temperature used for hybridisation.

500 mls:	4 x SET	0.1% SDS	0.1% NaPPi
500 mls:	2 x SET	0.1% SDS	0.1% NaPPi
500 mls:	1 x SET	0.1% SDS	0.1% NaPPi
500 mls:	0.5 x SET	0.1% SDS	0,1% NaPPi

After washing, the filters were marked with radioactive ink so that the autoradiographs could be aligned with the master plates at a later stage (Section 2.26.4). The filters were autoradiographed after each wash step to examine the hybridisation signal. If a

higher stringency wash was required, immersed filters in 0.1 x SET, 0.1% SDS, 0.1% NaPPi for 1 hour at the temperature used for hybridisation, then autoradiographed.

2.26.4 Purification of Putative Positive Clones from Plasmid Library

Aligned the autoradiograph with the master plate using the signals from the radioactive ink on the film and the asymmetric pattern of marks keyed in the plate. Identified the area of the plate that gave rise to the putative positive signal(s) and recovered it as a plug of agar using the wide end of a sterile Pasteur pipette. Inoculated 1 ml LB/ampicillin broth with the agar plug and vortexed briefly. Plated serial dilutions of the cell mixture onto 9 cm LB/ampicillin plates to determine the dilution required to give rise to approximately 200 colonies per plate. Repeated screening procedure as in Sections 2.26.1 - 3.

2.27 Plaque Hybridisation

2.27.1 Transfer of DNA to Nylon Membranes

Infected 600 µl *E. coli* XL1-Blue MRA (P2) (prepared as in Section 2.19.1) with 50,000 plaque forming units of the amplified phage library (Section 2.20). Plated this onto a 15 cm, dried, NZY plate and incubated for 6 - 8 hours until plaques were just touching. Stored plate overnight at 4 ^{o}C to harden the top agar. Overlaid plate with a pre-labelled, nylon membrane disc for 2 minutes. Keyed filter to plate in an asymmetric pattern using a sterile needle. Made duplicate filter by placing another pre-labelled membrane disc onto the master plate for 4 minutes. Keyed duplicate to master plate in the same asymmetric pattern as the original filter. Placed filters plaque side up on Whatmann 3MM paper saturated with denaturing solution (1.5 M NaCl, 0.5 M NaOH) for 9 minutes. Transferred filters in a horizontal position to 3MM paper soaked in neutralising solution (1.5 M NaCl, 1 M Tris/HCl pH 7.2, 2 mM EDTA) twice for 5 minutes each. Finally placed filters on 3MM paper saturated in 2 x SSPE solution (3 M

NaCl, 0.2 M NaH_2PO_4, 20 mM EDTA pH 8) for 5 minutes. Between each transfer, blotted membranes on dry 3MM paper to remove excess fluid. Dried filters for 1 hour at room temperature and fixed DNA at 80 ^{o}C for 2 hours.

2.27.2 Hybridisation of Radiolabelled DNA Probe to Phage Library

Soaked filters in 2 x SET (20 x SET: 0.4 mM Tris/HCl pH 8, 3 M NaCl, 20 mM EDTA pH 8) for 5 minutes then pre-washed in 200 mls 5 x SET, 0.5% SDS, 1 mM EDTA pH 8 for 1 hour at 50 ^{o}C. Gently removed bacterial debris from membranes with mediwipes. Prehybridised filters in 50 mls prehybridisation solution (4 x SET, 5 x Denhardt's solution, 0.1% SDS, 0.1% tetra sodium pyrophosphate, 100 µg/ml denatured salmon sperm DNA) for 2 - 16 hours at the required temperature. Prepared radiolabelled probes (Section 2.21) and hybridised to filters for 24 hours in 25 mls freshly made prehybridisation solution at the desired temperature. Filters were incubated in sealed hybridisation tubes that rotated slowly in a Techne hybridisation oven (model HB - 1D). Washed filters as in Section 2.26.3. After washing, the filters were marked with radioactive ink so that the autoradiographs could be aligned with the master plates at a later stage. Autoradiographed filters after each wash step to check hybridisation pattern. Washed filters to higher stringencies if necessary.

2.27.3 Purification of Putative Positive Clones from the Phage Library

Plaques giving rise to putative positive hybridisation signals were identified as in Section 2.26.4 and recovered as a 1 cm^2 agar plug using a sterile scalpel blade. Added agar plug to 1 ml SM buffer plus 20 µl chloroform and vortexed briefly. Titred phage (Section 2.19) to determine the dilution resulting in approximately 2000 plaques per 15 cm plate. Repeated screening procedure as in Sections 2.27.1 and 2. After a total of three screening rounds, single plaques could be isolated and the clones purified.

2.27.4 Purification of Lambda Phage DNA

Promega's Wizard lambda DNA kits were used to purify lambda phage DNA. Approximately 5 µg DNA was obtained using the protocol recommended by the manufacturer. DNA was of a suitable quality for restriction analysis.

2.28 Preparation of Recombinant Vectors for DNA Sequencing

2.28.1 Preparation of Single Stranded M13 Template DNA

Transformed cells containing recombinant M13 vector were selected by α-complementation for purification of ss template for DNA sequencing.

Inoculated 100 mls 2 x TY medium with 1 ml of an overnight culture of *E. coli* TG1 of SURE^{TM}. Aliquoted 1.5 mls of these diluted cells into sterile Bijoux bottles and inoculated each with a white plaque using a sterile Pasteur pipette. Incubated cultures at 37 ^{o}C for 5.5 hours then transferred to clean microfuge tubes. Centrifuged tubes for 5 minutes and retained supernatant taking care not to pick up any cells. Re-centrifuged supernatant twice more to ensure all the *E. coli* cells had been removed. Precipitated phage by mixing supernatant with 200 µl PEG/NaCl (20% polyethylene glycol 6000, 2.5 M NaCl). Left mixture at room temperature for 15 minutes. Centrifuged tubes for 5 minutes and discarded supernatant. Repeated centrifugation for a further 2 minutes to remove all remaining traces of PEG. Dissolved the viral pellet in 100 µl TE buffer then extracted with 50 µl phenol pH 8 by vortexing for 20 seconds then allowing to stand at ambient temperature for 15 minutes. Vortexed tubes again for 20 seconds before centrifuging for 3 minutes. The upper, aqueous layer was retained and extracted with 1 ml chloroform. Added 7.5 µl 4 M sodium acetate pH 7.5 and 250 µl ethanol to the pooled aqueous phase then precipitated DNA for 1 hour at -20 ^{o}C. To harvest precipitate, centrifuged tubes for 10 minutes and dissolved DNA pellet in 30 µl TE buffer. Analysed DNA for recombinants and purity by

electrophoresis through an agarose/ethidium bromide gel. Stored template DNA at -20 ^{o}C.

2.28.2 Preparation of ds Template DNA for Sequencing

This technique was used to denature ds template DNA for primer annealing and subsequent DNA sequencing.

Used 3 - 5 µg of ds DNA for each sequencing reaction. If template was too dilute, DNA was concentrated by ethanol precipitation. Precipitated DNA was resuspended in no more than 8 µl TE. Added 2 µl 1 M NaOH/1 mM EDTA pH 8 to solution to denature DNA and incubated at room temperature for 5 minutes. Added 3 µl 3 M sodium acetate pH 6, 17 µl water plus 30 µl isopropanol and kept at room temperature for 5 minutes to precipitate denatured DNA. Centrifuged tube at 12,000 rpm for 15 minutes, washed pellet with 70% ice - cold ethanol and dissolved denatured DNA in 7 µl TE buffer. For primer annealing step refer to Section 2.29.1b.

2.29 Dideoxy DNA Sequencing

Cloned PCR products or restriction fragments were sequenced using Sanger's dideoxy DNA sequencing method (1980). Sequencing reactions were performed using Sequenase® Version 2.0 enzyme and other reagents supplied in a kit from United States Biochemical, Cleveland, Ohio, U.S.A..

2.29.1 Annealing Reaction

For each set of four sequencing reactions, a single annealing reaction was performed.

(a) Annealing Primer to ss M13 Template DNA

Mixed 1 µg ss M13 DNA (Section 2.28.1) with 1 µl (3 ng) primer and 2 µl reaction buffer (40 mM Tris/HCl pH 7.5, 20 mM $MgCl_2$, 50 mM NaCl). Increased reaction volume to 10 µl with deionised, distilled water. Warmed tube to 65 ^{o}C for 2

minutes then cooled slowly for 30 - 40 minutes until temperature was below 30 °C. Annealed template had to be used within 4 hours.

(b) Annealing Primer to Denatured ds Template DNA

3 - 5 µg denatured ds template DNA was resuspended in 7 µl TE buffer (Section 2.28.2). Added 2 µl reaction buffer and 1 µl primer to denatured DNA as before (Section 2.29.1a) and incubated at 37 °C for 15 - 30 minutes.

2.29.2 Labelling Reaction

To the annealed template - primer added 1 µl 0.1 M DTT, 2 µl dGTP diluted labelling mix (5 x concentrated dGTP labelling mix contains 7.5 µM dGTP, 7.5 µM dCTP, 7.5 µM dTTP), 0.5 µl (5 µCi) [α-^{35}S]dATP and 2 µl diluted Sequenase® Version 2.0 (diluted 1:8 with enzyme dilution mix and used within 1 hour). Mixed contents thoroughly but gently and incubated for 3 minutes at 18 °C.

2.29.3 Termination Reaction

Aliquoted 2.5 µl of the appropriate termination mix (80 µM of the four deoxynucleotides, 8 µM of one of the 4 dideoxynucleotides, 50 µM NaCl) into four separate tubes labelled G, A, T, C and incubated at 37 °C for at least 1 minute to warm the solutions. 3.5 µl of the completed labelling reactions were transferred to each of the four tubes containing the dideoxynucleotides and mixed. Continued the incubations for a further 5 minutes at 37 °C before stopping the reactions with 4 µl stop solution (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol FF). The reactions were kept on ice before loading onto the sequencing gel but could be stored at -20 °C for up to 1 week. Immediately before electrophoresis, the DNA was denatured by heating at 80 °C for 10 minutes.

2.29.4 DNA Sequencing Gel Electrophoresis

Prepared 50 mls of de-gassed sequencing gel (6% acrylamide/bis-acrylamide (40:1), 8 M urea, 1 x TBE). To polymerise gel, added 300 µl freshly made 10%

ammonium persulphate and 50 µl TEMED. Poured acrylamide mix between sealed glass plates (20 cm x 40 cm x 0.4 cm) and allowed to set for at least 1 hour at room temperature. The gel was pre-run for 30 minutes at a constant current of 35 mA. Rinsed wells with 1 x TBE to remove urea. Loaded 2 - 3 µl of each sequencing reaction onto gel and electrophoresed for 1.5 to 4 hours at 35 mA constant current. Fixed DNA and removed urea by soaking gel in 10% (v/v) acetic acid, 10% (v/v) methanol for 25 minutes. Transferred gel to Whatmann 3MM paper and dried on a slab gel drier at 80 ^{o}C for 1 hour 30 minutes. Autoradiographed gel overnight to visualise DNA sequencing ladder.

2.30 Computer Programmes for the Analysis of DNA Sequences

The Wisconsin Genetic Computer Group package was used for the analysis and comparison of the various DNA sequences determined in this project. The package allowed the comparison of DNA sequences to those in Genbank and EMBL (European Molecular Biology Laboratory, Heidelberg, Germany) databases.

Chapter 3

Amplification of *Anabaena variabilis* Genomic DNA by PCR

3.1 Introduction

The aim of this project was to clone the EPSP synthase gene, *aro*A, of *Anabaena variabilis* ATCC 29413. This would enable us to determine the amino acid sequence for comparison to other EPSP synthases. The polymerase chain reaction (PCR) was performed to amplify a fragment of the *aro*A gene which could be used as a homologous probe for screening a genomic DNA library to isolate the entire gene. This chapter describes the work carried out and assesses the viability of PCR with regard to this research.

3.2 The Polymerase Chain Reaction

PCR is an *in vitro* method for the enzymatic amplification of defined nucleic acid sequences (for review see Bej *et al.*, 1991 and Arnheim and Erlich, 1992). By exponentially amplifying a target sequence, PCR enhances the probability of isolating specific genes from complex DNA mixtures. The reaction involves denaturation of template DNA followed by annealing and extension of two oligonucleotide primers which bind to the complementary sequences flanking the target site. DNA synthesis is performed by the thermostable enzyme, *Taq* DNA polymerase, (from *Thermus aquaticus*) which can withstand the temperatures required for strand separation and primer annealing. This has enabled PCR to be performed in automated thermocycling machines. Genomic, plasmid or complementary DNA (cDNA) can act as template for the reaction. The success of PCR is dependent on the ability of oligonucleotide primers to bind specifically to their target site. When information regarding the target sequence is incomplete, degenerate primers can be designed from conserved regions of the protein sequence. The first gene to be cloned by PCR was the human β globin gene (Saiki *et al.*, 1985). Since then a large variety of genes have been cloned using this method. Some examples, of particular relevance to this work, include the *Haemophilus influenzae* and *Lactobacillus lactis* EPSP synthase genes (Maskell, 1993 and Griffin and Gasson, 1995 respectively), the *Anabaena variabilis* ATCC 29413 pyruvate :

ferredoxin oxidoreductase gene (Schmitz *et al.*, 1993) and the α and β allophycocyanin genes of *Synechocystis* sp. PCC 6803 (DiMagno and Haselkorn, 1993).

3.3 Design of Degenerate PCR Primers

When the target sequence is unknown, PCR primers can be designed from conserved regions of amino acid sequence (for review see Dieffenbach *et al.*, 1993). Since the genetic code is degenerate, it is impossible to predict the exact nucleotide sequence from an amino acid sequence. To accommodate for this, different bases (or degeneracies) can be introduced at the same position within the primer. Comparing protein sequences of functional significance can aid in primer design. Additionally, codon usage tables help in the selection of bases at the third codon position. The 3' end of the primer is critical for the success of PCR because this region is complementary to the target sequence. Therefore, the number of mismatches must be kept to a minimum. To reduce the number of degeneracies, it is desirable to include a conserved methionine or tryptophan in the primer design. Each of these amino acids are encoded by one codon. Serine, leucine and arginine should be avoided as each are encoded by one of six codons. In positions with three or four fold degeneracy, 2-deoxyinosine can be substituted because it can bind to all four bases. Deoxyinosine-containing primers have been successfully used to isolate genes (Knoth *et al.*, 1988 and Patil and Dekker, 1990). The 5' end of the primer usually contains bases not complementary to the target, such as the restriction site tail required for subsequent cloning of the PCR product. Some restriction enzymes do not cleave efficiently at the end of a DNA fragment. Therefore, one, two or three non-specific bases are added to the 5' end of the restriction site to overcome this problem. The position of conserved regions of amino acid sequence from which degenerate primers are designed dictate the length of the PCR product synthesised. PCR products between 150-bp and 1000-bp are generally synthesised for the amplification of gene sequences, although products over 3-kb have been produced.

3.4. Strategy for PCR Amplification of *A. variabilis* Genomic DNA

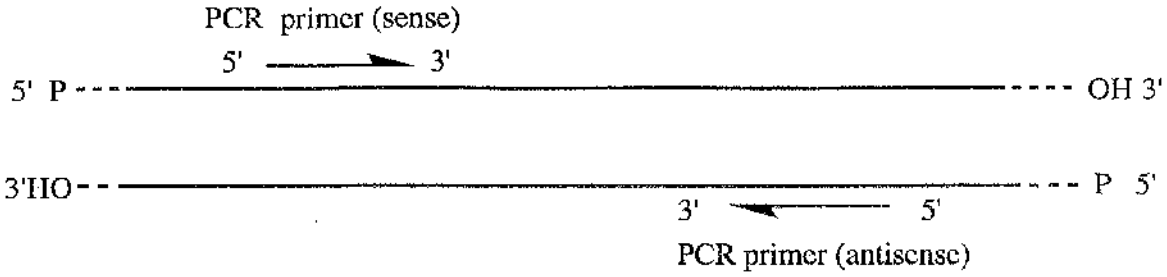
To isolate an internal fragment of the *A. variabilis aro*A gene, PCR was performed employing degenerate primers and *A. variabilis* genomic DNA. Figure 3.1 outlines the procedure used for PCR amplification, cloning and dideoxy DNA sequencing. *A. variabilis* was grown and the genomic DNA extracted by Dr. H.A. Powell and Dr. N.W. Kerby (Dundee University, Scotland).

3.4.1 Design of Degenerate *aro*A Primers

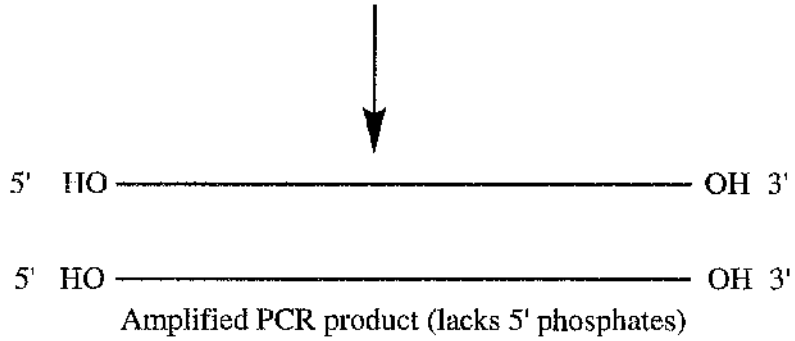
Sequence alignments have shown regions of conservation between EPSP synthase sequences from plants, bacteria, fungi and yeast. Two such regions were used to design degenerate oligonucleotide primers as illustrated in Figure 3.2 and Table 3.1. This strategy works on the assumption that the homology between different EPSP synthase sequences is shared by the *A. variabilis* enzyme. Therefore, primer 1 was designed from the putative glyphosate binding region (Padgette *et al.*, 1991) which is very well conserved (see Section 3.8.3 for further discussion). Primer 2 was designed from a region better conserved in Gram negative bacteria and higher plants than in fungi and yeast. The homology in this region was considered more likely to be shared with cyanobacteria since these organisms are photosynthetic prokaryotes. Since the codon preferences of *Anabaena variabilis* ATCC 29413 were unknown at the time this work was carried out, deoxyinosine residues were inserted at positions of degeneracy. Synthesis of oligonucleotides was performed as in Section 2.13.

3.4.2 PCR Amplification of *A. variabilis* Genomic DNA Using Primers 1 and 2

To obtain optimal conditions for the synthesis of a PCR product of the desired size, a number of different parameters were tested. This involved varying the amounts of genomic DNA from 10 to 1000 ng and primers from 40 to 100 pmoles per 100 µl



Exponential amplification of desired target with cycles of
DNA denaturation, primer annealing and DNA synthesis with *Taq* DNA ploymerase



Amplified PCR product (lacks 5' phosphates)

M13mp19RF
or pBluescript SK-

Restriction digest amplified
PCR product and cloning vector

Ligation of PCR product into vector

Transformation of ligation products into suitable host.

Selection and amplification of recombinant vectors.

Preparation of recombinant vectors for dideoxy DNA sequencing

Figure 3.1 Strategy for PCR amplification of genomic DNA and sequencing the synthesied products.

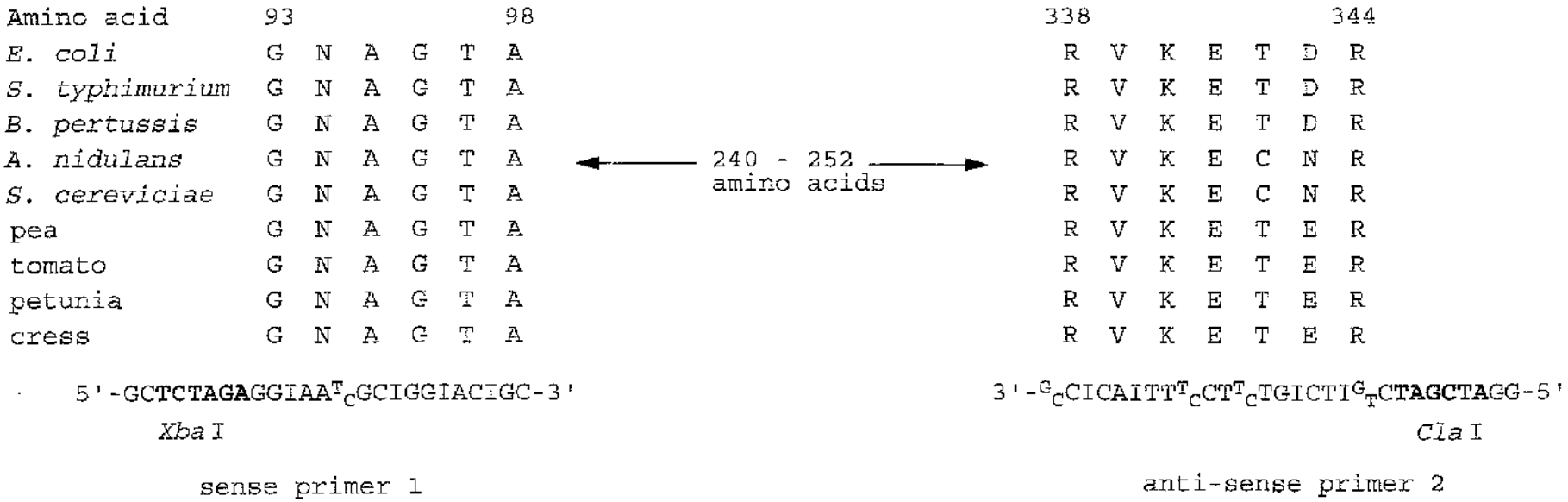


Figure 3.2 Design of degenerate primers 1 and 2 for the amplification of a segment of the *Anabaena variabilis aro* A gene.

Multiple sequence alignments of two conserved regions of EPSP synthase sequence from a variety of sources are shown. Degenerate primers are displayed beneath the sequence alignments. Restriction sites are shown in bold type. Two base extensions exist 5' to the restriction sites. The amino acid positions according to the *E. coli* EPSP synthase sequence are indicated.

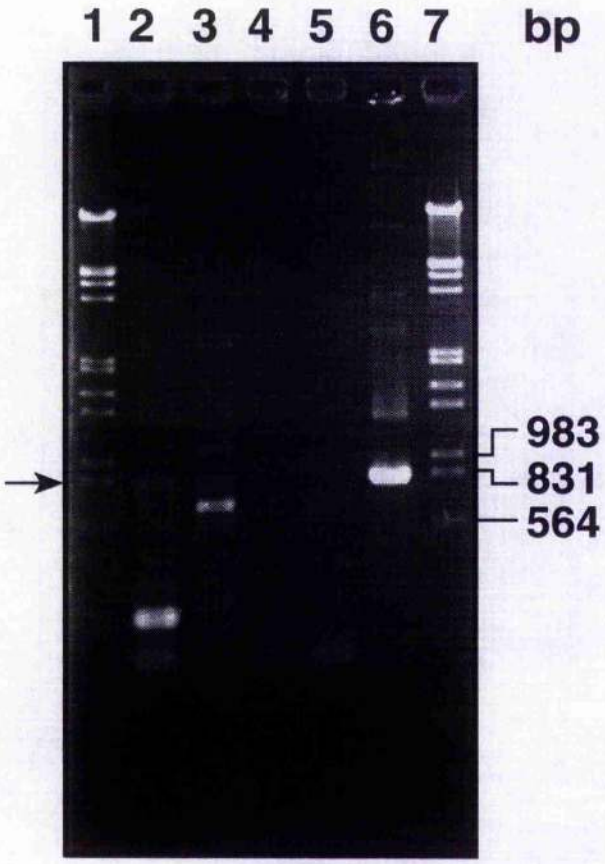
reaction. $MgCl_2$ concentrations between 1.5 and 5 mM were tested and the annealing temperature ranged between 42 - 50 °C. The reaction conditions described below yielded a PCR product of the size of interest.

PCR of 1 µg *A. variabilis* genomic DNA was performed as described in Section 2.14.1 in the presence of 1.5 mM $MgCl_2$, 200 µM of each dNTP, 40 pmoles primer 1 and 100 pmoles primer 2. More primer 2 was used to increase the probability of hybridisation to its target site since this primer was designed from a less well conserved region (Figure 3.2). The PCR parameters were as follows: denaturation step at 94 °C for 3 minutes followed by 30 cycles at 94 °C for 1 minute, 45 °C for 2 minutes and 72 °C for 2 minutes. An extension step at 72 °C for 5 minutes completed the reaction. Analysis of the reaction products by agarose gel electrophoresis (Section 2.14.2) revealed that synthesis of a DNA fragment of the predicted size, ~800-bp, had occurred (Lane 2, Figure 3.3).

Positive controls were performed in parallel with the above experiments using the clone, pSG5161, as template for the reaction. pSG5161 is a pUC18 derivative containing the pea *aro*A cDNA sequence on a 1.5-kb *Eco*RI fragment (Granger, 1989). A PCR product of the expected size (809-bp) was synthesised (Lane 6, Figure 3.3). A negative control confirmed that the products formed were not contaminants from PCR reagents or from automatic pipettes used to set up the experiment (Lane 5, Figure 3.3). Single primer controls performed with *A. variabilis* genomic DNA showed that primer 1 alone synthesised a fragment ~800-bp in length (Lane 3, Figure 3.3). This fragment appeared to be slightly smaller in size than the PCR product of interest. So the PCR reaction performed with both primers 1 and 2 may have synthesised different products from both the mixed primer pair as well as primer 1 alone.

The 800-bp DNA fragment (Lane 2, Figure 3.3) was then purified from a low melting point agarose / ethidium bromide gel (Section 2.12.3) and cloned into the phage vector, M13mp19RF. Dideoxy DNA sequencing was carried out to screen for the *aro*A fragment.

Figure 3.3 PCR of *Anabaena variabilis* genomic DNA with degenerate primers 1 and 2. A 10% volume of each reaction product was separated through a 1% agarose gel then stained with ethidium bromide. Lanes 1 and 7 are DNA markers: lambda phage DNA cut with *Hin*dIII and *Eco*RI. Lanes 2 - 4 are PCR products synthesised from 1 µg *A. variabilis* DNA. Lane: (2) primers 1 and 2, (3) primer 1 alone, (4) primer 2 alone. Lane 5 is a negative control with no template DNA and lane 6 is a positive control with 100 ng pSG5161 (a pea *aro*A cDNA clone). Both control experiments were performed with primers 1 and 2. See text for reaction conditions. In lane 2 the amplified DNA fragment of interest is ~800-bp in size and is indicated by an arrow. A band of similar size has been synthesised by primer 1 alone. Smears at the bottom of each lane are unincorporated primers.


1 2 3 4 5 6 7
bp
983
831
564

3.4.3 Cloning PCR Amplified Products

The 800-bp PCR product was digested with the restriction enzymes *Xba*I and *Cla*I (Section 2.9.4) and purified (Section 2.12.2). Digestion using the aforementioned enzymes would cut the restriction ends of the DNA insert whether it was synthesised from the mixed primer pair and / or primer 1 alone (see Section 3.4.2). Electrophoresis of the digested PCR product through an agarose gel (Section 2.11) showed that neither *Xba*I nor *Cla*I obviously cut within the insert DNA. The PCR product was then ligated into *Xba*I / *Acc*I cut M13mp19RF (Section 2.14.3) to selectively clone DNA fragments synthesised by the mixed primer pair (*Acc*I restriction ends are complementary to *Cla*I). Ligation products were then transformed into *E. coli* TG1 (Table 2.1 and Sections 2.171a/2/3a). Transformants putatively containing recombinant phage were selected by α-complementation (Section 2.17.4). Single stranded phage DNA was purified from these cells (Section 2.28.1) and analysed by agarose gel electrophoresis (Section 2.11). Twelve recombinant clones were recovered from directional cloning of the PCR product into *Xba*I, *Acc*I cut M13mp19RF. However, each of the purified recombinant clones were of a different size. This suggested that the DNA had been randomly cleaved.

All 12 recombinants were, nevertheless, partially sequenced on one strand using the universal primer by the dideoxy DNA chain termination method (Sanger *et al.*, 1980) as in Section 2.29. The amino acid sequence of each clone was subsequently deduced in six reading frames and compared to the appropriate regions of EPSP synthase sequences from various bacteria, fungi, yeast and plants. The deduced amino acid sequence of the insert from clone pAV35 showed considerable homology to other EPSP synthase sequences near the primer 1 target site. However, the amino acid sequences deduced from the other cloned PCR products did not show any homology to the published EPSP synthase sequences, so were not further analysed. The double stranded replicative form of pAV35 was purified (Section 2.6.1) and the residual DNA sequence determined using the reverse primer for M13. The insert sequence near the reverse primer had a high %GC content and so dITP was substituted for dGTP during the sequencing reaction. dITP minimises the occurrence of secondary structure which

would otherwise cause DNA fragments to run faster through the polyacrylamide sequencing gel. This would cause the spacing between the DNA bands in the gel to compress, so making the sequence difficult to read. After sequencing the insert DNA of pAV35, the amino acid sequence was deduced and aligned with other EPSP synthases. This multiple sequence alignment showed the insert from pAV35 shared considerable homology to EPSP synthases from other species (as illustrated in Figure 3.4).

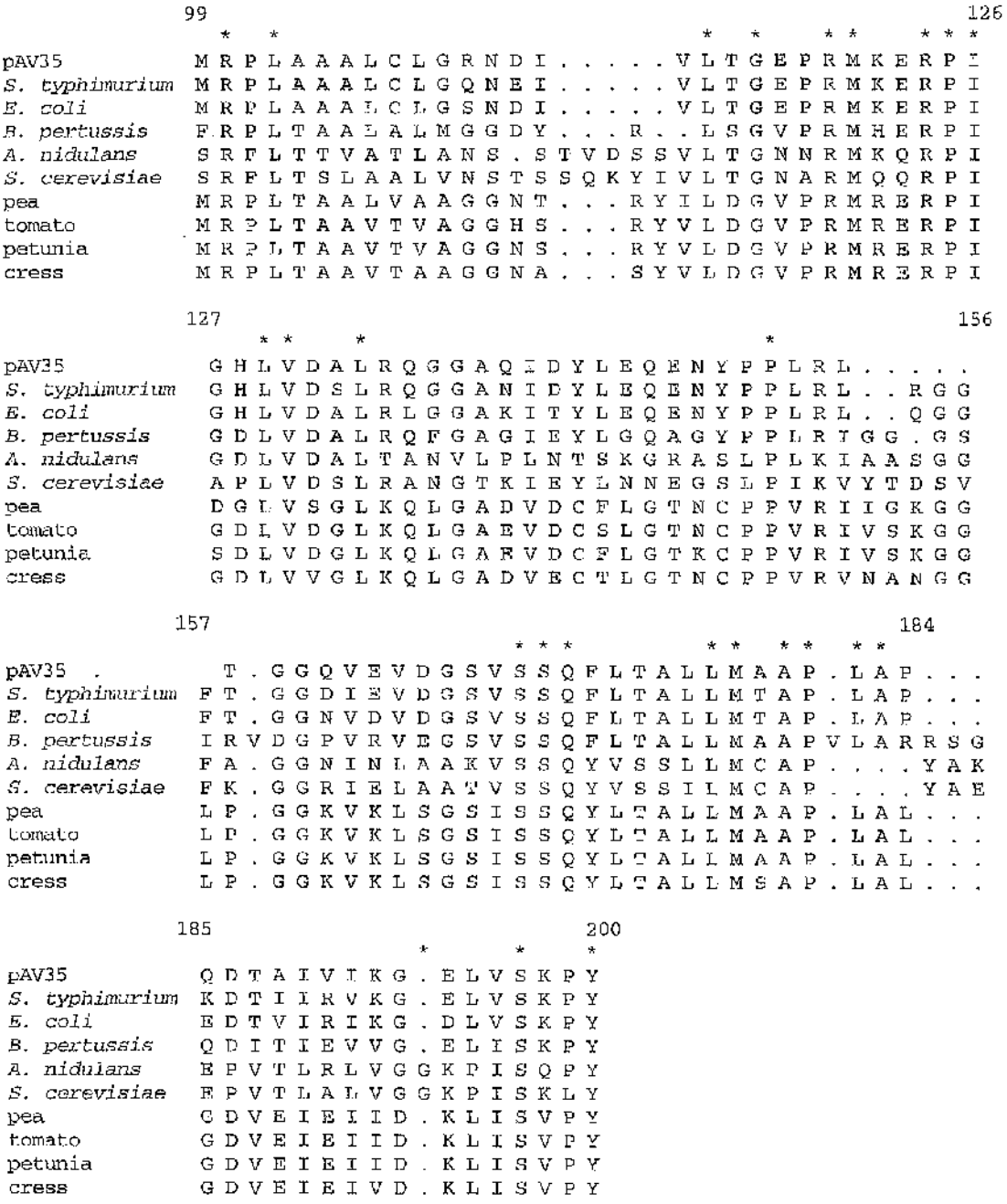
Unfortunately, a large portion of the 3' end of the PCR product of pAV35 had been deleted reducing the size of the insert from ~800-bp to 320-bp. DNA sequencing showed that the vector sequence was unaltered by the deletion event indicating that cleavage was exclusive to the cloned fragment. This accounted for the various sizes of inserts in the recombinants initially isolated. Attempts were then made to clone the full length PCR product.

3.4.4 Attempts to Clone the Full Length, PCR Amplified *aroA* Fragment

Cloning the 800-bp PCR product into M13mp19RF resulted in cleavage of the insert to a fragment 320-bp in size. As discussed in Section 3.4.3, the ~800-bp PCR product did not contain internal *Xba*I or *Cla*I sites. Agarose gel electrophoresis of the ligation products formed between the PCR product and M13mp19RF also indicated that insert DNA had not been altered prior to transformation into *E. coli* TG1. The reduction in insert size may have been due to secondary structure formed by homologous regions within the insert which were then deleted by the host's DNA recombination systems. To minimise the possibility of this occurring, *E. coli* SURE™ was employed instead of *E. coli* TG1 as the host for transformation of the ligation products. The key features of *E. coli* SURE™ are a series of mutations that significantly reduce the frequency of rearrangements and deletions of unstable DNA molecules (Greener, 1990 and Table 2.1).

Figure 3.4. EPSP synthase sequence alignment.

Deduced amino acid sequence of the putative *A. variabilis* EPSP synthase gene fragment (pAV35) is aligned with appropriate segments of *Salmonella typhimurium, E. coli, Bordatella pertussis, Aspergillus nidulans, Saccharomyces cerevisiae*, pea, tomato, petunia and cress EPSP synthase sequences. The identity of the putative *A. variabilis* EPSP synthase to the enzyme sequences of *E. coli, A. nidulans* and pea are 89.2%, 40.2% and 55.8% respectively. An asterisk represents complete identity amongst all the EPSP synthase sequences. Dots are used to optimise alignments. Amino acid positions of the *E. coli* EPSP synthase sequence are indicated.


99
126
pAV35 M R P L A A A L C L G R N D I . . . . . V L T G E P R M K E R P I
S. typhimurium M R P L A A A L C L G Q N E I . . . . . V L T G E P R M K E R P I
E. coli M R P L A A A L C L G S N D I . . . . . V L T G E P R M K E R P I
B. pertussis F R P L T A A L A L M G G D Y . . . R . . L S G V P R M H E R P I
A. nidulans S R F L T T V A T L A N S . S T V D S S V L T G N N R M K Q R P I
S. cerevisiae S R F L T S L A A L V N S T S S Q K Y I V L T G N A R M Q Q R P I
pea M R P L T A A L V A A G G N T . . . R Y I L D G V P R M R E R P I
tomato M R P L T A A V T V A G G H S . . . R Y V L D G V P R M R E R P I
petunia M R P L T A A V T V A G G N S . . . R Y V L D G V P R M R E R P I
cress M R P L T A A V T A A G G N A . . . S Y V L D G V P R M R E R P I
127
156
pAV35 G H L V D A L R Q G G A Q I D Y L E Q E N Y P P L R L . . . . .
S. typhimurium G H L V D S L R Q G G A N I D Y L E Q E N Y P P L R L . . R G G
E. coli G H L V D A L R L G G A K I T Y L E Q E N Y P P L R L . . Q G G
B. pertussis G D L V D A L R Q F G A G I E Y L G Q A G Y P P L R T G G . G S
A. nidulans G D L V D A L T A N V L P L N T S K G R A S L P L K I A A S G G
S. cerevisiae A P L V D S L R A N G T K I E Y L N N E G S L P I K V Y T D S V
pea D G L V S G L K Q L G A D V D C F L G T N C P P V R I I G K G G
tomato G D L V D G L K Q L G A E V D C S L G T N C P P V R I V S K G G
petunia S D L V D G L K Q L G A R V D C F L G T K C P P V R I V S K G G
cress G D L V V G L K Q L G A D V E C T L G T N C P P V R V N A N G G
157
184
pAV35 . T . G G Q V E V D G S V S S Q F L T A L L M A A P . L A P . . .
S. typhimurium F T . G G D I E V D G S V S S Q F L T A L L M T A P . L A P . . .
E. coli F T . G G N V D V D G S V S S Q F L T A L L M T A P . L A P . . .
B. pertussis I R V D G P V R V E G S V S S Q F L T A L L M A A P V L A R R S G
A. nidulans F A . G G N I N L A A K V S S Q Y V S S L L M C A P . . . . Y A K
S. cerevisiae F K . G G R I E L A A T V S S Q Y V S S I L M C A P . . . . Y A E
pea L P . G G K V K L S G S I S S Q Y L T A L L M A A P . L A L . . .
tomato L P . G G K V K L S G S I S S Q Y L T A L L M A A P . L A L . . .
petunia L P . G G K V K L S G S I S S Q Y L T A L L M A A P . L A L . . .
cress L P . G G K V K L S G S I S S Q Y L T A L L M S A P . L A L . . .
185
200
pAV35 Q D T A I V T K G . E L V S K P Y
S. typhimurium K D T I I R V K G . E L V S K P Y
E. coli E D T V I R I K G . D L V S K P Y
B. pertussis Q D I T I E V V G . E L I S K P Y
A. nidulans E P V T L R L V G G K P I S Q P Y
S. cerevisiae E P V T L A L V G G K P I S K L Y
pea G D V E I E I I D . K L I S V P Y
tomato G D V E I E I I D . K L I S V P Y
petunia G D V E I E I I D . K L I S V P Y
cress G D V E I E I V D . K L I S V P Y

The PCR product synthesised by the mixed primer pair PCR reaction (Lane 2, Figure 3.3) was cloned into M13mp19RF as described before (Sections 2.14.3 and 3.4.3) except using *E. coli* SURE^{TM} as host. Single and double stranded templates were prepared (Sections 2.6.1 and 2.28) and DNA sequenced (Section 2.29) using the universal and reverse primers respectively. The deduced amino acid sequence of five clones, pAV51 - 55, matched pAV35 exactly. So the same *aro*A-like fragment had been isolated. Unfortunately, the deletion event that took place in pAV35 re-occurred in pAV51 - 55.

Although attempts to clone the full length PCR product were unsuccessful, the 320-bp cloned insert was used to probe Southern blots of restriction digested *A. variabilis* genomic DNA. This would allow detection of DNA fragments carrying homologous sequences. The cloned insert was regarded as reliable for these experiments for a number of reasons. The deduced amino acid sequence of the insert had a strong homology to EPSP synthases from other organisms (Figure 3.4). Importantly however, this deduced amino acid sequence was sufficiently different from other EPSP synthases to suggest that the cloned insert originated from *A. variabilis* genomic DNA. A number of clones carrying the 320-bp insert containing the putative *A. variabilis aro*A gene fragment were sequenced. The resulting information showed that each clone was identical which indicated that no rearrangements had occurred within the cloned DNA.

3.5 Southern Blots of *Anabaena variabilis* Genomic DNA

In order to obtain the entire sequence of the putative *A. variabilis aro*A gene, a library of genomic DNA from this cyanobacterium had to be constructed. A Southern blot of restriction digested *A. variabilis* genomic DNA was initially performed to determine the size and number of fragments to which the putative *A. variabilis aro*A fragment hybridised.

A new stock of *A. variabilis* genomic DNA was used for these experiments and prepared as in Sections 2.3, 2.7 and 4.4.1. The 320-bp *aro*A-like fragment was

purified from a low melting point agarose / ethidium bromide gel (Section 2.12.3) and radiolabelled with [α-^{32}P]dCTP (Sections 2.21 and 2.22) to a specific activity of 1.6 x 10^9 cpm/μg DNA. The probe was hybridised to a Southern blot of 7 μg *Hin*dIII digested *A. variabilis* genomic DNA (Section 2.23) at 60 ^{o}C (Section 2.24). After the filters were washed from 2 x SET to 1 x SET at 60 ^{o}C to remove non-specifically bound probe (Section 2.24) and exposed to X-ray film for 24 hours at -80 ^{o}C, the hybridisation signals which resulted were not very strong in their intensity (Figure 3.5).

Since the probe was labelled to a high specific activity and a large amount of genomic DNA had been digested and bound to the filter and the filters were washed under non-stringent conditions and there was a long exposure time to X-ray film, the hybridisation signal should have been much stronger if the probe was binding to homologous DNA sequences. This experiment then raised considerable fears that the cloned *aro*A-like fragment had been synthesised not from *A. variabilis* genomic DNA but from contaminating DNA that may have been present in the original DNA solution used for the PCR experiments (Section 3.4). The contamination problem could not have originated from the *Hin*dIII digested DNA used for the Southern blot experiment. This was because this genomic DNA was extracted from a culture of *A. variabilis* cells known to be free from contamination (refer to Section 2.3).

To determine the strength of signal that would result from a probe binding to homologous prokaryotic genomic DNA, a control Southern blot was performed. Genomic DNA from the photosynthetic bacterium *Rhodobacter sphaeroides* and its light harvesting complex gene, LHCII, were the kind gifts of Professor R. Cogdell (Glasgow University, Scotland). The LHCII gene was radiolabelled to a specific activity of 3.6 x 10^9 cpm/μg and hybridised to 7 μg of *Eco*RI and *Sal*I digested *Rb. sphaeroides* DNA at 60 ^{o}C as described above. After washing the filters from 2 x SET to 1 x SET at 60 ^{o}C, strong hybridisation signals of the expected size were observed after 3.5 hours exposure to X-ray film at -80 ^{o}C (Figure 3.6).

Figure 3.5 Southern hybridisation of the 320-bp putative *aro*A gene fragment to 7 μg *A. variabilis Hin*dIII digested genomic DNA. Genomic DNA was digested with *Hin*dIII for 1 hr (lane 1), 3 hrs (lane 2), 5 hrs (lane 3) and 16 hours (lane 4). Hybridisation and washing conditions are outlined in the text. It should be noted that restriction digestion of the genomic DNA was not performed exactly as described in Section 2.9.1. The digestion reactions were not mixed thoroughly prior to incubation. Since mixing allows the genomic DNA to become more evenly distibuted throughout the reaction volume, complete digestion of *A. variabilis* DNA by *Hin*dIII at each restriction site could not be guaranteed. The hybridisation pattern may then reflect binding of the 320-bp probe to partially digested *A. variabilis* genomic DNA.

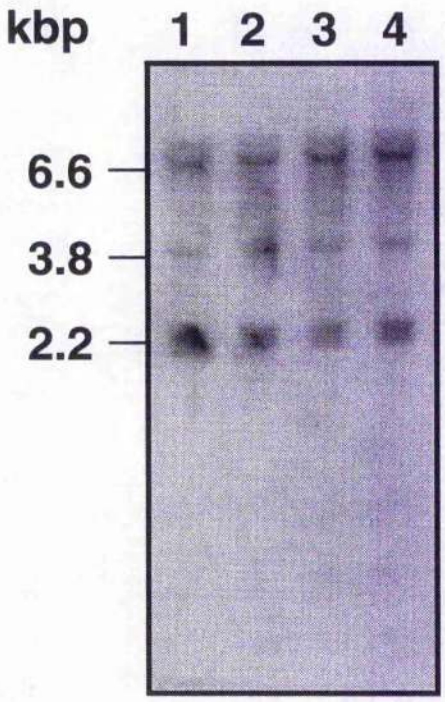
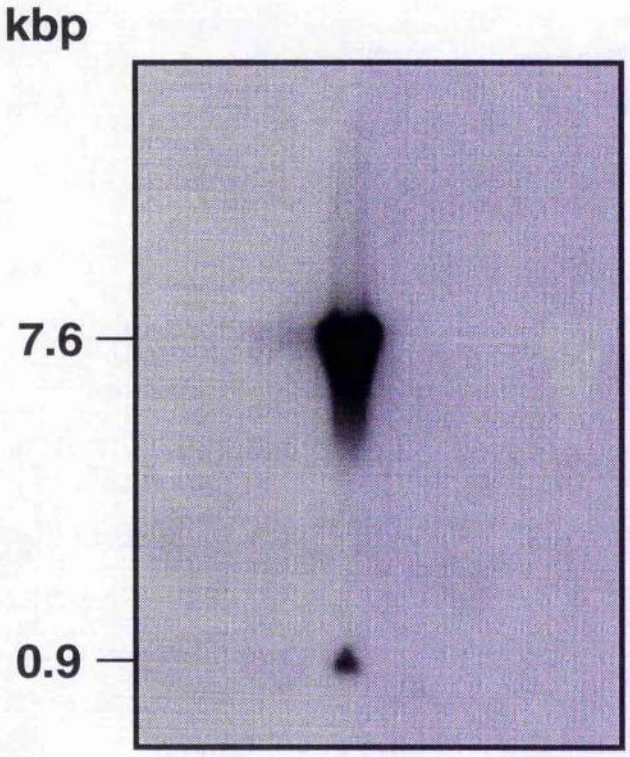

kbp
1
2
3
4
6.6
3.8
2.2

Figure 3.6 Southern hybridisation of *Rhodobacter sphaeroides* LCIII1 gene to *Rb. sphaeroides* genomic DNA. 7 μg genomic DNA was digested with *Eco*RI and *Sal*I (Section 2.9.4). Hybridisation and washing conditions are outlined in the text.


kbp
7.6
0.9

Comparison of the results obtained from the *A. variabilis* and *Rb. sphaeroides* Southern blots supported the hypothesis that the cloned *aro*A-like DNA fragment was not synthesised from *A. variabilis* DNA but from contaminating DNA.

3.6 New Strategy for PCR Amplification of an *Anabaena variabilis aro*A Gene Fragment.

Fresh cultures of *A. variabilis* were grown and checked for contamination (Section 2.3.1/2). Genomic DNA was extracted from these axenic cultures of *A. variabilis* cells as described in Section 2.7. This uncontaminated genomic DNA was used as a template for newly designed degenerate primers which were the kind gift of Drs. M. dalla Chiesa and D. Maskell (Imperial College of Science, Technology and Medicine, London, England). The new primers (numbered 3 and 4) were based on conserved regions of the EPSP synthase sequences of the Gram positive bacteria, *Bacillus subtilis* (Henner *et al.*, 1986) and *Staphylococcus aureus* (O'Connell *et al.*, 1993) (Figure 3.7 and Table 3.1). Primer 3 (as with primer 1) was designed from the putative glyphosate binding region (Padgette *et al.*, 1991; see Section 3.8.3 for further discussion). The uncontaminated genomic DNA was also used as a template for primers 1 and 2 as well as with mixed primer sets (e.g.. 1 and 4 ; 3 and 2) because their annealing temperatures and %GC content were compatible.

3.6.1 PCR Amplification of Uncontaminated *Anabaena variabilis* Genomic DNA with Degenerate Primers

The PCR was carried out with *A. variabilis* genomic DNA and all the primer sets as described in Section 2.14.1. The reaction conditions were varied to obtain the optimal conditions for synthesis of the desired product. This included using intact or digested template DNA; varying the amount of template DNA from 0.1 - 1 µg; modifying the denaturation time (between 3 and 7 minutes), changing the annealing temperature (between 45 - 55 ^{o}C) and time (between 1 and 2 minutes); altering the

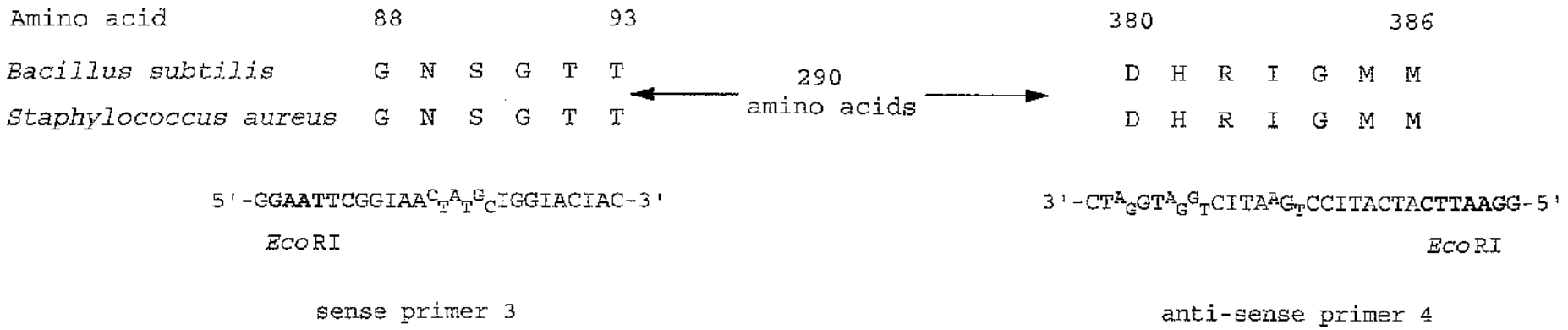


Figure 3.7 The degenerate primers 3 and 4 used for the amplification of an *Anabaena variabilis aro* A gene segment (courtesy of Drs. dalla Chiesa and D. Maskell, Imperial College, London). Multiple sequence alignments of two conserved regions of EPSP synthase sequence from *Bacillus subtilis* and *Staphylococcus aureus* are shown. Degenerate primers are displayed beneath the alignments. Restriction sites are shown in bold type. One base extensions exist 5' to the restriction sites. The amino acid positions according to the *B. subtilis* EPSP synthase sequence are indicated.

Table 3.1 PCR primers for amplification of *Anabaena variabilis aroA* gene fragment.

Sense primer 1 (25 mer, 512 fold, GC content = 72%, T_a = 57 °C)	5'-GC**TCTAGA**GGIAA(CT)GCIGGIACIGC-3' *Xba*I
Antisense primer 2 (28 mer, 4096 fold, GC content = 60%, T_a = 53 °C)	5'-GG**ATCGAT**C(GT)ITCIGT(TC)TC(TC)TTIACIC(GT)-3' *Cla*I
Sense primer 3 (24 mer, 2048 fold, GC content = 59%, T_a = 49 °C)	5'-**GGAATTC**GGIAA(CT)(AT)(GC)IGGIACIAC-3' *Eco*RI
Antisense primer 4 (27mer, 384 fold, GC content = 55%, T_a = 57 °C)	5'-**GGAATTC**ATCATNCC(AGT)ATNC(GT)(AG)TG(AG)TC-3' *Eco*RI

Primers are illustrated in their sense or antisense orientation. The lengths, %GC content and their estimated degeneracies (fold) and annealing temperatures (T_a) are shown. The degeneracies are calculated assuming 2-deoxyinosine can bind to all four bases equally well. T_a of the sequence specific part of the primer is determined with the %GC content indicated using the equation in Bej *et al.* (1991). Restriction sites are shown in bold type with a one or two base overhangs at the 5' end. Brackets indicate a mixture of bases at one base position. I = 2-deoxyinosine.

$MgCl_2$ concentration between 1.5 and 4 mM, the amount of oligonucleotide between 40 and 100 pmoles and the cycle number between 30 and 40.

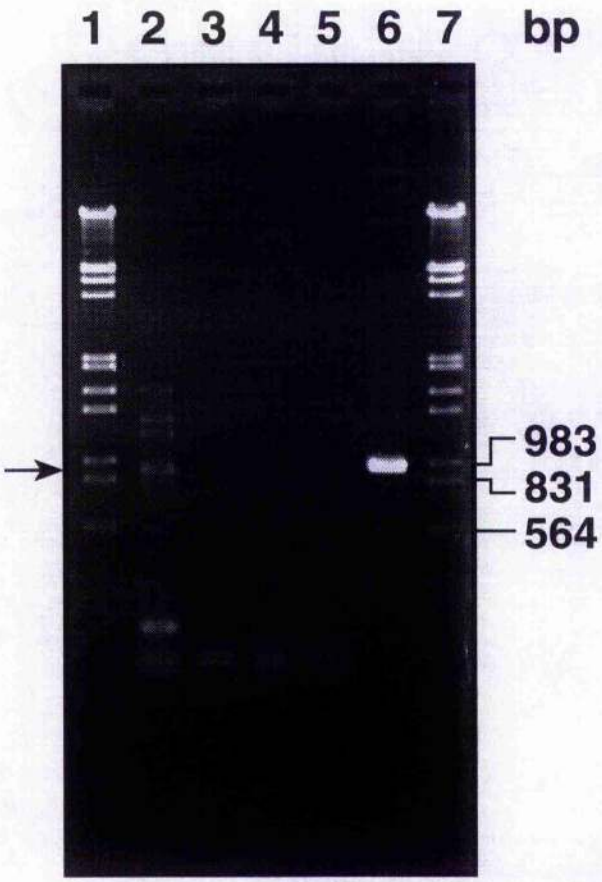
Negative results were obtained with all the primer sets under all the experimental conditions tested except with primers 3 and 4 (Figure 3.8). Primers 3 and 4 yielded a product of the desired size (~900-bp) with 1 μg intact genomic DNA, 3 mM $MgCl_2$, 100 μM dNTP and 100 pmoles oligonucleotides. The PCR parameters were as follows: denatured DNA at 94 °C for 5 minutes followed by 30 cycles at 94 °C for 1 minute, 50 °C for 1 minute and 72 °C for 1.5 minutes. After 30 cycles samples were incubated at 72 °C for 5 minutes to complete the reaction.

A positive control experiment was performed using pSM01, an *aro*A clone from the unicellular cyanobacterium *Synechocystis* 6803 (Mayes *et al.*, 1993). This showed that primers 3 and 4 synthesised a product of the desired size (927-bp). A negative control confirmed the purity of the reagents used. Single primer controls showed the product synthesised resulted from the mixed primer pair, 3 and 4, and not from one primer alone (Figure 3.8).

3.6.2. Cloning and DNA Sequencing of the 900-bp PCR Product Amplified with Primers 3 and 4

The 900-bp PCR product synthesised with primers 3 and 4 was purified from a low melting point agarose / ethidium bromide gel (Section 2.12.3) and digested with *Eco*RI (Section 2.9.3) prior to cloning into the phagemid vector, pBluescript SK- (Section 2.14.3). Transformation of the ligated DNA into *E. coli* DH5α (Section 2.17.1b/2/3b) did not, however, yield any recombinant clones. Analysis of the ligation products by agarose gel electrophoresis (Section 2.11) indicated that only one end of the PCR product had ligated into pBluescript SK- forming a linearised DNA molecule. Since linear DNA is transformed into *E. coli* at a much reduced efficiency in comparison to circular plasmid DNA (Sambrook *et al.*, 1989), no recombinant clones were isolated. It was assumed that the restriction site of one of the primers used to synthesise the 900-bp PCR product could not be digested by *Eco*RI and had remained

Figure 3.8 PCR of uncontaminated *Anabaena variabilis* genomic DNA with degenerate primers 3 and 4. A 10% volume of each reaction product was separated through a 1% agarose gel then stained with ethidium bromide. Lanes 1 and 7 are DNA markers: lambda DNA cut with *Hin*dIII and *Eco*RI. Lanes 2 - 4 are PCR products generated from 1 µg *A. variabilis* genomic DNA and primers 3 and 4 (lane 2), primer 3 alone (lane 3) and primer 4 alone (lane 4). Lane 5 is a negative control with no template DNA and lane 6 is a positive control with pSM01 (a *Synechocystis aro*A clone). Both control experiments were performed with primers 3 and 4. See text for reaction conditions. In lane 2 the desired fragment (~900-bp in size) is indicated by an arrow. Smears at the bottom of each lane are unincorporated primers.


1 2 3 4 5 6 7 bp
983
831
564

blunt ended. Therefore, the PCR product was cloned into pBluescript SK- cut with *Eco*RI and *Sma*I to accommodate the *Eco*RI restriction site tail and blunt end of the insert. Ten recombinant clones were isolated after these ligation products were transformed into *E. coli* DH5α (Section 2.5.1).

A high quality stock of each recombinant clone was prepared (Section 2.6.2) and approximately 100-bp of the insert DNA was sequenced (Section 2.29) using the KS and SK primers specific for pBluescript SK-. Of the ten recombinants isolated, four different sequences were determined. So the PCR primers had not specifically hybridised to one target sequence in the *A. variabilis* genomic DNA. The partial sequence of each insert was translated in six reading frames and compared to other translated *aro*A sequences available in the genembl data base. This was accomplished using the, "bestfit", command of the Wisconsin Genetic Computer Group package (Section 2.30). Not one of the cloned PCR products shared any homology to the available EPSP synthase sequences. A general homology search was then performed using the, "tfasta", command to determine if the translated insert sequences had any significant homology to other sequences logged in the genembl data-base (Section 2.30). Interpretation of these results proved very difficult because of the limited sequence information available with respect to the cloned PCR products. This problem could have been overcome by sequencing more of the insert DNA, but since the principal aim of this project was to clone the *A. variabilis aro*A gene, this work was put to one side.

3.7 Southern Blot of *Anabaena variabilis* Genomic DNA Using 900-bp PCR Product as Probe

PCR of genomic DNA using degenerate primers can yield a product of the desired size that contains a mixture of different DNA fragments (as discussed in Section 3.6.2). To determined if the 900-bp PCR product synthesised with primers 3 and 4 contained an *aro*A-like sequence, it was used as a probe for hybridisation to different restriction digestions of *Anabaena variabilis* genomic DNA. The hybridisation pattern

was then compared to *A. variabilis* DNA digested with the same enzymes and probed with the 618-bp *Synechocystis* sp. 6803 *aro*A gene fragment. As discussed in Section 4.3.2/3, the *Synechocystis aro*A probe binds specifically to *A. variabilis* genomic DNA. The signal obtained from hybridisation of the 900-bp PCR product to the cyanobacterial genomic DNA would differ from that of the *Synechocystis aro*A probe. This is because DNA sequencing had revealed that the 900-bp PCR product contained a family of different fragments. However, if the signal from the PCR product contained a pattern that could be matched to the *Synechocystis aro*A signal then at least one *aro*A-like fragment may have been synthesised and further DNA sequencing would be required.

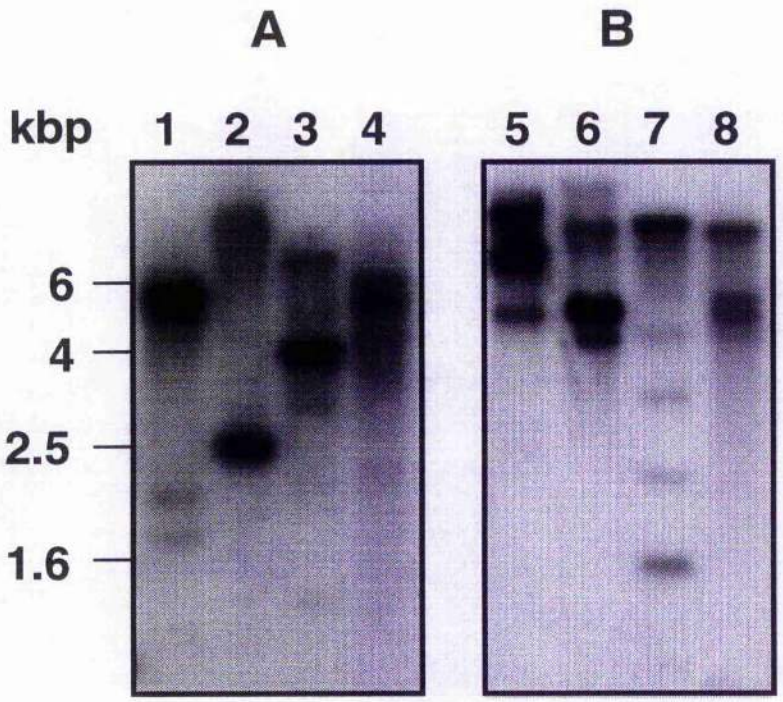
A Southern blot of *A. variabilis* genomic DNA, digested with the restriction enzymes indicated in Figure 3.9, was prepared.(Section 2.23). The 900-bp PCR product and *Synechocystis aro*A probe were radiolabelled (Sections 2.21 and 2.22) and hybridised to the genomic DNA at 65 ^{o}C and 50 ^{o}C respectively (Section 2.24). Each membrane was washed at the appropriate temperature to remove non-specifically bound probe prior to autoradiography for 16 hours at -80 ^{o}C. The 900-bp probe could bind to *A. variabilis* DNA under stringent hybridisation conditions so was probably synthesised from the genomic DNA of this cyanobacterium (Figure 3.9, B). However, distinct differences were seen on comparison of the resulting hybridisation signals from each probe (Figure 3.9, A and B). Therefore, it seemed very likely that the PCR product amplified by primers 3 and 4 did not contain the *A. variabilis aro*A gene fragment.

3.8 Discussion

3.8.1 PCR and Contamination

The first approach to isolate the *Anabaena variabilis aro*A gene by PCR using degenerate primers 1 and 2 was unsuccessful. The PCR experiment amplified a

Figure 3.9 Southern hybridisation of the (A) *Synechocystis aro*A probe and (B) 900-bp PCR product to 2 μg *Anabaena variabilis* genomic DNA. Genomic DNA was digested with *Eco*RV (lanes 1 and 5), *Eco*RI (lanes 2 and 6), *Acc*I (lanes 3 and 7) and *Cla*I (lanes 4 and 8). The hybridisation conditions are detailed in the text.


A
B
kbp
1
2
3
4
5
6
7
8
6
4
2.5
1.6

product ~800-bp in size, however, it was unclear if this fragment had been generated by the mixed primer pair or primer 1 alone (Section 3.4.2). The 800-bp product was subsequently reduced in size by a deletion event upon cloning into the phage vector, M13mp19RF (Section 3.4.3). DNA sequencing 12 recombinants revealed that clone pAV35 contained a 320-bp insert whose deduced amino acid sequence was homologous to other EPSP synthases. Sequencing the insert DNA showed that primer 1 had bound specifically to its target site but could not establish the same for primer 2 since the deletion event had removed the 3' end of the cloned PCR product. Nevertheless, the 320-bp insert of pAV35 was used as a probe to hybridise to Southern blots of *A. variabilis* genomic DNA under stringent hybridisation conditions (Section 3.5). The resulting signal was quite faint which suggested that the amplified product was synthesised from contaminating DNA that may have been present in the original stock solution of genomic DNA (Section 3.4). Contamination probably occurred whilst growing the *A. variabilis* cultures. Cyanobacteria grow to very low densities and are generally cultured in large volumes over a 3 - 4 week period, so the risk of contamination is very high. The degenerate primers that synthesised the ~800-bp PCR product bound more specifically to the contaminating DNA than to *A. variabilis* DNA. This suggested that one or both of the primer target sites in *A. variabilis* are sufficiently different from Gram negative bacteria and higher plant EPSP synthase sequences from which the primers were designed (Figure 3.2). This was supported when primers 1 and 2 failed to amplify a DNA fragment of the desired size from uncontaminated *A. variabilis* genomic DNA (Section 3.6.1).

The identity of the contaminating *aro*A-like fragment has not yet been determined. The deduced amino acid sequence of this product is not identical to any EPSP synthase so far isolated from bacteria, fungi or plants (shown in Figure 3.4 or logged in the genembl data-base). The source of contamination may have been one of the various organisms regularly used in the laboratory that originally cultured the cyanobacterium. Southern or dot blots of genomic DNA from different organisms

could be performed using the 320-bp *aro*A-like fragment as a probe to identify the contaminant.

3.8.2 Cleavage of the 800-bp PCR Amplified Product

The 800-bp PCR product was site specifically cleaved when cloned into M13mp19RF, resulting in an insert 320-bp in length (Section 3.4.4). Cloned inserts of foreign DNA are capable of forming non-standard secondary structures which are then deleted by the recombination and repair systems of the *E. coli* host. As such, *E. coli* SURE™ was chosen for cloning the unstable ~800-bp PCR product instead of *E. coli* TG1. As well as carrying mutations eliminating all known restriction systems, *E. coli* SURE™ contains a series of mutations inactivating certain recombination and repair systems. These changes in characteristics of the host minimises the occurrence of DNA rearrangements (Greener, 1990 and Section 3.4.5). Despite the advantages of using *E. coli* SURE™, the PCR product was again site specifically cleaved whilst cloning the insert into M13mp19RF.

Deletions or rearrangements can occur when insert DNA is cloned into the polycloning site of the phage vector series M13mp. This is because the polycloning site is positioned in a region which contains initiation and termination sites for (+) and (-) strand DNA replication and transcription start and stop sites (Messing, 1983). Thus, foreign DNA could interfere with the proper functioning of this region. Since cloned DNA is not essential for phage propagation, large recombinants are selected against because of their slower growth rates upon transformation into their *E. coli* host. Deletion of recombinant DNA would lessen the burden on the host and lead to an increase in the cell growth rate. So the proportion of cells containing the original, undeleted recombinant would decrease with every cell division. However, rearrangements usually occur with larger inserts (> 5-kb) not with those DNA fragments in the optimal size range between 100 - 900-bp. So deletion of the 800-bp PCR product for this reason would have been unusual.

3.8.3 Primer Design

This chapter describes the various attempts made to amplify a fragment of the *aro*A gene from *A. variabilis* genomic DNA. Apart from the problems encountered with contamination (Section 3.8.1), the lack of success in reaching this goal may have been due to less than favourable reaction conditions. Optimisation of PCR conditions is an empirical process. Various parameters were tested in an attempt to amplify the desired product (Sections 3.4.2 and 3.6.1) but without success.

The design of the PCR primers could have affected the specificity of the amplification reaction (this has been illustrated by Schmitz *et al.*, 1993). *Anabaena variabilis* is naturally tolerant to glyphosate because the target enzyme, EPSP synthase, is resistant to inhibition by the herbicide (Powell *et al.*, 1992 and Section 1.9.2). Primers 1 and 3 were designed from the putative glyphosate binding region (Figures 3.2 and 3.7) of the available EPSP synthase sequences. Resistance of *A. variabilis* to glyphosate may result from a specific amino acid sequence at this region which is significantly different to those EPSP synthases already isolated from glyphosate sensitive organisms. This could be the reason why an *aro*A-like sequence was amplified from contaminating DNA and not *A. variabilis* DNA using the degenerate primers 1 and 2 (designed from glyphosate sensitive, Gram negative bacteria and higher plant sequences). Also, this may be why PCR with primers 3 and 4 (designed from glyphosate sensitive, Gram positive bacterial sequences) did not amplify an *A. variabilis aro*A fragment.

Primers 3 and 4, nevertheless, were originally used to amplify an *aro*A gene fragment from the glyphosate tolerant cyanobacterium, *Synechocystis* 6803 (dalla Chiesa *et al.*, 1994). The deduced EPSP synthase sequence of *Synechocystis* is similar to the *B. subtilis* and *S. aureus* enzymes especially at the putative glyphosate binding region (Mayes *et al.*, 1993) from which primer 3 was designed. The cause of glyphosate tolerance in *Synechocystis* has not yet been established but could be due to production of a glyphosate tolerant EPSP synthase or overproduction of the enzyme (Powell *et al.* 1991). If the former hypothesis is true, glyphosate tolerance in

Synechocystis 6803 would not be caused by changes in amino acid sequence at the putative binding site of the herbicide. So the EPSP synthases of the cyanobacteria, *Synechocystis* and *A. variabilis*, could feasibly share significant homology, especially at the putative glyphosate binding region. If this were the case, primers 3 and 4 should have bound to their respective target sites in the *A. variabilis aro*A gene.

The specificity of the amplification reaction could have been further improved if the codon bias of *A. variabilis* DNA had been incorporated into the primer design. The primers used for the experiments described in this chapter were designed to include all possible sequence permutations at positions of degeneracy (Figures 3.2 and 3.7). At the time these experiments were performed the information available regarding codon bias was very limited (Tandeau de Marsac and Houmard, 1987; Campbell and Gowri, 1990). In the former study, the authors acknowledged more genes had to be analysed to draw more significant conclusions regarding codon usage in cyanobacteria. Despite the lack of codon bias in their design, primers 3 and 4 efficiently amplified the *Synechocystis aro*A gene fragment (dalla Chiesa *et al.*, 1994). This success could be attributed to the type of template DNA used in the PCR experiment. The *Synechocystis* PCR product was amplified from pSM01, a pUC18 derivative containing a 5.5-kb insert of genomic DNA carrying the EPSP synthase gene, not from genomic DNA. The high copy number of *aro*A sequences in this mixture of plasmid DNA would probably have simplified the PCR reaction. When attempting to amplify the *A. variabilis aro*A gene fragment, primers 3 and 4 may have amplified non-target sequences under the various reaction conditions tested because of the complex nature of the genomic DNA. Incorporating codon bias into the primer design could have increased the specificity of primer annealing and reduced the frequency of mispriming. A recent study by Krishnaswamy and Schanmugasundraram (1995) analysed the codon usage in a large number of genes from various cyanobacterial species, including *Anabaena*.. The information from that research would be valuable for any future experiments.

Chapter 4

Hybridisation of the *Synechocystis aroA* Probe to an *Anabaena variabilis* Genomic DNA Library

4.0 Common Introduction to Chapters 4, 5 and 6

One of the most common gene cloning strategies is the construction of a library of DNA fragments containing the gene of interest. An ideal genomic DNA library is a collection of independent, overlapping clones representative of the genome (Kaiser and Murray, 1985). These clones must be maintained in a stable form and contain fragments large enough to carry whole genes but small enough to be easily mapped by restriction analysis. The following chapters describe the construction of several *A. variabilis* genomic DNA libraries which were screened for the *aro*A gene using two methods: (i) nucleic acid hybridisation using a variety of *aro*A probes from various sources and (ii) complementation of an *aro*A mutation of the auxotroph, *E. coli* AB2829. The intricacies and virtues of each technique with regard to this work are described and discussed in this and the two succeeding chapters.

4.1 Heterologous Hybridisation

Heterologous hybridisation was the first strategy ever used to clone a cyanobacterial gene (Mazur, Rice and Haselkorn, 1980). This technique utilises a previously characterised gene from one source as a probe to isolate the same gene from another source. The success of the technique is dependent on the sequence similarity and binding specificity between the probe and target DNA. Therefore, the probe should be a purified internal fragment of the gene corresponding to a functionally and / or structurally conserved part of the encoded protein. To isolate the gene of interest, a DNA library is constructed and screened under the appropriate hybridisation conditions using a radiolabelled, heterologous DNA probe to detect complementary sequences. Successive rounds of screening are performed to isolate a pure, positively hybridising clone. A large number of cyanobacterial genes have been cloned using bacterial, algal chloroplast, higher plant chloroplast and cDNA sequences as heterologous probes (Tandeau de Marsac and Houmard, 1987). Recently, cyanobacterial genes have been used as heterologous probes due to an increase in the number of genes cloned from these organisms. Cyanobacterial probes have been used to clone the fatty acid desaturase and vanadium dependent nitrogenase genes of *A. variabilis* (Sakamoto *et al.*, 1994 and Thiel, 1993) and the phosphoenolpyruvate carboxylase gene of *Anabaena* sp. PCC 7120 (Luinenburg and Coleman, 1992) from different phage libraries.

In this chapter, the construction of an *A. variabilis* genomic DNA library in the phagemid vector, pBluescript SK-, is described. A probe derived from the *aro*A gene of the unicellular cyanobacterium, *Synechocystis* sp. PCC 6803, was used to screen the library for the clone of interest. This strategy has been successfully used to clone other cyanobacterial genes. For instance, the genes encoding the *Synechocystis* reaction centre proteins of photosystem I (*psa*AB) and the *Anabaena* sp. PCC 7120 gene for flavodoxin were isolated from plasmid libraries using cyanobacterial genes as probes (Smart and McIntosh, 1991 and Leonhardt and Straus, 1989).

4.2 Strategy for Isolating the *Anabaena variabilis aro*A Gene By Heterologous Hybridisation

The *aro*A gene of the unicellular cyanobacterium *Synechocystis* sp. PCC 6803 (dalla Chiesa *et al.*, 1994) was chosen as the heterologous probe in an attempt to isolate the *A. variabilis* EPSP synthase gene. Southern blots of restriction digested *A. variabilis* genomic DNA were probed with an internal fragment of the *Synechocystis aro*A gene to determine the optimum hybridisation conditions and binding specificity of the probe to the target DNA. To detect the *aro*A gene, a DNA library representing the *A. variabilis* genome was constructed in pBluescript SK- and screened with the *Synechocystis* probe.

4.3 Southern Hybridisation of the *Synechocystis* sp. PCC 6803 *aro*A Probe to *Anabaena variabilis* Genomic DNA

4.3.1 Preparation of the *Synechocystis* sp. PCC 6803 *aro*A Probe

As stated in Section 4.1, a heterologous probe must be a pure, internal gene fragment preferably corresponding to a conserved part of the enzyme with functional and / or structural significance. This tends to minimise problems encountered with non-specific hybridisation. The *aro*A gene has been cloned from various bacteria, fungi and plants (Section 1.7.4). During the course of this work, the EPSP synthase gene from the unicellular cyanobacterium, *Synechocystis* sp. PCC 6803 was cloned by Mayes *et al.* (1993). A number of probes derived from unicellular cyanobacteria have been used to isolate genes from different *Anabaena* species (Thiel, 1993; Sakamoto *et al.*, 1994). Therefore, an internal fragment of the *Synechocystis aro*A gene was used as a probe in an attempt to isolate the EPSP synthase sequence from an *A. variabilis* genomic DNA library.

A 927-bp fragment of the *Synechocystis aro*A gene, accounting for approximately 70% of the coding sequence, was used for these hybridisation

experiments (Figure 4.1). The deduced amino acid sequence of this gene fragment shared homology along its length to other bacterial, fungal and plant EPSP synthases. Therefore, hybridisation between the 927-bp probe and target DNA would probably be due to several homologous regions. This would result in the formation of a stable duplex and a strong hybridisation signal that would be advantageous for screening experiments. Employing a large segment of the *aro*A gene as a probe was considered more prudent than using small fragments. Hybridisation of smaller probes, derived from different regions of the *Synechocystis aro*A gene, to *A. variabilis* DNA would depend on the degree of homology shared at one particular region. This would probably not result in such a reliable hybridisation signal.

The 927-bp *Synechocystis aro*A fragment was amplified by the PCR using primer set 3 and 4 (Table 3.1 and Figure 3.7) and the plasmid, pSM01, a pUC18 derivative which carries a 5.5-kb fragment of the *Synechocystis* genome containing the *aro*A gene (Mayes *et al.*, 1993). Both primers 3 and 4 and pSM01 were the kind gifts of Drs. M. dalla Chiesa and D. Maskell (Imperial College of Science, Technology and Medicine, London, England). PCR of 100 ng pSM01 was performed as described in Section 2.14.1 in the presence of 1.5 mM $MgCl_2$, 200 μM dNTP and 40 pmol primers 3 and 4. The PCR parameters were as follows: denaturation step at 94 °C for 3 minutes ; 30 cycles at 94 °C for 1 minute, 45 °C for 2 minutes and 72 °C for 2 minutes, lastly an extension step at 72 °C for 5 minutes completed the reaction. The amplified 927-bp *aro*A fragment was purified from a low melting point agarose gel as in Section 2.12.3 for subsequent use.

4.3.2 Optimal Conditions for Hybridisation of *Synechocystis* sp. PCC 6803 *aro*A Probe to *Anabaena variabilis* Genomic DNA

A strong signal produced from hybridisation of the *Synechocystis aro*A probe to *A. variabilis* genomic DNA was essential for screening a DNA library to isolate a clone carrying the EPSP synthase gene.

The strength of a hybridisation signal is dependent on the degree and extent of complementarity shared between the two sequences that form the DNA duplex (Anderson and Young, 1985). In other words, the stability of the hybrid DNA. This can be measured by the T_m (or melting temperature), the temperature at which the strands are half dissociated. The T_m can be calculated for homologous sequences using the equation of Meinkoth and Wahl (1984):

$$T_m = 81.5 + 0.41(\%G+C) + 16.6\log[Na^+] - 0.61(\%\ \text{formamide}) - 500 / n \quad [1]$$

where G + C is the percentage guanine plus cytosine present in the DNA, $[Na^+]$ is the molarity of sodium ions in solution and n is the length of the shortest chain of the duplex. Hybridisation reactions are usually carried out at temperatures 25 ^oC below the duplex T_m. This is because at 0 ^oC the hybridisation rate is very slow. But as the temperature increases so does the hybridisation rate until it reaches a broad maximum approximately 25 ^oC below the duplex T_m (Marmur and Doty, 1961). This relationship applies to the formation of heterologous as well as homologous duplexes. However, the T_m of nucleic acid hybrids can be depressed by base mismatching. A 1% mismatch can reduce the T_m between 0.5 and 1.4 ^oC (Anderson and Young, 1985). Therefore, the optimum hybridisation temperatures for poorly matched duplexes tend to be lower than that for well matched duplexes.

The Meinkoth and Wahl (1984) equation illustrates that the ionic strength of the hybridisation solution can affect the stability of the DNA hybrid. High ionic strength solutions can stabilise poorly matched duplexes. The ionic strength of a solution can also affect the hybridisation rate (Wetmur and Davidson, 1968). At low ionic strength, homologous or heterologous sequences hybridise slowly. As the ionic strength increases so does the hybridisation rate. Therefore, high ionic strength solutions are used for hybridisation reactions and to maintain the stability of heterologous duplexes.

The %GC content of DNA can influence the melting temperature of homologous and heterologous hybrids. A G-C pair has three hydrogen bonds whereas an A-T pair shares only two. Therefore, hybrid DNA formed from DNA with a high %GC content is very stable (Wetmur and Davidson, 1968). However, the %GC content of *A. variabilis* and *Synechocystis* DNA is moderate at 42.5 and 47% respectively.

Formamide is a denaturing agent that can reduce the melting temperature of nucleic acid hybrids and ,therefore, affect the incubation temperature for hybridisation reactions (as shown in equation one). Therefore, solutions containing formamide allow hybridisations to be performed at lower temperatures than would otherwise be the case (McConaughy *et al.*, 1969). However, this denaturing agent was not used in the experiments described in this chapter so did not influence duplex formation between the *Synechocystis* probe and target DNA.

Since the homology shared between the *Synechocystis* and *A. variabilis aro*A genes was unknown, the optimum hybridisation conditions had to be determined empirically. The temperature of hybridisation had to be low enough to produce a strong signal but not so low that non-specific hybridisation occurred. The salt concentration of the hybridisation buffer had to be high enough to stabilise heterologous DNA duplexes during the hybridisation and the initial washing steps. The stringency of the washing conditions was increased by gradually decreasing the salt concentration of the wash buffer. After each wash step the filters were autoradiographed to analyse the specificity and strength of the resulting signal. This eventually revealed duplexes formed from closely related sequences. Various experimental conditions were tested. The optimal hybridisation conditions are outlined below.

25 ng of the 927-bp *Synechocystis aro*A probe was radiolabelled to a specific activity of 2.3 x 10^9 cpm / µg, purified (Sections 2.21 and 2.22) and hybridised to 1 µg *Hin*dIII and *Eco*RI digested *A. variabilis* genomic DNA in 4 x SET at 50 ^{o}C (Sections 2.23 and 2.24). The membranes were subsequently washed from 2 x SET to 1 x SET at 50 ^{o}C and autoradiographed for 16 - 18 hours. Figure 4.2(b) shows the resulting

hybridisation signal. Southern blots performed at 55 to 60 °C or washing the membranes in 0.1 x SET at 50 °C produced signals which were too faint for heterologous screening experiments.

4.3.3 Binding Specificity of the *Synechocystis* sp. PCC 6803 *aro*A Probe to *Anabaena variabilis* Genomic DNA

As well as a strong signal, specific hybridisation of the *Synechocystis aro*A probe to *A. variabilis* genomic DNA was also essential to isolate the EPSP synthase gene by heterologous screening of a DNA library. Therefore, the 927-bp *Synechocystis* probe was cut into fragments and hybridised to different digests of *A. variabilis* genomic DNA. The resulting signal patterns of each *aro*A gene fragment were compared to verify that their binding specificities were identical.

The 927-bp *Synechocystis aro*A probe was digested separately with *Bgl*II or *Hin*cII (Section 2.9.3) to generate fragments 618 and 275-bp in size respectively. Figure 4.1 illustrates the positions of the *Bgl*II and *Hin*cII fragments in relation to the full length *Synechocystis aro*A gene. The 618 and 275-bp *aro*A probes were purified from a 2% l.m.p. agarose gel (Section 2.12.3). 25 ng of each probe was radiolabelled to specific activities of 4.6 - 5.8 x 10^8 cpm / µg DNA respectively (Sections 2.21 and 2.22). Both probes were hybridised to *Hin*dIII and *Eco*RI digests of *A. variabilis* genomic DNA at 50 °C (as in Section 4.4.2). Figure 4.2 compares the resulting hybridisation patterns of the 618 and 275-bp probes to the 927-bp *aro*A probe.

The 927, 618 and 275-bp *Synechocystis aro*A probes all hybridised to a 2.5-kb *Eco*RI fragment. The 275-bp probe did not produce a signal from the *Hin*dIII digest, yet the 927 and 618-bp *aro*A fragments hybridised strongly to a 3.8-kb fragment. Assuming the *Synechocystis aro*A probe is specifically hybridising to the *A. variabilis* EPSP synthase gene, these results suggest that the *A. variabilis aro*A gene may have an internal *Hin*dIII site. If this is the case, the restriction map of the *A. variabilis aro*A gene differs from the *Synechocystis* gene which lacks sites for *Hin*dIII.

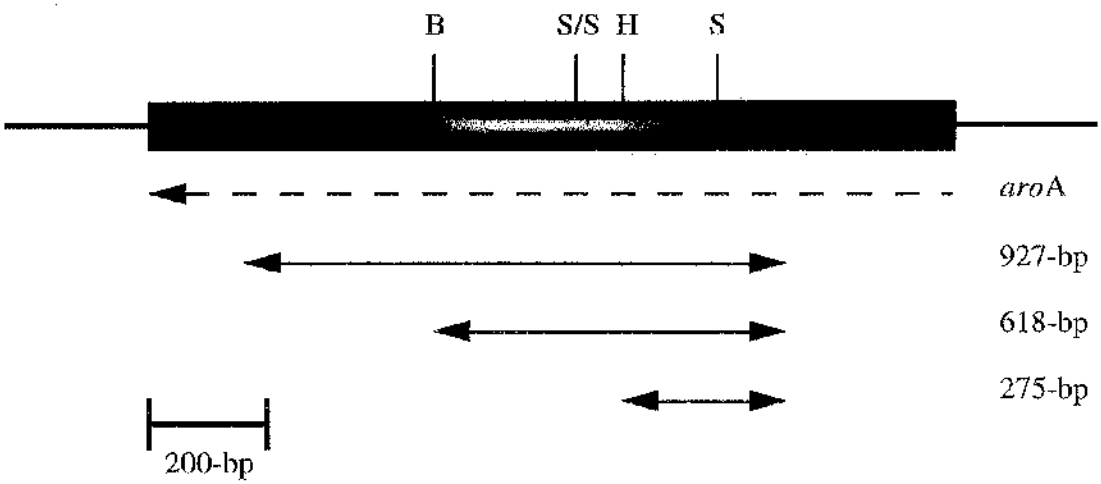
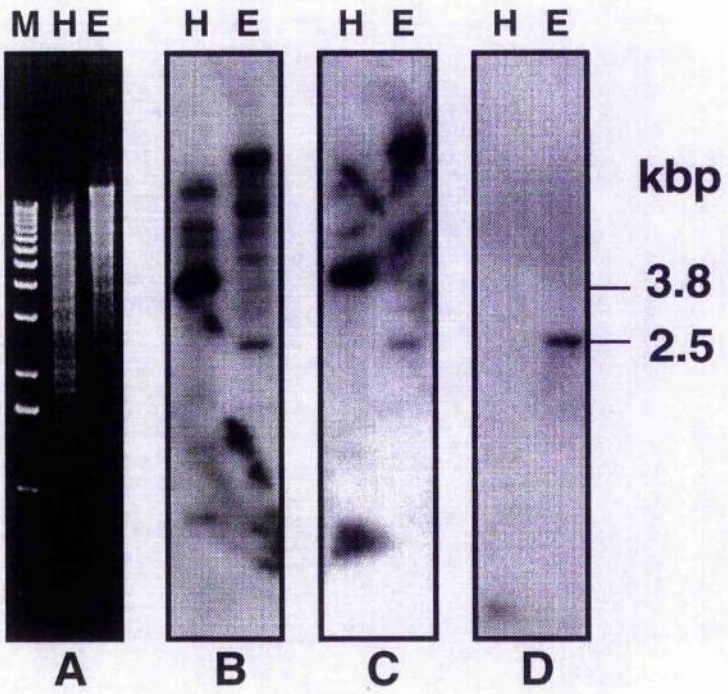


Figure 4.1 The positions of the *Synechocystis aro*A probes relative to the full length EPSP synthase gene. The boxed region denotes the *Synechocystis aro*A gene. The thin lines either side of the *aro*A gene represent *Synechocystis* genomic DNA. Double arrowed lines represent the *aro*A gene fragments used as probes. The dotted arrow signifies the direction of transcription of the *aro*A gene from the 5' to the 3' end. B = *Bgl*II, S/S = *Sma*I/*Stu*I, H = *Hin*cII and S = *Sma*I.

Figure 4.2 Hybridisation of different sized *Synechocystis aro*A gene fragments to *Anabaena variabilis* genomic DNA. (A) Ethidium bromide stained gel of 1 μg *Anabaena variabilis* genomic DNA digested separately with *Hin*dIII (H) or *Eco*RI (E). M = DNA size markers: 1-kb DNA ladder from Gibco BRL, Paisley, Scotland. Southern hybridisation of the *Synechocystis* 927-bp (B), 618-bp (C) and 275-bp (D) *aro*A probes to the restriction digested *A. variabilis* genomic DNA shown in (A). Hybridisation and washing conditions are described in the text. The sizes of the hybridising bands are indicated.


M H E
H E
H E
H E
kbp
3.8
2.5
A
B
C
D

The hybridisation patterns illustrated that different fragments of the *Synechocystis aro*A gene were able to bind specifically to *A. variabilis* genomic DNA. Therefore, the *Synechocystis* probe must share homology along its length to the target sequence within the *A. variabilis* DNA as predicted in Section 4.3.1. The identical hybridisation pattern of each *aro*A fragment to the target DNA suggested that the *Synechocystis* probe was binding to genomic fragments carrying the *A. variabilis aro*A gene. Therefore, this heterologous probe was considered suitable for screening an *A. variabilis* genomic DNA library to isolate the EPSP synthase gene.

4.4 Construction of the *Anabaena variabilis* Genomic DNA Library for Heterologous Screening

An *A. variabilis* genomic DNA library was constructed in the phagemid vector, pBluescript SK-. To increase the probability of isolating the *A. variabilis aro*A gene, genomic DNA was digested with *Hin*dIII and size fractionated prior to cloning into pBluescript. *Hin*dIII was used to generate donor fragments because hybridisation of the 927 and 618-bp heterologous *aro*A probes to *Hin*dIII digested *A. variabilis* DNA suggested that at least part of the EPSP synthase gene was located on a 3.8-kb fragment (Figure 4.2). pBluescript SK- was chosen for this work because, like other plasmid vectors, it is ideal for cloning DNA fragments less than 10-kb in size. Also, pBluescript replicates to a high copy number within *E. coli*. The large number of recombinant clones within the transformed cell should allow easier detection of the *A. variabilis aro*A clone by colony hybridisation.

4.4.1 Preparation of *Anabaena variabilis* Genomic DNA for Library Construction

A. variabilis cells were cultured and checked for contaminating growth as in Section 2.3.1/2. The extraction of genomic DNA was performed essentially as in Section 2.7 with some minor changes to improve the yield of DNA. Since construction of the *A. variabilis* DNA library in pBluescript SK- involved cloning fragments up to

10-kb in size, isolation of high molecular weight genomic DNA (> 50-kb) was unnecessary. So more vigorous extraction techniques than those outlined in Section 2.7 were employed to obtain higher yields of lower molecular weight DNA. This involved increasing centrifugation speeds from 3-krpm to 10-krpm to harvest cells more efficiently and to achieve better separation of organic and aqueous phases after more vigorous phenol / chloroform extractions. These modifications to the protocol yielded 500 µg of *A. variabilis* genomic DNA, 20 - 30-kb in size.

20 µg of *A. variabilis* genomic DNA was digested with *Hin*dIII (Section 2.9.1), electrophoresed through a l.m.p. agarose gel and size fractionated from 2.5 to 12-kb as described in Section 2.15. The DNA was quantitated (Section 2.4.2) and the concentration adjusted to 100 ng / µl with TE buffer.

4.4.2 Ligations of *Anabaena variabilis* genomic DNA into pBluescript SK-

Theoretically it is possible to calculate the required amounts of DNA which favour the formation of recombinant clones (Sambrook *et al.*, 1989). In practice, however, problems arise due to damaged restriction ends which occur when preparing DNA and the slight degree of error when quantitating DNA. So the optimal concentrations of vector and donor DNA required to encourage the formation of recombinant molecules were determined empirically. Several test reactions were carried out to find this ratio due to the reasonably large quantity of donor DNA that had been prepared (Section 4.4.1).

pBluescript SK- was digested with *Hin*dIII and dephosphorylated for the ligation reactions (Section 2.15). Dephosphorylation reduces the occurrence of recircularisation of vector DNA and promotes formation of recombinants. The reaction conditions outlined in Section 2.15 were used to ligate 100 ng of pBluescript separately to 100, 200 and 300 ng of *Hin*dIII cut, size fractionated *A. variabilis* genomic DNA. Additionally, 400 ng of pBluescript was ligated to 100 ng of donor DNA.

To ascertain the proportion of recombinant to non-recombinant clones present in each test library, the ligation products were transformed into competent *E. coli* DH5α (Section 2.15). Transformants harbouring recombinant clones were identified by α-complementation (or blue / white colony selection, Section 2.17.4). The number of recombinant clones in each library is shown in Table 4.1. However, α-complementation can give rise to false-positive results (Yanisch - Perron *et al.*, 1985). Therefore, phagemid DNA from each library was purified from 12 randomly selected white colonies (Section 2.5.1) and subsequently digested with *Hin*dIII (Section 2.9.1) to release the insert DNA. Analysis of the digested DNA by agarose gel electrophoresis (Section 2.11) showed the proportion of white colonies harbouring recombinant clones (Table 4.1). This information allowed a more precise estimation of the ratio of recombinant to non-recombinant clones in each test library (Table 4.1).

Table 4.1. Number of recombinant clones in each *A. variabilis* test library made in pBluescript SK-.

Library [1]	% Recombinants α-complementation	% Recombinants in white colonies	Estimated % recombinants [2]
1:1	82	98	80
1:2	89	100	89
1:3	92	100	92
4:1	61	82	50

[1] The libraries are represented as a ratio of the nanogram quantities used for each reaction. [2] The estimated number of recombinants was calculated for each library by taking the results obtained from blue/white colour selection and the phagemid preparations into consideration.

To establish if the libraries were of a sufficient size and complexity to contain the *aro*A gene, the number of clones to be screened to have an arbitrary probability of containing any particular sequence was determined using the equation of Clarke and Carbon (1976):

$$N = \ln (1 - p) / \ln (1 - x/y) \qquad [2]$$

Equation 2 states that N number of clones will have a probability (p) of containing any one specific DNA sequence when x is the insert size (in bp) and y is the size of the haploid genome (in bp). The equation assumes complete randomness of sequence representation in the library and that each inserted fragment is of identical size. In practice, these conditions could not be guaranteed, therefore, the equation was a guide to determining the size of the library required for screening purposes.

Analysis of the recombinant clones of each test library by restriction digestion and agarose gel electrophoresis allowed the average insert size for each library to be determined (Table 4.2). The size of the *A. variabilis* genome is 5.4 x 10^6-bp (Herrero *et al.*, 1984). Therefore, the number of individual recombinant clones with a 99% probability of representing the sequence of interest was calculated for each clone bank (Table 4.2). When determining N, the proportion of non-recombinant clones present in each library had to be considered. For example, 20% of the 1:1 library consisted of non-recombinant vector DNA. As a result a further 1243 colonies harbouring recombinants would have to be screened to have a 99% probability of isolating one particular sequence (Revised library size, Table 4.2).

Table 4.2. The number of recombinant clones to be screened to have a 99% probability of isolating the *A. variabilis aro*A gene

Library	Average insert size (kbp)	Library Size (N)	Revised library Size (N) *
1:1	4	6216	7459
1:2	3.49	7187	7977
1:3	3.23	7796	8419
4:1	2.87	8679	13,018

* Revised library size takes the proportion of non-recombinant clones in each library into consideration.

The 1:1 library was selected for heterologous screening because the average insert size (4-kb) increased the probability of cloning the 3.8-kb *Hin*dIII fragment which putatively carries the *A. variabilis aro*A gene. Also, the size of the library calculated to give a 99% chance of containing one copy of a specific sequence was, in practice, a reasonable number of clones to screen.

The Clarke and Carbon equation (1976) estimated the number of clones required to screen a library representative of the entire *A. variabilis* genome. Since *Hin*dIII digested *A. variabilis* DNA had been size fractionated from 2.5 - 12-kb before ligation into pBluescript, the frequency of occurrence of a particular sequence in that size range was expected to be higher. Therefore, the *A. variabilis* library was screened using a homologous probe to more accurately determine its size and quality.

4.5 Evaluation of the *Anabaena variabilis* Genomic DNA Library in pBluescript SK-

The quality of the size fractionated *A. variabilis* DNA library was evaluated to determine if it was likely to be of a sufficient size and complexity to contain the *aro*A gene. This was initially indicated when recombinant clones of the *A. variabilis* library were restriction digested and 72% were shown to carry unique restriction patterns (Section 4.4.2). To verify this result, the library was screened using the genes encoding the vanadium dependent nitrogenase of *A. variabilis* ATCC 29413 (Thiel, 1993). The *vnf*DGK operon of *A. variabilis* occurs only once per genome. It also lacks *Hin*dIII sites. So the operon would not be split into separate clones in the *A. variabilis* library which had been constructed from *Hin*dIII genomic fragments. According to the calculation made in Section 4.4.2, the probe was expected to hybridise to approximately one clone in every 8000. However, due to size fractionation of the donor DNA, the size of the library would probably be reduced.

4.5.1 Preparation of the *Anabaena variabilis vnf*DGK Probe

The *vnf*DGK operon is carried on a 3.3-kb *Cla*I fragment of *A. variabilis* genomic DNA in the pUC119 derivative, pRSI, which was kindly donated by Dr. Teresa Thiel, University of Missouri, St. Louis, Missouri, U.S.A. 5 µg pRSI was digested with an excess of *Cla*I and *Sca*I to release the *vnf* operon from pUC119 (Section 2.9.4). Digestion with *Cla*I alone produced two restriction fragments 3.3-kb (insert DNA) and 3.2-kb (vector DNA) in size which could not be separated by agarose gel electrophoresis. The second digestion with *Sca*I cleaved pUC119 into two smaller fragments, allowing good electrophoretic separation of the *vnf* probe from plasmid DNA through a l.m.p. agarose gel. The *vnf* probe was subsequently purified as in Section 2.12.3.

4.5.2 Colony Hybridisation of the Homologous *vnf*DGK Probe to the *Anabaena variabilis* Genomic DNA Library

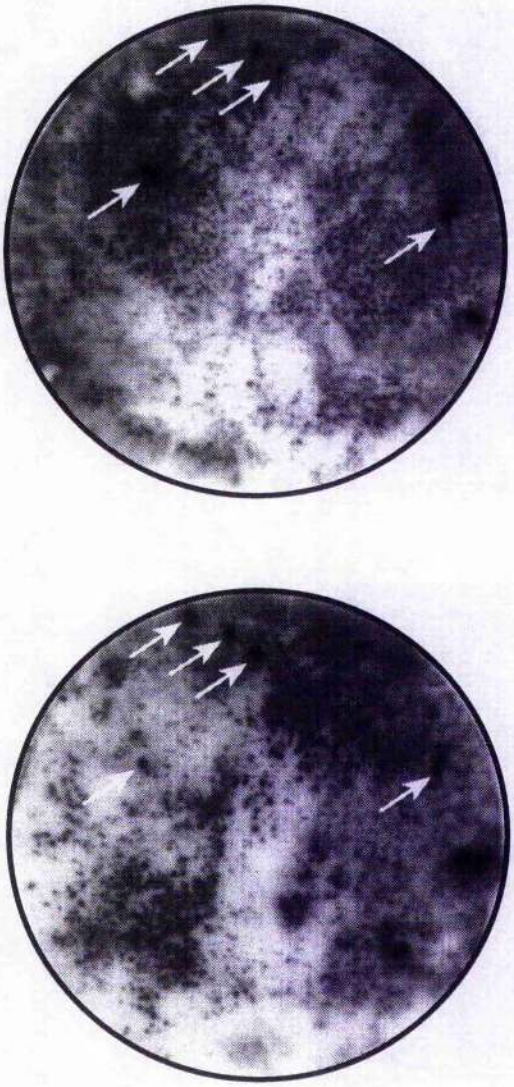
The *A. variabilis* genomic DNA library was transformed into *Epicurian Coli*® XL2-Blue MRF' (Sections 2.15 and 4.6, Jerspeth *et al.*, 1992 and Table 2.1 for details). Approximately 20,000 transformants were plated onto 15 cm diameter LB / ampicillin plates. The transformants were replicated onto nylon membranes from which duplicates were made. DNA was then liberated, denatured and immobilised onto the membranes as in Section 2.26.1. 100 ng of the *A. variabilis vnf* probe was radiolabelled to a specific activity of 5 x 10^8 cpm/μg and hybridised to the library at 65 ^oC (Section 2.26.2). Washing the filters from 2 x SET to 0.1 x SET at 65 ^oC (Section 2.26.3), followed by an 18 hour exposure to X-ray film produced the strong, duplicate hybridisation signals shown in Figure 4.3. Since the probe bound to five clones in every 20,000 transformants, there was one copy of the *vnf*DGK operon present for approximately every 4000 clones.

The analysis of insert size and homologous screening using the *vnf*DGK probe confirmed that the *A. variabilis* library, constructed from size fractionated *Hin*dIII genomic fragments, was likely to be of a satisfactory size and complexity to contain the *aro*A gene.

4.6 Screening the *Anabaena variabilis* Genomic DNA Library in pBluescript with the *Synechocystis aro*A Probe

The experiments described previously showed that the *Synechocystis aro*A probe hybridised strongly to specific restriction fragments of *A. variabilis* genomic DNA. Also, the library of size fractionated *Hin*dIII fragments constructed in pBluescript SK- was likely to be of a sufficient size and complexity to contain the *A. variabilis aro*A gene. Therefore, the *A. variabilis* library was screened with the *Synechocystis aro*A probe in an attempt to isolate an EPSP synthase clone.

Figure 4.3 Autoradiogram illustrating hybridisation of the *Anabaena variabilis vnf*DGK probe to replicate membranes carrying approximately 20,000 clones of the *A. variabilis* genomic DNA library in pBluescript (grown in *E. coli*® XL2-Blue MRF'). The hybridisation and washing conditions are outlined in the text. Duplicate signals are indicated by white arrows.



Southern blot experiments showed that both the 927 and 618-bp fragments of the *Synechocystis aro*A gene hybridised to a *Hin*dIII fragment of *A. variabilis* genomic DNA, 3.8-kb in size (Section 4.3.3). However, the 618-bp fragment produced a much cleaner signal upon hybridisation to *A. variabilis* genomic DNA than the 927-bp probe (Figure 4.2). Therefore, the smaller probe was chosen for screening the phagemid library.

Epicurian Coli® XL2-Blue MRF' (Table 2.1 and Jerspeth *et al.*, 1992) was transformed with the *A. variabilis* library and approximately 20,000 ampicillin resistant transformants were grown on LB / ampicillin plates 9 or 15 cm in diameter as described in Section 4.5.2. XL2-Blue MRF' was the host of choice for these experiments for several reasons. The cells are described as ultra-competent which means higher transformation efficiencies can be attained. This was ideal for screening the *A. variabilis* library since more colonies would be obtained per transformation reaction. Additionally, *E. Coli*® XL2-Blue MRF' is deficient in all known restriction systems. So *A. variabilis* genomic DNA was less likely to be a substrate for cleavage by the host's defence mechanisms (Jerspeth *et al.*, 1992 and Section 6.6). Lastly, the mutation of the host's *rec*A gene would help to prevent integration of closely related foreign DNA sequences into the host chromosome and improve insert stability. This maintained the integrity and copy number of transformed vector DNA.

After transformation of the *A. variabilis* library into *E. Coli*® XL2-Blue MRF' (Section 2.15), the transformants were replicated onto nylon membranes and duplicated. Cell lysis, DNA denaturation and immobilisation were performed exactly as described previously (Section 2.26.1). 100 ng of the *Synechocystis* 618-bp *aro*A probe was radiolabelled to a specific activity of between 4 - 5 x 10^8 cpm / µg DNA and hybridised to the *A. variabilis* library at 50 ^{o}C (Section 2.26.2). The optimal conditions for hybridisation of the *Synechocystis* probe to *A. variabilis* DNA (determined in Section 4.3.2) were used for this experiment. Membranes were washed under conditions of increasing stringency from 4 x SET at ambient temperature to 0.5 x SET at 50 ^{o}C to remove as much background signal as possible (Section 2.26.3).

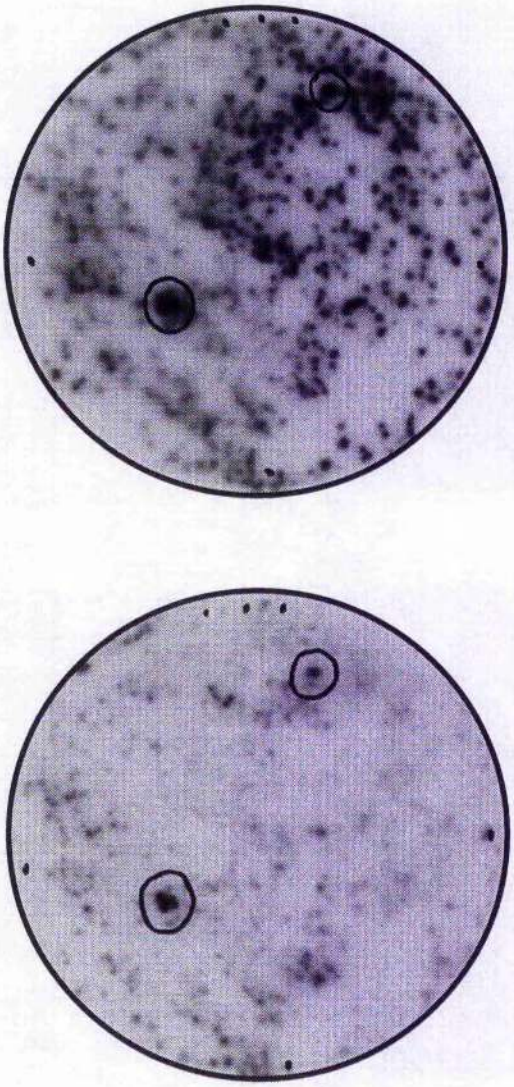
After each washing step the membranes were autoradiographed and the resulting signal examined to ensure hybridisation of the probe to target DNA.

According to the experiments performed in Section 4.5.2, one clone in every 4000 should contain the sequence of interest. Under the conditions used in these experiments, the *Synechocystis aro*A fragment bound to a larger number of clones than the predicted library size. Therefore, the probe must have hybridised non-specifically to an array of different sequences. This produced signals of varying intensity. Only colonies exhibiting strong, duplicate signals were picked and re-screened (Figure 4.4). The intensity of these signals suggested that the corresponding clones contained sequence homologous to the *Synechocystis aro*A probe.

Each clone was analysed in a second round of screening using identical hybridisation conditions (Section 2.26.4). However, the absence of any strong, duplicate signal from the re-screened putative positive clones showed that the *Synechocystis* probe had not bound specifically to the sequence of interest. In total, 60,000 colonies (approximately equivalent to 20 copies of the size fractionated library) were carefully screened. All clones producing convincing signals were re-screened. In spite of this, the *Synechocystis aro*A probe did not bind strongly to any of the re-screened clones. So the 3.8-kb *Hin*dIII fragment, putatively carrying the *A. variabilis aro*A gene, had not been isolated from the phagemid library.

Since the *Synechocystis aro*A probe was not binding specifically to the target *A. variabilis* sequence in the phagemid library, the false positive signals could have been caused by hybridisation of the heterologous probe to the *E. coli* host genetic material. The varying intensity of the hybridisation signal produced from the transformants harbouring the library could have been due to the different rates of growth exhibited by some cells. The stronger signals might have been produced from *E. coli* cells, containing smaller clones, able to replicate more efficiently than those transformants containing larger recombinants.

Figure 4.4 Autoradiogram showing hybridisation of the *Synechocystis* 618-bp *aro*A probe to duplicate membranes containing 500 clones of the *Anabaena variabilis* genomic DNA library in pBluescript (grown in *E. coli*® XL2-Blue MRF'). Hybridisation conditions are described in the text. Membranes were washed at 1 x SET at 50 °C and autoradiographed for 16 hours at -80 °C. Strong duplicate signals have been circled.



4.7 Discussion

When attempting to isolate genes from a DNA library by heterologous hybridisation, the probe being used must bind specifically to the sequence of interest (as illustrated in Smart and McIntosh, 1991, Luinenburg and Coleman, 1992 and Thiel, 1993). Southern blot experiments showed that the *Synechocystis aro*A probe hybridised strongly to specific fragments of restriction digested *A. variabilis* DNA (Section 4.3). In particular, a 618-bp *Bgl*II fragment of the *aro*A gene hybridised cleanly to a *Hin*dIII fragment, 3.8-kb in size. An *A. variabilis* library of size fractionated, *Hin*dIII fragments was then constructed in the phagemid, pBluescript SK- (Section 4.4). The quality of the library was assessed to show that it was likely to be of a sufficient size and complexity to contain the *A. variabilis aro*A gene (Section 4.5). One clone in 4000 was expected to contain the sequence of interest. In spite of this, when the library was screened with the *Synechocystis* probe under the pre-determined optimal hybridisation conditions, there was a succession of false positive results (Section 4.6). After screening the equivalent of 20 copies of the size fractionated library, the *A. variabilis aro*A clone had not been isolated. This was probably due to non-specific hybridisation of the heterologous probe to the genomic DNA of *E. Coli*® XL2-Blue MRF'.

Increasing the stringency of the hybridisation and / or washing conditions could have reduced the level of non-specific hybridisation, but may not have increased the sensitivity of the screening method. Southern blots performed at 55 and 60 ^{o}C produced weaker hybridisation signals than at 50 ^{o}C, the temperature used for screening purposes (Section 4.3.2). DNA / DNA hybridisations are usually performed about 25 ^{o}C below the melting temperature of a hybrid duplex but the T_m can be reduced by base mismatching (Anderson and Young, 1985 and Section 4.3.2). The number of mismatches shared between the *Synechocystis aro*A sequence and *A. variabilis* target sequence is unknown, however, the higher hybridisation temperatures may have been too close to the T_m of the heterologous duplex causing the hybrid to partially dissociate. Washing the membranes at 0.1 x SET at 50 ^{o}C, instead of at a

higher salt concentration, also produced faint signals. This was probably due to the instability of the heterologous duplex under low ionic conditions. This would also have been influenced by the degree of mismatching shared between the *Synechocystis* probe and *A. variabilis* target sequence.

Southern hybridisation of the *Synechocystis aro*A probe to the purified recombinants that comprised the *Anabaena variabilis* library could have shown if the clone bank contained the 3.8-kb sequence of interest. However, future attempts to clone the *A. variabilis aro*A gene using the *Synechocystis* or other heterologous probes would still require a reduction in the high level of non-specific hybridisation described in Section 4.6. So the amount of *E. coli* genomic DNA immobilised to the nylon membrane support had to be reduced. As discussed in Chapter 5, an *A. variabilis* genomic DNA library was constructed in the phage vector, lambda FIX® II, to counter this problem.

The homology shared between the *Synechocystis* and *A. variabilis aro*A sequences is unknown. Analysis of other genes cloned from these two cyanobacterial species show that the degree of homology can be reasonably extensive. For example, the fatty acid desaturase (*des*A) gene (Sakamoto *et al.*, 1994, Wada *et al.*, 1990), the glyceraldehyde-3-phosphate dehydrogenase (*gap*2x) gene (Martin *et al.*, 1993; Valverde *et al.*, 1995), and the *psa*AB genes encoding photosystem I reaction centre proteins (Nyhus *et al.*, 1994; Smart and McIntosh, 1991) have been cloned from *Synechocystis* PCC 6803 and *A. variabilis*. The nucleotide sequences, containing the open reading frames of these genes from the two cyanobacteria, were compared using the, "gap", function of the Wisconsin genetic computer group package (Section 2.30). The *des*A and *gap*2x genes of the two cyanobacterial species shared a 55% and 62% identity respectively and the *psa*AB genes were 74% identical.

The homology shared between other *Synechocystis* and *A. variabilis* sequences plus the results obtained from the Southern hybridisation experiments described previously (Section 4.3) highlight the value of using the *Synechocystis aro*A probe in these and future screening experiments.

Chapter 5

Screening an *Anabaena variabilis* Genomic DNA Library with *aro*A Probes from Different Sources

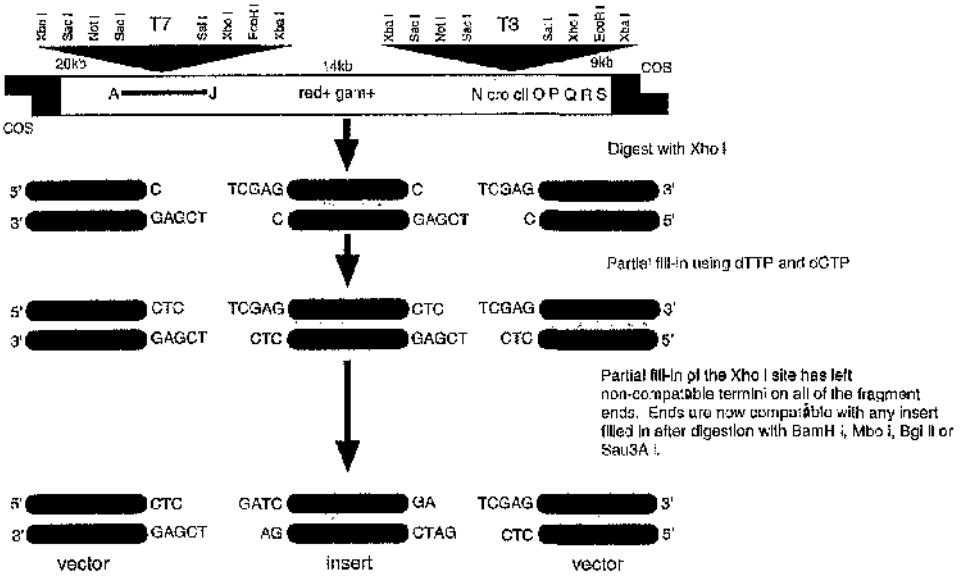
5.1 Introduction

The high level of background hybridisation that resulted from screening the *A. variabilis* library in pBluescript with the *Synechocystis aro*A probe prevented isolation of the *A. variabilis* EPSP synthase gene (Chapter 4). In an effort to improve the sensitivity of the screening procedure, an *A. variabilis* library was constructed in the phage vector, lambda FIX® II (Stratagene Limited, Cambridge, England and Table 2.2). Control experiments showed that plaque hybridisation was a more sensitive technique than colony hybridisation. As such, the phage library of *A. variabilis* DNA was screened using heterologous *aro*A probes from a variety of sources, namely: *Synechocystis* sp. PCC 6803, *E. coli* and pea. Screening the library with a number of different heterologous probes was performed to maximise the possibility of cloning the *A. variabilis* EPSP synthase gene. *Synechocystis, E. coli* and pea probes have been successfully used to isolate other genes from various *Anabaena* species, for example, the fatty acid desaturase gene of *A. variabilis* (Sakamoto *et al.*, 1994) and the glutamine synthetase and ribulose-1,5-bisphosphate carboxylase small sub unit genes of *Anabaena* sp. PCC 7120 (Fisher *et al.*, 1981 and Nierzwicki-Bauer *et al.*, 1984) respectively.

5.2 Strategy for the Construction of the *Anabaena variabilis* Genomic DNA Library in Lambda FIX® II

Figure 5.1 outlines the strategy used for the construction of the *A. variabilis* genomic DNA library in lambda FIX® II. This phage vector ligates DNA fragments between 9 - 23-kb in size. Southern hybridisation of the *Synechocystis aro*A probe to blots of *Sau*3AI digested *A. variabilis* DNA showed that the heterologous probe hybridised to fragments in this size range (Section 5.4). Therefore, high molecular genomic DNA was extracted from uncontaminated *A. variabilis* cells and partially digested with *Sau*3AI to generate donor fragments of the appropriate size. The resulting restriction ends were partially filled-in with dGTP and dATP to create

Figure 5.1 DNA library construction in lambda FIX® II. The partial fill-in technique as described in Section 5.3. (Diagram reproduced from Stratagene Catalogue (1995), Stratagene Ltd., Cambridge, England). After lambda FIX II is digested with *XhoI* and the restriction ends partially filled-in with dTTP and dCTP, the stuffer fragment of the vector is removed. Vector arms can now ligate to insert DNA prepared as described in the text.


Xba I Sac I Not I Sac I T7 Sal I Xho I EcoR I Xba I
Xba I Sac I Not I Spe I T3 Sal I Xho I EcoR I Xba I
20kb
14kb
9kb
COS
A J
red+ gam+
N cro cII O P Q R S
COS
Digest with Xho I
5' C
3' GAGCT
TCGAG C
C GAGCT
TCGAG 3'
C 5'
Partial fill-in using dTTP and dCTP
5' CTC
3' GAGCT
TCGAG CTC
CTC GAGCT
TCGAG 3'
CTC 5'
Partial fill-in of the Xho I site has left non-compatible termini on all of the fragment ends. Ends are now compatible with any insert filled in after digestion with BamH I, Mbo I, Bgl II or Sau3A I.
5' CTC
3' GAGCT
GATC GA
AG CTAG
TCGAG 3'
CTC 5'
vector
insert
vector

compatible termini for ligation into the replacement vector, lambda FIX® II (Sections 5.3 and 5.5; Figure 5.1). The ligation products were packaged into infectious phage particles before introduction into the host bacteria. The quality of the newly constructed library was then evaluated prior to screening with the heterologous *aro*A probes (Section 5.6). This established that the library was sufficiently large and complex to have a good chance of containing the *A. variabilis* EPSP synthase gene.

5.3 Reasons for Using the Replacement Vector, Lambda FIX®II

Construction of a library in the majority of phage vectors requires the initial preparation of several hundred micrograms of high molecular weight genomic DNA. The large quantity of DNA is required for size fractionation (Kaiser and Murray, 1985). This procedure enriches insert DNA in the optimum size range for the formation of recombinant phage genomes that can be packaged into viable virus particles. However, only 50 - 100 μg genomic DNA was extracted from *A. variabilis* cells using the protocol in Section 2.7. This quantity of *A. variabilis* DNA was too low to perform the size fractionation procedure. To overcome this problem, lambda FIX II was used to construct an *A. variabilis* DNA library. Lambda FIX II was supplied by Stratagene Limited (Cambridge, England) pre-digested with *Xho*I. The restriction ends of the vector were partially filled-in with dTTP and dCTP forming 5'-TC-3' overhangs. This ensured the vector would not self-ligate. The central stuffer fragment was removed from the vector DNA to increase the probability of creating recombinant phage. Digestion of *A. variabilis* genomic DNA with *Sau*3AI and partially filling-in the resulting restriction ends with dGTP and dATP would create 5'-GA-3' overhangs. Donor DNA prepared in this way would ligate to lambda FIX® II but would not self-ligate. This would prevent the formation of multiple inserts which was important because the juxtapositioning of normally non-contiguous fragments within the *A. variabilis* genome would have created problems when characterising recombinant DNA molecules. Size fractionation of donor DNA was unnecessary because only

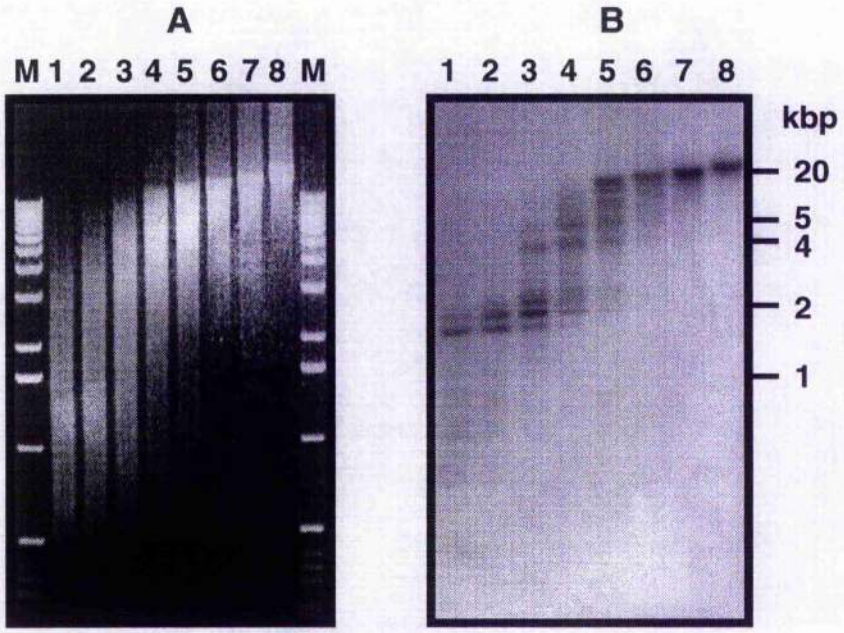
recombinants containing insert DNA between 9 and 23-kb in size would produce viable phage DNA able to be packaged into infectious phage particles. Phage containing insert DNA outside this size range would not be packaged efficiently.

5.4 Southern Hybridisation of the *Synechocystis aroA* Probe to *Sau*3AI Digested *Anabaena variabilis* Genomic DNA

Lambda FIX II clones DNA between 9 and 23-kb in size. In order to produce donor DNA in this size range, *A. variabilis* DNA was digested with *Sau*3AI. Southern blots of *Sau*3AI digested DNA were probed with the *Synechocystis aroA* gene to ensure the probe hybridised to target DNA in the appropriate size range.

The stock of genomic DNA used for this experiment was extracted from *A. variabilis* cells as in Section 4.4.1. A total of 10 µg of genomic DNA was digested with varying dilutions of *Sau*3AI indicated in Figure 5.2 (A) (Section 2.9.2). The digested DNA was electrophoresed through a 0.8% agarose gel from which Southern blots were made (Section 2.23). The 618-bp, *Bgl*II fragment of the *Synechocystis aroA* gene was prepared as in Section 4.3.3. 25 ng of this probe was radiolabelled (Sections 2.21 and 2.22) to a specific activity of 2 x 10^8 cpm/µg then hybridised to the Southern blots of *Sau*3AI digested *A. variabilis* DNA at 50 °C in 4 x SET (Section 2.24). The pre-determined optimal conditions for hybridisation of the *Synechocystis aroA* probe to *A. variabilis* DNA were used for this experiment (Section 4.3.2). After the membranes were washed from 2 x SET to 0.5 x SET at 50 °C and autoradiographed for 16 hours, the hybridisation signal shown in Figure 5.2 (B) was obtained. When genomic DNA was digested with 0.062 and 0.031 units *Sau*3AI per µg DNA, the majority of fragments were between 4 and 20-kb in size. The *Synechocystis aroA* probe hybridised strongly to a band 15 - 20-kb in size when the *A. variabilis* DNA was digested in this size range.

Figure 5.2 Hybridisation of the *Synechocystis aroA* probe to *Sau*3AI digested *Anabaena variabilis* genomic DNA. (A) Ethidium bromide stained agarose gel of 1 μg *A. variabilis* genomic DNA digested with the following amounts of *Sau*3AI: 2 units (lane 1), 1 unit (lane 2), 0.5 units (lane 3), 0.25 units (lane 4), 0.125 units (lane 5), 0.0625 units (lane 6), 0.031 units (lane 7) and no enzyme (lane 8). M = DNA markers: 1-kb DNA ladder from Gibco BRL. (B) Southern hybridisation of the *Synechocystis* 618-bp *aroA* probe to the *Sau*3AI digested *A. variabilis* genomic DNA shown in (A). Hybridisation and washing conditions are described in the text.


A
M 1 2 3 4 5 6 7 8 M
B
1 2 3 4 5 6 7 8
kbp
20
5
4
2
1

5.5 Construction of the *Anabaena variabilis* Genomic DNA Library in Lambda FIX® II

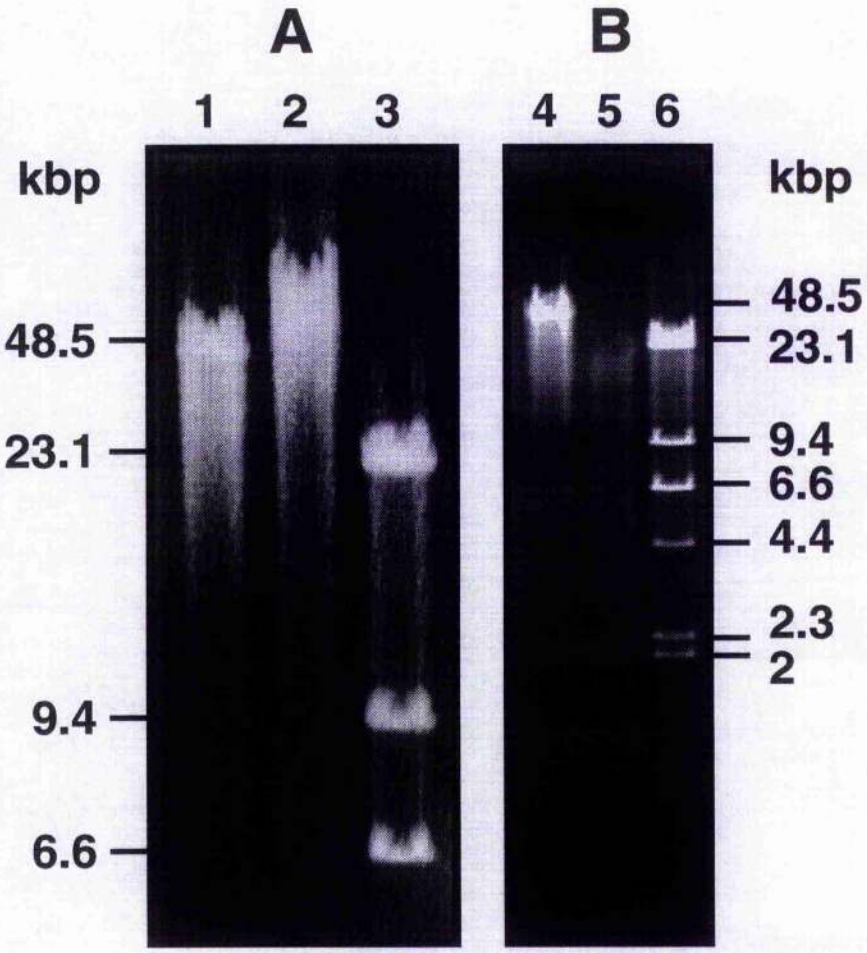
5.5.1 Preparation of *Anabaena variabilis* Genomic DNA for Ligation into Lambda FIX® II

Since lambda FIX II clones DNA between 9 and 23-kb in size, high molecular weight genomic DNA (at least 50-kb in size) had to be extracted from *A. variabilis* cells. Genomic DNA of a lower molecular weight was not suitable for library construction because it would have contained a large proportion of fragmented DNA with sheared ends unable to be cloned into the phage vector.

A. variabilis cells were grown and checked for contamination as in Section 2.3.1/2. 50 - 100 µg of genomic DNA was extracted from the cells using the procedure outlined in Section 2.7. Analysis by electrophoresis through a 0.3% agarose gel (Section 2.11) showed that the majority of the genomic DNA was over 50-kb in size (Figure 5.3 (A)). Although the yield of DNA was quite low, its size and purity (as measured in Section 2.4.2) was more than adequate for library construction. A number of problems were encountered when culturing *A. variabilis* cells for the extraction of genomic DNA and are discussed more fully in Section 5.8.1.

The *Sau*3AI digested genomic DNA used for the Southern blot experiment described in Section 5.4 indicated that 0.062 and 0.031 units *Sau*3AI per µg DNA generated restriction fragments the majority of which were between 4 and 20-kb in size. Pilot reactions were performed using the same concentrations of *Sau*3AI on the high molecular weight DNA to determine if similar sized fragments were produced (Section 2.9.2). Since the stock of high molecular weight DNA was very precious, only 0.5 µg DNA was used for each digestion. Electrophoresis of the digestion products through a 0.3% agarose gel (Section 2.11) showed that the majority of restriction fragments between 5 and 23-kb in size were generated using 0.031 units *Sau*3AI per µg DNA (Figure 5.3 (B)).

Figure 5.3 Preparation of *Anabaena variabilis* genomic DNA for phage library construction. DNA was analysed by electrophoresis through 0.3% agarose / ethidium bromide gels. (A) high molecular weight *A. variabilis* genomic DNA (lane 2) and (B) 0.5 µg *A. variabilis* genomic DNA digested with 0.015 *Sau*3AI (equivalent to 0.031 units *Sau*3AI per µg DNA) (lane 5). Lanes 1 and 4 contain uncut lambda DNA. Lanes 3 and 6 contain lambda DNA digested with *Hin*dIII.


A
B
1
2
3
4
5
6
kbp
48.5
23.1
9.4
6.6
kbp
48.5
23.1
9.4
6.6
4.4
2.3
2

The pilot reactions were scaled up to make a stock of donor DNA for cloning purposes. This was performed as in Section 2.18.1. 10 µg of the high molecular weight genomic DNA was digested with 0.031 units *Sau*3AI per µg DNA under the same conditions used for the pilot reactions (Section 2.9.2). In previous attempts at library construction, 50 µg of genomic DNA was used to make a stock of donor fragments. This amount was approximately the average yield of DNA obtained from the protocol used to extract high molecular weight genomic DNA from *A. variabilis* cells (Section 2.7). However, due to over-digestion of the DNA, the restriction fragments generated were too small for cloning into λ FIX II. In case such problems re-occurred, only 10 µg (a fraction of the stock of high molecular weight genomic DNA) was used to generate the donor fragments. Analysis of the digestion products by agarose gel electrophoresis showed that fragments 5 - 23-kb had been generated. The digested DNA was purified by phenol / chloroform extraction and ethanol precipitated as in Section 2.18.1 then dissolved in TE buffer.

The restriction ends of the *Sau*3AI digested DNA were partially filled-in with dATP and dGTP (Section 2.18.1) to create 5'-GA-3' overhangs complementary to the restriction ends of λ FIX II (Section 5.3). The DNA was then purified and re-dissolved in TE buffer.

Since only 10 µg of genomic DNA was used to prepare the donor fragments for library construction, a good recovery of DNA after purification from the *Sau*3AI digestion and partial fill-in reactions was essential. This was achieved when the DNA was ethanol precipitated at -20 ^{o}C for 12 - 16 hours and centrifuged at 4 ^{o}C for 30 minutes.

After purification, the *Sau*3AI digested, partially filled-in DNA was quantitated using a DNA fluorimeter type TKO 100 (Section 2.18.2). Under the correct assay conditions this machine could accurately measure as little as 10 ng/ml DNA, so was ideal for quantitating the small amount of donor DNA used in these experiments. The concentration of donor DNA was measured as 525 ng per µl. This was verified by agarose gel electrophoresis which also showed that the insert DNA had not been further

degraded during the partial fill-in reaction or the purification procedures. A total of 5.25 µg of donor DNA had been recovered from these manipulations which was sufficient for library construction in λ FIX II.

5.5.2 Ligation of *Sau*3AI Cut, Partially Filled-In *Anabaena variabilis* DNA to Lambda FIX® II

When performing ligations, the optimal conditions for the formation of recombinant molecules is usually determined empirically. For phage library construction, ligation conditions favouring the formation of long concatenated DNA molecules that are packaged more efficiently into infectious particles are required.

The amount of *A. variabilis* DNA available for library construction was very limiting, so a large number of test ligations could not be carried out. Ligations involving replacement vectors, such as λ FIX II, are usually most efficient when roughly equal molarities of vector and donor DNA are present in the reaction mixture (Kaiser and Murray, 1985). Since the size of genomic DNA fragments generated by *Sau*3AI was between 5 - 23-kb, the average insert size of the *A. variabilis* donor DNA was assumed to be 14-kb. Therefore, 0.4 µg donor DNA was ligated to 1 µg λ FIX II (Section 2.18.3). A control ligation was also performed using a test insert supplied by Stratagene, Cambridge, England. pMF is a 12-kb DNA fragment with *Bam*H1 restriction ends partially filled-in to leave 5'-GA-3' overhangs. 0.3 µg of pMF was ligated 1 µg λ FIX II as recommended by the manufacturers (Section 2.18.3).

0.28 µg and 0.26 µg of the products obtained from the *A. variabilis* and control ligation reactions were packaged into infectious phage particles (Section 2.18.4). The packaging extracts, (Stratagene Limited, Cambridge, England) were prepared from *E. coli* cells lacking all known restriction systems. As discussed later in Section 6.6, *A. variabilis* DNA could act as a substrate for digestion by the various restriction systems present in *E. coli*. Therefore, restriction minus packaging extracts were used to improve the packaging efficiency and the representation of *A. variabilis* sequences in the phage library.

To determine the efficiency of the *A. variabilis* and control ligation and packaging reactions, the titres of the packaged phage were determined (Section 2.19). *E. Coli*® XL1-Blue MRA (P2) (Table 2.1) was used as a host strain for phage infection. XL1-Blue MRA (P2) contains a P2 phage lysogen that supports the growth of recombinant λ FIX II by spi (sensitive to P2 infection) selection. This reduces the occurrence of non-recombinant phage. Section 2.19.3 describes spi selection in more detail. The *E. Coli*® strain XL1-Blue MRA (P2), like the packaging extracts, lacks all known restriction systems which should enhance the stability of *A. variabilis* sequences in this heterologous host (Section 6.6).

XL1-Blue MRA (P2) cells were prepared for infection with varying dilutions of packaged phage DNA from the *A. variabilis* and control reactions (Sections 2.19.1/2). The cell mixtures were then grown on NZY medium at 37 °C for 8 hours. The number of plaques were counted and the titres were calculated (using the equation in Section 2.19.2) to be 7.4 x 10^4 and 1 x 10^6 PFU/μg DNA for the *A. variabilis* and control reactions respectively.

5.5.3 Amplification of the *Anabaena variabilis* Genomic DNA Library in Lambda FIX® II

The number of infectious particles obtained from packaging 0.28 μg of the *A. variabilis* library was 20,800. Assuming the average insert size was 14-kb (Section 5.5.2), 1840 recombinant clones would have to be screened to have a 99% probability of isolating one particular sequence (Clarke and Carbon, 1976 and Section 4.4.2). So there should have been approximately 11 copies of the *A. variabilis* genome present in the recombinant phage obtained from one packaging reaction. Nevertheless, it was deemed prudent to amplify the *A. variabilis* library to make a high titre stock of recombinant phage because previous heterologous hybridisation experiments (Chapter 4) had shown that a large number of clones would have to be screened to maximise the possibility of isolating the *aro*A gene. Amplification also eliminated the need for repeated ligation and packaging reactions and saved the time and effort required to

prepare *A. variabilis* donor DNA needed for phage library construction. The high titre phage library of *A. variabilis* DNA could also act as a source from which a variety of different sequences could be isolated.

In general, growth of recombinant phage DNA is usually unaffected by the sequence content of the insert DNA but in a minority of cases differential growth of recombinant phage can occur. If this were to happen during amplification, the sequence content of the library would be affected. This would increase the number of recombinant clones to be screened to give a 99% probability of isolating the sequence of interest (Kaiser and Murray, 1985). Sequence misrepresentation is exacerbated by the competitive growth of recombinants in liquid culture or at high plaque densities. So when amplifying the *A. variabilis* library, phage were plated at relatively low densities on agar plates as in Section 2.20. The titre of the *A. variabilis* library after the amplification procedure was 1 x 10^{10} PFU/ml (Section 2.19.2). This was a more than adequate stock of recombinant phage for screening purposes.

5.6 The Quality of the Amplified *Anabaena variabilis* Genomic DNA Library in Lambda FIX® II

The quality of the amplified *A. variabilis* library was assessed to establish if it was sufficiently large and complex enough to contain any one particular sequence. To accomplish this, a number of experiments were performed.

A random selection of clones were purified and restriction mapped to give an indication of the distribution of *A. variabilis* sequences in the library. *E. Coli*® XL1-Blue MRA (P2) was infected with approximately 1000 PFU from the amplified *A. variabilis* library. The cell mixture was grown on 15 cm NZY plates (Sections 2.19.1/2). Six well isolated plaques were picked at random and the phage DNA was purified as in Section 2.27.4. 500 ng of each phage clone was digested with *Not*I and *Hin*dIII (Section 2.9.4). Digestion with *Not*I at the polycloning site of λ FIX II (Figure 5.1) released the insert DNA from the vector. *Hin*dIII lacked sites within the vector arms but cut the *A. variabilis* insert DNA into smaller pieces. This allowed better

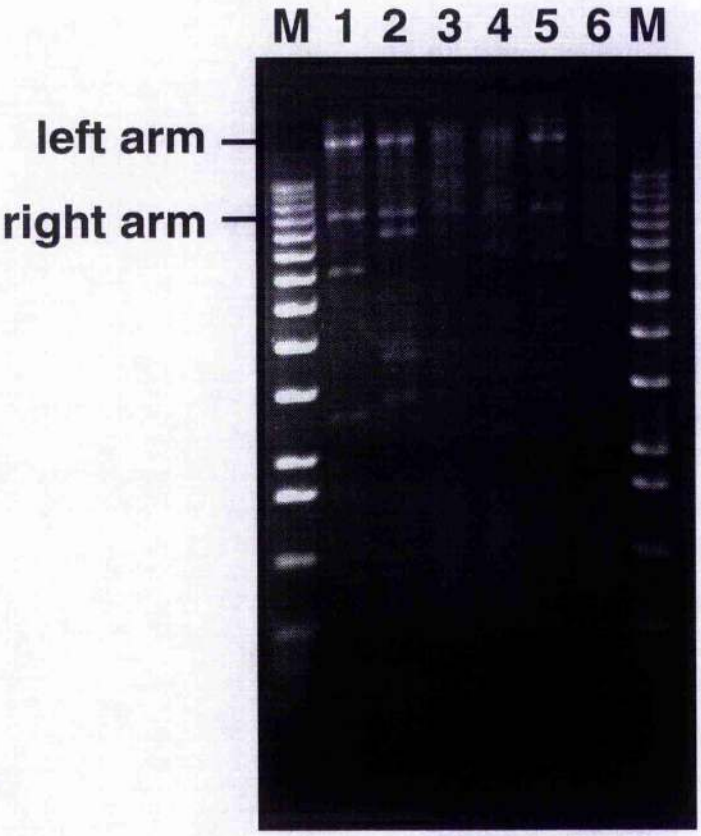
separation of the cloned DNA from the λ FIX II arms when the digested phage clones were electrophoresed through a 0.7% agarose gel (Section 2.11), as shown in Figure 5.4. Analysis of the digestion products showed that each clone had a unique restriction map which suggested that the library contained an array of different *A. variabilis* sequences. This experiment also showed that each clone contained insert DNA indicating that spi selection (Section 2.19.3) for recombinant phage must have been working efficiently.

The average size of insert DNA from the six purified clones was 14.4-kb. This information was used to estimate the number of clones that had an arbitrary probability of representing any particular sequence (Clarke and Carbon, 1976 and Section 4.4.2). Since the size of the *A. variabilis* genome is 5.4 x 10^6-bp, 1769 individual recombinant clones were calculated to have a 99% probability of containing any specific sequence.

To further assess the sequence distribution, the newly constructed library was screened with a homologous probe. The *vnf*DGK operon, which encodes the vanadium dependent nitrogenase of *Anabaena variabilis* ATCC 29413, was used for this purpose (Thiel, 1993 and Section 4.5.2). The *vnf*DGK genes are carried on a 3.3-kb *Cla*I fragment of *A. variabilis* genomic DNA in the pUC119 clone, pRSI, which was the kind gift of Dr. Teresa Thiel, University of Missouri, St. Louis, Missouri, U.S.A.

After purification from pRSI (Section 4.5.1), 100 ng of the *vnf*DGK probe was radiolabelled (Sections 2.21 and 2.22) to a specific activity of 5 x 10^8 cpm/μg. *E. Coli*® XL1-Blue MRA (P2) was infected with 50,000 PFU from the amplified *A. variabilis* library (Sections 2.27.1). The cell / phage mixture was grown on a 15 cm NZY plate until the resulting plaques were of a small, uniform size. The plaques were replicated onto 2 nylon membrane discs. The DNA on these duplicate membranes was

Figure 5.4 Restriction analysis of randomly selected clones from the amplified *Anabaena variabilis* library in lambda FIX II. Each clone (numbered 1 - 6) was digested with *Not*I and *Hind*III and the reaction products were analysed by electrophoresis through a 0.7% agarose - ethidium bromide gel. M = DNA markers: 1-kb DNA ladder (supplied by Gibco BRL). DNA bands corresponding to vector arms are indicated.


M 1 2 3 4 5 6 M
left arm
right arm

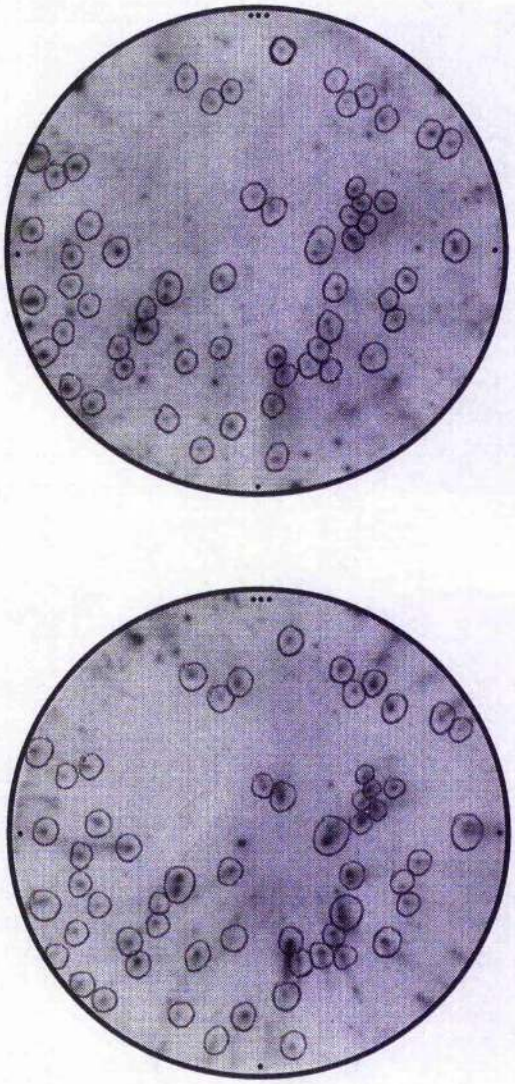
denatured and immobilised as described in Section 2.27.1. The radiolabelled *vnf*DGK probe was hybridised to the library at 65 °C in 4 x SET (Section 2.27.2). After 24 hours, the membranes were washed stepwise from 2 x SET to 0.1 x SET at 65 °C then autoradiographed overnight. The strong, duplicate hybridisation signals seen in Figure 5.5 were produced.

The Clarke and Carbon equation predicted that 1769 individual recombinant clones would have to be screened to have a 99% probability of isolating any particular sequence from the *A. variabilis* library. The *vnf*DGK probe hybridised to 64 clones in every 50,000. Therefore, this sequence occurred in approximately two out of every 1769 clones. This indicated that the library was probably large enough to contain any one sequence.

5.7 The Heterologous Probes Used to Screen the Amplified *Anabaena variabilis* Genomic DNA Library in Lambda FIX®II

As discussed in Chapter 4, the *Synechocystis aro*A gene hybridised strongly to specific fragments of restriction digested *A. variabilis* genomic DNA but failed to isolate the clone of interest from the pBluescript library of *A. variabilis* sequences. This was due to the high level of background noise caused by non-specific hybridisation of the probe to the genetic material of the host strain harbouring the library. Both the newly constructed λ FIX II library and the pBluescript library have been screened using the homologous *vnf*DGK probe under the same hybridisation conditions (Sections 5.6 and 4.5.2 respectively). The resulting hybridisation signals were compared and showed that plaque hybridisation was more sensitive than colony hybridisation (Figures 5.5 and 4.3). This was due to the decreased level of background hybridisation of the probe to the genetic material of the host strain harbouring the phage library. For this reason, it was deemed prudent to screen the λ FIX II library with the *Synechocystis aro*A probe in an attempt to clone the *A. variabilis* EPSP synthase gene.

Figure 5.5 Autoradiogram resulting from hybridisation of the *Anabaena variabilis vnf*DGK probe to replicate membranes carrying approximately 50,000 clones of the *A. variabilis* genomic DNA library in lambda FIX II (grown in *E. coli*® XL1-Blue MRA (P2)). The hybridisation and washing conditions are outlined in the text. The homologous probe hybridised to 64 clones out of approximately 50,000. Positively hybridising, duplicate signals have been circled.



In order to maximise the possibility of isolating a positively hybridising clone from the *A. variabilis* library, *aro*A sequences derived from different sources were used as probes. As mentioned in Section 5.1, *E. coli* genes and pea cDNA sequences have been used to isolate various *Anabaena* genes from phage libraries. Both the *E. coli* K12 and pea *aro*A sequences have been cloned (Duncan and Coggins, 1984 and Granger, 1989) in this laboratory and were available to use as probes for screening purposes. Section 5.8.2 discusses in more detail the reason for using these particular *aro*A genes as probes.

One drawback when using the *E. coli aro*A probe to screen the *A. variabilis* library housed in the *aro*A$^+$ host, *E. Coli*® XL1-Blue MRA (P2), was the high background hybridisation due to the extensive homology shared between the probe and the residual host DNA in the plaques. The background noise could have been eliminated using a *aro*A$^-$ deletion mutant of *E. coli*. However, the only *aro*A$^-$ mutant strains available contained point mutations in the essential gene. If these mutants were used as host strains for the *A. variabilis* library, the *E. coli aro*A probe could still hybridise to the unaltered *aro*A sequence of the host. Since an *aro*A$^-$ deletion mutant was not available, *E. Coli*® XL1-Blue MRA (P2) was used as the host strain for these experiments. It was thought that if the *E. coli* probe hybridised to a clone containing the *A. variabilis aro*A sequence, a much stronger signal would result compared to that produced from the background hybridisation.

5.7.1 Preparation of the *Synechocystis*, *E. coli* and Pea *aro*A Probes

The 618-bp, *Bgl*II fragment of the *Synechocystis aro*A gene was prepared as in Section 4.3.1/3 and used for screening purposes for the same reasons given in Section 4.6.

The *E. coli aro*A gene is carried on the plasmid, pKD501, a pAT153 derivative containing a 4.6-kb *Pst*I fragment of *E. coli* K12 genomic DNA (Duncan and Coggins, 1984). A 766-bp fragment of the *aro*A gene was amplified from pKD501 by

PCR using primers 1 and 2 (Table 3.1 and Figure 3.2) as in Section 2.14.1. PCR was performed using 100 ng plasmid DNA, 1.5 mM $MgCl_2$, 200 µM dNTP and 40 pmoles of each primer. The incubation conditions were as follows: denaturation step at 94 °C for 5 minutes, cycle step at 94 °C for 1 minute, 50 °C for 1 minute and 72 °C for 1.5 minutes for 30 cycles and an extension step at 72 °C for 5 minutes completed the reaction. The amplified 766-bp product was purified from a l.m.p. agarose gel as in Section 2.12.3.

pSG5161 is a pUC18 derivative containing the pea *aro*A sequence on an *Eco*RI cDNA fragment 1.5-kb in length (Granger, 1989). The enzymes, *Pst*I and *Bam*HI were used to cut out a 921-bp segment of the *aro*A sequence (Section 2.9.4). The digestion products were electrophoresed through a l.m.p. agarose gel from which the 921-bp pea *aro*A fragment was purified (Section 2.12.3).

The *Synechocystis*, *E. coli* and pea *aro*A probes represented approximately 70%, 60% and 66% respectively of the coding sequences. The reasons for using large segments of the *aro*A genes as probes are the same as those stated in Section 4.3.1.

5.7.2 Screening the Amplified *Anabaena variabilis* Genomic DNA Library in Lambda FIX® II with the *Synechocystis*, *E. coli* and Pea *aro*A Probes

E. Coli® XL1-Blue MRA (P2) was infected with the amplified *A. variabilis* library and approximately 50,000 PFU were grown on each of three 15 cm NZY plates (Section 2.27.1) until small plaques of uniform size appeared. The plaques from each plate were replicated onto nylon membranes where the DNA was then denatured and immobilised. Duplicate filters were made for the identification of authentic hybridisation signals. 100 ng of each of the *Synechocystis*, *E. coli* and pea *aro*A probes were radiolabelled (Sections 2.21 and 2.22) to specific activities of 6.0×10^8, 5.9×10^8 and 5.3×10^8 cpm/µg DNA. The three probes were hybridised to separate sets of duplicate filters containing 50,000 clones each at 50 °C in 4 x SET (Section

2.27.2). The membranes were sequentially washed from 4 x SET at room temperature to 2 x SET at 50 °C and autoradiographed after each wash step.

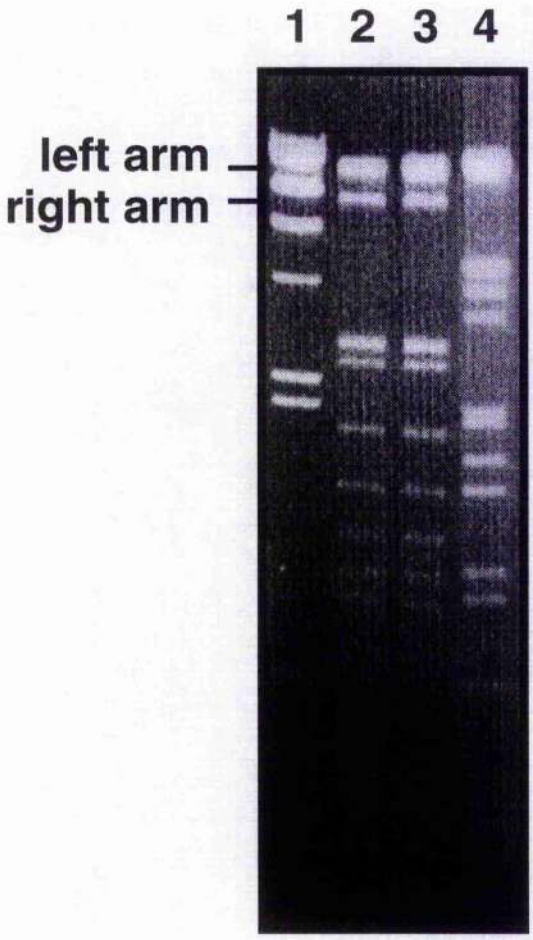
Hybridisation of the pea *aro*A probe to the *A. variabilis* library did not produce a convincing signal. Nine duplicate signals resulted from hybridisation of the *Synechocystis* probe to the library whilst the *E. coli* probe produced ten putative positive signals even after washing the filters in 0.5 x SET at 50 °C. The proportion of clones eliciting duplicate signals after hybridisation with the *Synechocystis* and *E. coli* probes was considerably smaller than the predicted library size (Section 5.6) but worth further investigation.

The clones corresponding to the duplicate signals were picked and re-screened with the appropriate heterologous probe under the same hybridisation conditions (Section 2.27.3). All the clones screened with the *Synechocystis* probe were false positives. However, two out of the ten clones detected with the *E. coli* probe were purified to homogeneity over two more rounds of screening. After the tertiary screen, phage DNA was purified from a selection of well isolated plaques from a number of different plates (Section 2.27.4). 500 ng of the purified phage DNA (λ E5-2 and λ E5-4) was digested with *Not*I and *Hin*dIII to free the insert DNA from the vector arms (Section 2.9.4). Agarose gel electrophoresis showed that each clone had the same restriction map and had an insert size of approximately 15-kb (Figure 5.6). This signified that the *E. coli* probe had isolated identical clones from the *A. variabilis* library.

5.7.3 Analysis of the Purified Recombinant Clones Isolated from the *Anabaena variabilis* Library

To ensure the *E. coli aro*A probe hybridised specifically to the insert DNA of λ E5-2 and λ E5-4, a Southern blot of the gel in Figure 5.6 was carried out (Section 2.23). 25 ng of the *E. coli aro*A probe was radiolabelled (Sections 2.21 and 2.22) to a specific activity of 6 x 10^8 cpm / μg DNA and hybridised to the Southern blot of *Not*I, *Hin*dIII digested phage DNA at 50 °C in 4 x SET (Section 2.24). The membranes

Figure 5.6 Restriction analysis of purified clones isolated from the *Anabaena variabilis* library using the *E. coli aro*A probe. Ethidium bromide stained gel of clones λ E5-2 (lane 2) and λ E5-4 (lane 3) digested with *Not*I and *Hin*dIII. Lanes 1 and 4 are DNA markers: lambda DNA digested with *Hin*dIII and *Eco*RI / *Hin*dIII respectively. DNA bands corresponding to vector arms are indicated.


1 2 3 4
left arm
right arm

were washed from 4 x SET at room temperature to 0.5 x SET at 50 ^{o}C and autoradiographed after each wash step. However, even after repetition of this experiment, the probe did not hybridise to the insert DNA of the recombinant clones. This suggested that the isolated recombinant clones did not contain insert sequence complementary to the *E. coli* probe which contradicted the result obtained from the screening experiment.

To resolve this issue, separate preparations of *E. Coli*® XL1-Blue MRA (P2) could have been infected with λE5-2, λE5-4 and a negative control phage lacking *aro*A sequence (such as lambda FIX II containing the test insert pMF (Section 5.5.2)). If the *E. coli* probe hybridised to plaques containing λE5-2 and λE5-4 but not the negative control phage, then the purified clones would harbour insert DNA complementary to the *E. coli aro*A sequence.

A possible reason why the *E. coli aro*A probe could have isolated two identical clones from the *A. variabilis* library without specifically binding to the insert DNA of either recombinant is not clear. The Southern hybridisation experiment showed that the probe did not bind to the λ FIX II arms of either clone, so it may have hybridised to the genetic material of the host DNA. As discussed in Section 5.5.3, some recombinant phage are capable of replicating more efficiently than others. If the purified clones were able to replicate faster than the vast majority of other recombinants in the library, it would have given rise to a larger plaque. A more intense hybridisation signal would have resulted if the *E. coli* probe bound to the larger quantity of host genetic material present in these plaques. This could have been responsible for the false positive result. Yet during the screening procedure described in Section 5.7.2 there did not appear to be any significant difference in plaque size between λ E5-2, λ E5-4 and the other clones of the library.

5.8 Discussion

5.8.1 Overview of Phage Library Construction and Screening Procedures

The work in this chapter describes the construction of a library of *A. variabilis* genomic DNA in the replacement vector, λ FIX II (Section 5.5). As mentioned in Section 5.5.1, a number of problems were encountered with regard to DNA extraction from *A. variabilis* cells. *A. variabilis* cells were cultured to their late log phase of growth in several 100 ml volumes of BG11o medium in 250 ml conical flasks (Section 2.3.1). Due to the ease of handling of such small flasks, a good aseptic technique could be performed. This allowed uncontaminated *A. variabilis* cultures to be grown which was essential for extracting DNA for gene cloning experiments. 50 - 100 μg of good quality genomic DNA approximately 50-kb in size was produced using the gentle extraction procedures outlined in Section 2.7. The yield of DNA could have been much improved if a larger volume of cells (for example 5 - 10 litres) was grown. However, the facilities available made growing cyanobacterial cultures in flasks 10 - 20 litres in size difficult. Contamination was a great problem since it was very awkward to handle large flasks in flow hoods that could not accommodate their size. A higher yield was achieved when DNA was purified from cells that had reached their stationary phase of growth. But the genomic DNA from these older cells was more degraded and, therefore, unsuitable for phage library construction. As a result of the problems outlined above and those described in Section 5.5.1, a small quantity of good quality *A. variabilis* DNA was used for phage library construction.

Screening experiments were performed using the homologous *vnf*DGK probe to evaluate the quality of the newly constructed *A. variabilis* library in lambda FIX II. This work showed that the library was of a satisfactory size and should have a good possibility of containing the *aro*A gene (Section 5.6). EPSP synthase genes from *Synechocystis* sp. PCC 6803, *E. coli* K12 and pea were used as probes to screen the *A. variabilis* library for the clone of interest. The pea probe did not hybridise to any

clone in the library (Section 5.7.2). Duplicate hybridisation signals were produced when the library was screened with the *Synechocystis* probe but these were shown to be false positives upon further analysis. When the library was screened with the *E. coli* probe, two out of ten recombinant clones were purified to homogeneity. Both of the isolated clones had identical restriction maps (Section 5.7.3) but the *E. coli* probe did not hybridise to the recombinant DNA when Southern blots were performed. This signified that the *A. variabilis aro*A sequence had not been isolated. The possible reason(s) why the *E. coli* probe was able to isolate identical clones from the variety of different recombinants present in the *A. variabilis* library without hybridising to the insert DNA is unclear, but is discussed in Section 5.7.3. Although the isolated clones did not contain the *A. variabilis aro*A gene, the insert DNA sequence could contain some valuable information. Deduction of the cloned DNA sequence and its subsequent comparison to those sequences logged in the Wisconsin GCG data base (Section 2.30) would perhaps result in the tentative identification of genes not yet isolated from *A. variabilis*.

5.8.2 Justification for Using the *Synechocystis*, *E. coli* and Pea *aro*A Genes as Heterologous Probes

The use of heterologous probes to screen libraries of *A. variabilis* DNA was a major strategy especially after the *aro*A gene from the unicellular cyanobacterium, *Synechocystis* sp. PCC 6803 was isolated (Mayes *et al.*, 1993). It was presumed that the homology shared between the *Synechocystis* and *A. variabilis* genes would be high since both organisms are cyanobacteria. Southern blot experiments (Section 4.3.3) and comparison of other nucleotide sequences isolated from both cyanobacterial species (Section 4.7) seemed to support this assumption. Additionally, genes from unicellular cyanobacteria had already been used as probes to isolate sequences from various *Anabaena* species (Section 4.1 and 5.1). This is why both screening strategies using the phagemid and phage libraries (Chapters 4 and 5) involved the *Synechocystis aro*A probe.

The *E. coli* and pea *aro*A probes were chosen for heterologous screening after examining the homologies shared between the EPSP synthase sequences already cloned. The similarities between many of the EPSP synthases have been examined (Griffin and Griffin, 1991; Section 1.7.4). This analysis showed that the enzyme sequences could be assembled into four groupings: Gram negative bacteria, Gram positive bacteria, higher plants and fungi. The *Synechocystis* enzyme sequence is most homologous to the Gram positive bacteria (Mayes *et al.*, 1993). So when the phage library was screened (Chapter 5), *aro*A genes from *E. coli* and pea were also used as probes to represent EPSP synthase sequences from the Gram negative bacteria and higher plants groupings. Additionally, *E. coli* and pea sequences had previously been used as probes to isolate *Anabaena* genes from DNA libraries. The glutamine synthetase (*gln*A) gene of *E. coli* and the cDNA encoding the small sub unit of ribulose-1,5-bisphosphate carboxylase (*rbc*S) of pea have previously been used to isolate the equivalent genes from various *Anabaena* species (Fisher *et al.*, 1981; Nierzwicki-Bauer *et al.*, 1984). The cyanobacterial *rbc*S gene shares 55% homology with its respective probe (refer to previous reference). On comparison of the nucleotide sequences containing the *gln*A gene using the, "gap", function of the Wisconsin GCG package (Section 2.30), the *Anabaena* sequence was found to be 54% identical to the *E. coli* sequence. If the *E. coli* and pea *aro*A probes shared this homology with the *A. variabilis* target sequence, it would be possible to isolate the gene of interest by screening a DNA library using the hybridisation conditions described in the text (Section 5.7.2). This supposition was based on information taken from the work of Sakamoto *et al.* (1994) who had isolated the *A. variabilis* fatty acid desaturase gene from a phage library using a DNA probe 55% homologous to the target sequence and employing hybridisation conditions very similar to that described in Section 5.7.2.

Since no examples could be found in the literature describing the isolation of cyanobacterial genes using fungal or yeast probes, an *aro*A probe derived from these organisms was not used to screen the *A. variabilis* library.

5.8.3 Reasons Why the *Anabaena variabilis aro*A Gene was not Isolated from the Phage Library

Screening the *A. variabilis* DNA library using the *E. coli* and pea *aro*A probes was an entirely speculative approach and was carried out to maximise the possibility of isolating the cyanobacterial EPSP synthase gene. A Southern blot experiment demonstrating that these heterologous *aro*A probes could bind to *A. variabilis* DNA under the hybridisation conditions used for screening purposes were not performed. This experiment would have determined if these *E. coli* and pea probes could hybridise specifically to *A. variabilis* DNA and would have been the more definitive route to take.

Southern blot experiments showed that the *Synechocystis aro*A probe bound strongly to specific fragments of restriction digested *A. variabilis* genomic DNA (Sections 4.3.1/3) using the optimal hybridisation conditions described in Section 4.3.2. Under the same conditions, the cyanobacterial probe was unable to isolate the *aro*A clone from the *A. variabilis* library constructed in λ FIX II (Section 5.7.2). This problem did not arise from a high level of background hybridisation as was experienced with colony hybridisation (Section 4.6). Indeed, the *Synechocystis* probe did not even hybridise to the expected number of recombinant clones (Section 5.7.2). These experiments seemed to indicate that the phage library was not large and complex enough to contain the *A. variabilis aro*A gene, despite the results described in Section 5.6 that indicated the contrary.

Chapter 6

Screening an *Anabaena variabilis* Genomic DNA Library by Phenotypic Complementation

6.1 Complementation

When attempting to clone any gene of the shikimate pathway by phenotypic complementation, the well characterised *aro*$^-$ auxotrophic mutants of *E. coli* can be used (Pittard and Wallace, 1966). These mutants lack one essential activity in the shikimate pathway and so cannot grow on the minimal medium (deficient in aromatic amino acids) which can support the growth of wild type *E. coli*. Growth of an *aro*$^-$ *E. coli* strain on minimal medium requires the expression of an appropriate functional enzyme to complement the mutation. This provides a method of direct selection for the gene of interest. The required gene can be isolated by transforming the *E. coli* mutant with a library of genomic DNA and selecting for transformed cells able to grow on minimal medium. Plasmid DNA is purified from these cells then re-introduced into the mutant host. This confirms if the functional protein expressed by the cloned DNA is responsible for the phenotypic change from auxotrophy to prototrophy. A number of shikimate pathway genes have been cloned using this strategy, for example, the EPSP synthase and 3-dehydroquinate synthase genes of *E. coli* (Duncan and Coggins, 1984) and the chorismate synthase gene of *Staphylococcus aureus* (Horsburgh, 1995). Some cyanobacterial genes have also been isolated by complementation in *E. coli* including the *rec*A gene of *Anabaena variabilis* ATCC 29413 (Owttrim and Coleman, 1987), the $NADP^+$-isocitrate dehydrogenase gene of *Anabaena* sp. PCC 7120 (Muro-Pastor and Florencio, 1994) and the adenylate cyclase gene of *Anabaena cylindrica* (Katayama *et al.*, 1995).

6.2 Strategy for Isolating the *A. variabilis aro*A Gene by Complementation

Cloning the *A. variabilis aro*A gene by phenotypic complementation required the use of the *aro*A auxotrophic mutant, *E. coli* AB2829 (Pittard and Wallace, 1966). An *A. variabilis* genomic DNA library was constructed in pBluescript SK- for transformation into *E. coli* AB2829. Plasmid DNA was purified from transformants able to grow on minimal medium containing ampicillin and back

transformed into the mutant host. This was to confirm that growth of the mutant was due to the expression of an active EPSP synthase from the cloned *A. variabilis* DNA.

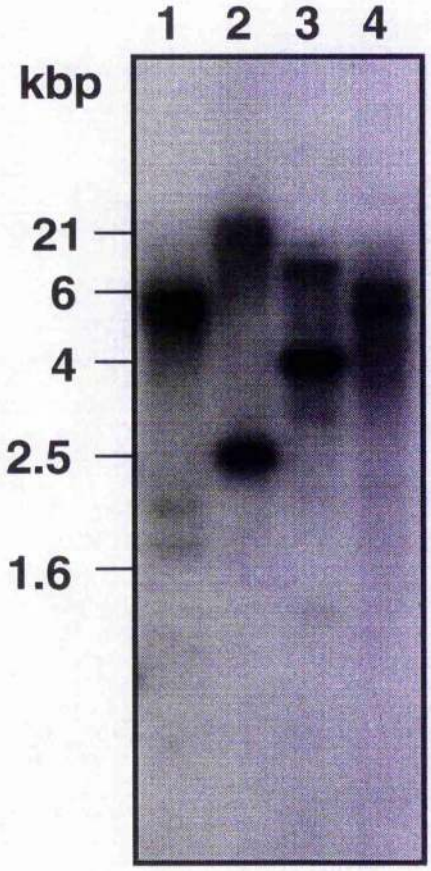
6.3 Construction of an *Anabaena variabilis* Genomic DNA Library in pBluescript SK- for Complementation

6.3.1 Preparation of *A. variabilis* Genomic DNA for Library Construction

Expression of a functional EPSP synthase was essential to clone the *A. variabilis aro*A gene by complementation in the *aro*A mutant *E. coli* AB2829. The *A. variabilis* library constructed in lambda FIX II was, therefore, unsuitable for complementation experiments. Growth of lambda FIX II clones in *E. coli* results in cell death because this phage vector exclusively enters the lytic growth cycle. In general, an important consideration when constructing libraries for complementation experiments was the choice of restriction enzyme used for generating donor fragments. Cleavage of the *aro*A gene would, in most cases, produce an inactive EPSP synthase unable to complement the *aro*A mutation of *E. coli* AB2829.

Southern hybridisations of the *Synechocystis aro*A probe (Section 4.3) to restriction digests of *A. variabilis* genomic DNA indicated the size and number of fragments upon which the *aro*A gene was reputedly located. The pattern of hybridisation obtained from *Hin*dIII digests of *A. variabilis* genomic DNA, using the 927, 618 and the 275-bp *Synechocystis aro*A gene fragments as probes, suggested the *aro*A gene had an internal restriction site for this enzyme (Section 4.3.3). Additionally, hybridisation of the *Synechocystis* 618-bp *aro*A probe to *A. variabilis* DNA, digested with *Eco*RV, *Acc*I and *Cla*I, indicated that the *aro*A gene could be carried on more than one restriction fragment (Figure 6.1). So genomic DNA digested with these restriction enzymes was considered unsuitable for making a library for complementation experiments. However, hybridisation of the 927, 618 and

Figure 6.1 Southern hybridisation of the *Synechocystis* 618-bp *aro*A probe to *Anabaena variabilis* genomic DNA. Genomic DNA was digested with *Eco*RV (lane 1), *Eco*RI (lane 2), *Acc*I (lane 3) and *Cla*I (lane 4). Experimental conditions are described in the text.


1 2 3 4
kbp
21
6
4
2.5
1.6

275-bp *Synechocystis aro*A gene fragments to *A. variabilis* DNA digested with *Eco*RI signified that the entire *aro*A gene could be located on a single fragment, 2.5-kb in size (Figure 4.2). This was entirely feasible since the average size of all bacterial *aro*A genes and plant cDNA sequences so far isolated is approximately 1.35-kb. So *A. variabilis* genomic DNA digested with *Eco*RI was used for library construction.

An uncontaminated stock of *A. variabilis* genomic DNA, prepared as described in Section 4.4.1, was used for construction of the library required for complementation. 5 µg of this genomic stock was digested with *Eco*RI, size fractionated from 1 to 10-kb and purified (as in Section 2.16). After quantitation (Section 2.4.2), the concentration of the digested DNA was adjusted to 100 ng / µl with TE buffer.

6.3.2 Ligation of Size Fractionated *Eco*RI Digested *Anabaena variabilis* DNA into pBluescript SK-

100 ng of *Eco*RI digested, size fractionated *A. variabilis* genomic DNA was ligated to 100 ng of *Eco*RI cut, dephosphorylated pBluescript SK-. 50 ng of the ligation products was transformed into *E. coli* AB2829 (Section 2.16). To determine the transformation efficiency, dilutions of the cell mixture were plated onto LB plates containing ampicillin to select for plasmid transformed cells. 50 ng of the *Eco*RI ligation yielded 1000 to 1200 ampicillin resistant transformants.

6.3.3 Quality of the *A. variabilis Eco*RI Genomic DNA Library

The quality of the *A. variabilis Eco*RI library was checked. This was to ensure the library represented the donor DNA used for its construction. *E. coli* AB2829 was transformed with the cyanobacterial library and grown on LB / ampicillin medium (Section 2.16). Plasmid DNA was purified from a random selection of 12 *E. coli* AB2829 transformants (Section 2.5.1), digested with *Eco*RI (Section 2.9.3) and analysed by agarose gel electrophoresis (Section 2.11). This

indicated that 72% of transformants harboured recombinant clones with an average insert size of 3.42-kb.

To give an indication of the number of clones to be screened, the size of the *A. variabilis Eco*RI library that had an arbitrary probability of containing a particular sequence was calculated using the Clarke and Carbon equation (1976 and Section 4.4.2). Since the genome size of *A. variabilis* is 5.4 x 10^6-bp (Herrero *et al.*, 1984) and the average insert size of the library was 3.42-kb, 7301 individual clones would have to be screened to give a 99% probability of isolating one particular sequence. This figure was increased by 28% to 9345, to take account of the number of non-recombinant clones present in the library. The actual size of the library was likely to be smaller since the donor DNA had been size fractionated. Nevertheless, the figure calculated by the Clarke and Carbon equation (1976) was used as a guide for screening purposes.

Only 1000 to 1200 ampicillin resistant colonies were obtained from transformations of 50 ng DNA from the *A. variabilis Eco*RI library into *E. coli* AB2829 (Section 6.3.2). Therefore, multiple transformations were performed to allow the appropriate number of clones to be screened so that there was a 99% chance of finding an *aro*A clone.

6.4 Complementation of the Auxotrophic Mutant, *E. coli* AB2829, with the Cloned EPSP Synthase Gene of *E. coli* K12

The stability of the *aro*A mutation of *E. coli* AB2829 was essential to isolate the *A. variabilis* EPSP synthase gene by phenotypic complementation. To test this stability, the auxotrophic host was transformed with pKD501 (Section 2.17.1b,2,3b), a pAT153 derivative carrying the *aro*A gene of *E. coli* K12 on a 4.6-kb *Pst*I genomic fragment (Duncan and Coggins, 1984). The cell mixture was plated onto minimal medium containing ampicillin which established if cells transformed with plasmid DNA expressed an active EPSP synthase able to complement the *aro*A mutation of the host. Untransformed *E. coli* AB2829 cells were cultured on minimal medium

supplemented with aromatic amino acids (Section 2.2d) to determine if the end products of the shikimate pathway supported the growth of the mutant. As a negative control, untransformed *E. coli* AB2829 was grown on minimal medium alone to establish if growth was inhibited under these conditions. All the cells were incubated at 37 °C for 2 days.

E. coli AB2829 cells grew successfully when transformed with pKD501 and when grown on minimal medium supplemented with aromatic amino acids. However, the untransformed mutant did not grow on unsupplemented minimal medium. These results established that auxotrophy of *E. coli* AB2829 was due to a mutation in the *aro*A gene of the shikimate pathway. Lack of growth of the mutant on minimal medium also showed that the reversion rate from the mutant to wild type phenotype was very low. Since the *aro*A mutation was proven to be stable, *E. coli* AB2829 could be used as a host in attempts to isolate the *A. variabilis aro*A gene by phenotypic complementation.

6.5 Complementation of *E. coli* AB2829 with the *Anabaena variabilis Eco*RI Genomic DNA Library

To test if a recombinant from the *A. variabilis* library complemented the deficiency of *E. coli* AB2829, 50 ng of the clone bank was transformed into the host cell and grown on minimal medium plates containing ampicillin. This medium selected plasmid transformed cells encoding an active EPSP synthase (Section 2.16). These cells were incubated at 37 °C for 2 to 3 days to allow for growth. To check the transformation efficiency, 100 µl of the transformed cell mixture was grown on LB / ampicillin. Approximately 1000 - 1200 transformants resulted per transformation reaction. Multiple transformations of the *A. variabilis Eco*RI library were performed to screen the calculated number of clones that gave a 99% probability of isolating the *A. variabilis aro*A gene (Section 6.3.3).

A total of six colonies grew on minimal medium / ampicillin plates. Plasmid DNA was purified from these transformants (Section 2.5.2) and digested with *Eco*RI

to release the insert. Analysis of digested plasmid DNA by agarose gel electrophoresis (Section 2.11) confirmed that two out of the six colonies contained recombinant DNA. These clones were back transformed into *E. coli* AB2829 and grown on minimal medium / ampicillin plates. This ensured that growth was due to expression of an active EPSP synthase from the cloned DNA. The second transformation step was crucial since it proved that both the isolated clones did not complement the *aro*A mutation of the host and were not responsible for prototrophic growth.

Since successive attempts did not yield a positive result, an *aro*A clone was not isolated by phenotypic complementation of *E. coli* AB2829 with the library of *A. variabilis Eco*RI genomic DNA fragments.

6.6 Discussion

Attempts were made to isolate the *A. variabilis aro*A gene by complementation of the *aro*A auxotrophic mutant, *E. coli* AB2829. An *A. variabilis* library of *Eco*RI genomic fragments was constructed in the phagemid vector, pBluescript, and transformed into *E. coli* AB2829. Transformed cells expressing an active EPSP synthase from a cloned piece of DNA would complement the *aro*A mutation of the host. This would result in growth of these transformants on minimal medium. *A. variabilis* DNA was digested with *Eco*RI for library construction because of evidence obtained from Southern blot experiments (Section 6.3.1). The signal pattern produced from hybridisation of *Synechocystis aro*A gene fragments to *Eco*RI digested *A. variabilis* DNA suggested the entire *aro*A sequence was carried on one restriction fragment, 2.5-kb in size. If this were the case, expression of a functional *A. variabilis* EPSP synthase able to complement the *aro*A mutation of *E. coli* AB2829 could have occurred. Yet, subsequent experimentation showed the library of *A. variabilis Eco*RI genomic fragments did not complement the *aro*A mutation of *E. coli* AB2829 (Section 6.5) and the *A. variabilis* EPSP synthase gene had not been found.

There is a feasible explanation for the occurrence of false positive clones that failed to complement *E. coli* AB2829 upon re-transformation (Section 6.5). This mutant host cell has an active recombination system. If a transformed plasmid contained a truncated *A. variabilis aro*A gene, complementation may have arisen from recombination of the plasmid-borne allele and the non-functional *aro*A gene of the host. If this occurred the plasmid-borne allele must have integrated into the host genetic material, since re-transformation of the putative positive clones into AB2829 did not yield a positive result. Owttrim and Coleman (1987) highlighted the problems caused by recombination when cloning a *rec*A-like gene by phenotypic complementation in *E. coli* HB101. In some instances the cloned *A. variabilis* DNA that expressed a RecA-like protein could not be recovered because the plasmid had integrated into the host genetic material.

The complementation strategy was chosen for a number of reasons. The size and catalytic properties of EPSP synthase from *A. variabilis* were shown to be very similar to the enzyme in *E. coli* (Powell *et al.*, 1992). So there was a good chance that the cyanobacterial EPSP synthase could have functionally complemented the deficiency of *E. coli* AB2829. Additionally, cyanobacterial genes previously cloned by complementation or expressed in *E. coli* suggested similarities existed between the structure of cyanobacterial and *E. coli* promoters. For example, the phosphoenolpyruvate carboxylase gene and the *arg*C gene, encoding N-acetylglutamate-5-semialdehyde dehydrogenase, of *Anabaena* sp PCC 7120 (Luinenburg and Coleman, 1992, Floriano *et al.*, 1992) plus the *rec*A gene of *A. variabilis* ATCC 29413 (Owttrim and Coleman, 1987) have each complemented a mutation in *E. coli*. The transcription factors of the heterologous host were able to express each cyanobacterial gene from their respective promoters. Cyanobacterial genes that cannot be expressed in *E. coli* tend to encode proteins with specialised functions. In some cases, expression of these proteins is dictated by unique signal / response mechanisms. An example of this is the transcription of the glutamine synthetase gene (*gln*A) of *Anabaena* sp. PCC 7120 which is involved in nitrogen

assimilation. Turner *et al.* (1983) reported that the *gln*A gene is transcribed from an *E. coli*-like promoter when *Anabaena* PCC 7120 cells were grown in the presence of ammonia. However, under nitrogen fixing conditions the glutamine synthetase gene was transcribed from a promoter that differed significantly from the *E. coli* consensus promoter. The low level of EPSP synthase activity in *A. variabilis* suggested that specialised signals were not required to induce the expression of the *aro*A gene. Indeed, the gene appeared to be constitutively expressed (Powell, 1991). Therefore, there was a reasonable possibility that the *A. variabilis aro*A gene had an *E. coli*-like promoter which would be recognised by the transcription machinery of *E. coli* AB2829.

The expression of the previously mentioned cyanobacterial genes in *E. coli* suggested that the cyanobacterial transcripts carried translation initiation sites able to bind the host ribosomes. Also, that the genetic codes and codon usages of the two organisms must have been compatible.

The various pieces of evidence previously mentioned indicated that expression of the *A. variabilis aro*A gene in *E. coli* was likely to be successful. The possible reasons explaining why the *A. variabilis Eco*RI library failed to complement the *aro*A mutation of *E. coli* AB2829 are also varied.

The absence of essential sequences within or adjacent to the EPSP synthase gene in the library could account for the lack of success. Digestion of *A. variabilis* DNA with *Eco*RI could have damaged the promoter region or sequences involved in transcription termination of the *aro*A gene. This would have affected gene expression. Alternatively, *Eco*RI digestion could have produced a truncated open reading frame that resulted in the expression of an inactive EPSP synthase. In these circumstances complementation would not have occurred. If the *aro*A gene had been cleaved by *Eco*RI, it was not detected by Southern hybridisation of the *Synechocystis aro*A probe to *A. variabilis* genomic DNA digested with this enzyme (Section 6.3.1). The *Synechocystis aro*A gene fragments used as probes may not have recognised the cleaved part of the EPSP synthase gene that may have contained a different part of the

sequence. If *Eco*RI had cleaved essential *aro*A sequences, libraries of *A. variabilis* genomic DNA digested with other restriction enzymes would have to be constructed. Screening these libraries by complementation using *E. coli* AB2829 may yield an *aro*A clone containing all the essential sequences required for the expression of active *A. variabilis* EPSP synthase.

Alternatively, the complementation experiments could have failed due to the existence of an extensive restriction system in the *E. coli* host strain, AB2829. The *E. coli* restriction systems are EcoK, McrA, McrBC and Mrr. EcoK cleaves DNA unprotected by adenine methylation at a specific recognition site (Bickle, 1993). The genomic DNA extracted from *A. variabilis* cells used for these cloning experiments would not have been methylated by the EcoK modification system. So any unprotected EcoK restriction sites present in the cyanobacterial DNA would have been a substrate for cleavage by this enzyme. McrA, McrBC and Mrr systems restrict DNA at specific sequences containing methylated adenine and cytosine residues (Raleigh *et al.*, 1991). Filamentous cyanobacterial DNA is known to have a high content of methylated adenine and especially cytosine residues (Padhy *et al.*, 1988). Therefore, *A. variabilis* DNA could have been cleaved by such methylation requiring restriction systems.

An example of the problems encountered when cloning *Anabaena* DNA in *E. coli* strains containing active restriction systems was illustrated in the work of Black and Wolk (1994). The authors noticed differences in plating efficiencies when phage libraries of *Anabaena* sp. PCC 7120 DNA were grown in the *E. coli* strains KW251 (*mcr*A, *mcr*B) and LE392 (*mcr*A, *mcr*B$^+$). Plating efficiencies were approximately 150 times higher when KW251 was used as the host strain for the cyanobacterial library than when LE392 was used. However, infection of both strains with recombinant phage containing a control insert resulted in equal plating efficiencies. The greater representation of cyanobacterial sequences in the *mcr*B$^-$ strain, KW251, than in the *mcr*B$^+$ strain, LE392 suggested that cleavage of the cloned *Anabaena* DNA by the McrB restriction system of *E. coli* LE392 was quite extensive.

If the effects of the McrB restriction system of *E. coli* LE392 on *Anabaena* PCC 7120 genomic DNA were mirrored in *E. coli* AB2829 with *A. variabilis* ATCC 29413 DNA, the quality of the *A. variabilis* library would be reduced. Therefore, there would be less chance of isolating an *aro*A clone by phenotypic complementation.

If the *A. variabilis* library was initially prepared in an *E. coli* strain containing an active EcoK modification system and inactive restriction systems prior to transformation into *E. coli* AB2829, the cloned cyanobacterial DNA would become methylated at the appropriate sites while being protected from restriction enzyme cleavage. Also, the methylation pattern endogenous to the cyanobacterial DNA would be reduced in recombinants purified from these *E. coli* cells. Upon transformation into *E. coli* AB2829, the *A. variabilis* library would be protected from the host's restriction systems. This would improve the quality of the *A. variabilis* library and the probability of isolating an *aro*A clone.

However, persistent screening of the *A. variabilis Eco*RI library in *E. coli* AB2829 could have yielded a positive result. Even if a clone able to express an active EPSP synthase had been subjected to digestion by the restriction systems of AB2829, some recombinants may have been left intact within the cell. This could have permitted low level expression of the cyanobacterial EPSP synthase gene which might have been sufficient to complement the metabolic deficiency of the host.

Chapter 7

General Discussion

7.1 The Objectives of this Work

The aim of the work described in this thesis was to isolate the *Anabaena variabilis* gene encoding the shikimate pathway enzyme, EPSP synthase, and deduce its amino acid sequence. Achieving this objective would have permitted the comparison of this cyanobacterial enzyme sequence to EPSP synthase sequences from other microorganisms and plants. This would have allowed the identification of conserved amino acids potentially involved in enzyme function. Large quantities of *A. variabilis* EPSP synthase could have been obtained by overexpressing the isolated *aro*A gene in a suitable host. To accomplish this, the cyanobacterial gene would have been placed under the control of a strong promoter contained in an appropriate vector system. This would have provided a sufficient amount of enzyme to perform numerous studies. The kinetic information gathered by Powell *et al.* (1992) could have been verified. Collection of the original kinetic data, especially the specific activity, was impeded by the small quantities of enzyme employed during these experiments. Small amounts of enzyme were used because of the low yield of EPSP synthase originally purified from *A. variabilis* cells. The overexpressed EPSP synthase would also have allowed chemical modification studies to proceed which might have identified reactive residues possibly involved in enzyme catalysis. The 3-dimensional positions of the atoms in *A. variabilis* EPSP synthase could have been determined by performing X-ray crystallography. The overexpressed enzyme would have provided enough protein to carry out experimental trials to obtain crystals for X-ray diffraction analysis. This could have indicated how the substrates, S3P and PEP, and the inhibitor, glyphosate, interact at the active centre. None of these objectives were accomplished because of the problems encountered when attempting to clone the *A. variabilis aro*A gene. The different cloning strategies employed were: amplification of an *aro*A gene fragment from *A. variabilis* genomic DNA using degenerate oligonucleotide primers by PCR; screening genomic DNA libraries using heterologous *aro*A probes and by phenotypic complementation in an *aro*A$^-$ auxotroph of *E. coli*. The reasons for using these

techniques and their benefits and drawbacks with specific regard to this work are discussed in Chapters 3 - 6.

7.2 Alternative Strategies for Cloning the *Anabaena variabilis aroA* Gene

At the outset of this project other gene cloning strategies were considered but were not pursued for various reasons.

The amino acid sequence of a protein or a peptide provides valuable information for the design of oligonucleotides for use as primers for PCR or probes for screening DNA libraries. A process called Edman degradation can determine the N-terminal sequence of a protein or peptide (Matsudaira, 1990). Approximately 10 - 100 pmoles of the protein of interest is required for this procedure. This would have been the equivalent of 0.5 - 5 μg of the 49 kDa protein EPSP synthase from *A. variabilis*. The average yield of protein obtained from 25 g *A. variabilis* cells (or one EPSP synthase preparation) was 60 ng or 1.2 pmoles (Powell *et al.*, 1992). This was an insufficient amount of protein to obtain sequence information by Edman degradation. To acquire enough protein for this process, several *A. variabilis* EPSP synthase preparations would have to be performed . Since fifty litres of *A. variabilis* cells had to be grown over a 3 - 4 week period to yield approximately 25 g of cells for one protein preparation, this cloning strategy was considered too time consuming and so was not explored.

The detection of clones expressing a particular protein using an antibody probe is one of the most direct methods of screening. In most cases, this strategy necessitates the expression of the desired protein in its functional form. To achieve this, foreign DNA is cloned into a plasmid or phage vector downstream of a strong, regulatable promoter and ribosome binding site. This allows the expression of the cloned DNA in a suitable *E. coli* host. To raise an antibody to the *A. variabilis* EPSP synthase, different techniques were considered. A conventional method produces polyclonal antiserum from an immunised animal, such as a mouse, rat or rabbit.

Polyclonal antiserum contains many different antibody specificities to various epitopes of the structurally complex immunogen (in this case EPSP synthase). Alternatively, monoclonal antibodies can be raised by immunising an animal with the protein. The spleen cells of the immunised animal are subsequently isolated and cloned to produce high levels of antibody specific for a single antigen determinant of the immunogen. For both techniques, a large amount of protein is required to immunise the animal. In mice, at least 50 μg of immunogen is needed to elicit an immune response (Johnstone and Thorpe, 1982). The low yield of EPSP synthase extracted from *A. variabilis* cells meant that raising antibodies using these procedures was unrealistic. A different approach involves raising antibodies to peptide sequences rather than the whole protein. Edman degradation could have provided the amino acid sequence information necessary for the construction of peptides for this purpose. Due to the potential difficulties involved in determining the amino acid sequence of the *A. variabilis* EPSP synthase (described above), this approach was not investigated. A polyclonal antiserum against the *E. coli* EPSP synthase has been raised (Likidlilid, 1989). Western blot experiments would have revealed if the *E. coli* EPSP synthase antibody (anti-E5) could cross-react with a 49 kDa protein corresponding to the *A. variabilis* EPSP synthase. If so, this antibody could have been used to screen an *A. variabilis* expression library to isolate an EPSP synthase clone. Since this DNA library would have been expressed in *E. coli* cells, anti-E5 would have reacted with the EPSP synthase expressed by the host. To eliminate this problem, a mutant form of *E. coli* unable to express EPSP synthase that is recognised by anti-E5 would have to be constructed for these screening experiments.

The huge development in gene transfer techniques in cyanobacteria and the ability to clone and inactivate their genes has advanced the study of these organisms at the genetic level (Thiel, 1994). The construction of mutant forms of cyanobacteria, such as *Anabaena* sp. PCC 7120, and complementation of those mutations by wild type genes has provided a valuable route towards the isolation and identification of novel genes (Wolk *et al.*, 1988; Buikema and Haselkorn, 1991a and b; Floriano *et al.*, 1992). To apply this method to the isolation of the *A. variabilis* EPSP synthase gene, an *aro*A

auxotrophic mutant of *A. variabilis* would have to be constructed by chemical or U.V. light mutagenesis. Also, a suitable gene transfer system would have to be developed. Conjugation has been the method of choice for DNA transfer in filamentous cyanobacteria. DNA is transferred by cell to cell contact. This technique is based on mobilisation of DNA from one bacterium (usually *E. coli*) to another via a broad host range conjugative plasmid. Shuttle vectors that can efficiently replicate in both *E. coli* and *A. variabilis* cells have yet to be developed. The future construction of such vectors and *aro*A auxotrophic *A. variabilis* cells would allow the *A. variabilis* EPSP synthase gene to be isolated by phenotypic complementation in a homologous host. This would avoid problems regarding gene expression that have previously been encountered when expressing cyanobacterial genes in heterologous hosts such as *E. coli* (Porter *et al.*, 1985 and Section 6.6).

7.3 Reasons for Cloning the *Anabaena variabilis aro*A Gene

As discussed in Section 1.8, many glyphosate tolerant transgenic plants have been constructed using mutated EPSP synthase genes. The majority of these herbicide tolerant plants are not commercially viable because they cannot withstand the levels of glyphosate used in agriculture. The poor growth of these transgenic plants may be attributed to the inability of the mutant EPSP synthase to function efficiently in the plant cell. This may be due to the decreased binding efficiency of the enzyme for its substrate, PEP, which appears to be a trait of most glyphosate tolerant EPSP synthases. Some of the mutant EPSP synthase genes used to construct the glyphosate tolerant plants were of bacterial origin. In these cases, poor growth might have been caused by inadequate expression of the foreign gene in the plant cells. The gene encoding the glyphosate tolerant *A. variabilis* EPSP synthase may be more appropriate for constructing herbicide resistant transgenic plants for a number of reasons. The cyanobacterial enzyme has a greater binding efficiency for PEP than other variant, glyphosate tolerant EPSP synthases. Also, the *A. variabilis aro*A gene may be more efficiently expressed in plant cells than other genes of bacterial origin. This is because

the cyanobacteria are thought to share a common ancestry with higher plant chloroplasts (Giovannoni *et al.* 1988) and *Anabaena* species have a codon usage similar to that found in these plastids (Krishnaswamy and Schanmugasundraram, 1995). This is a subject that merits further investigation but, as yet, there are no examples in the literature that explore the stability of cyanobacterial genes expressed in plant cells.

It is generally believed that the plastids of higher plant and algal cells arose from endosymbiosis of an ancestral cyanobacterium into an ancient eukaryotic cell. This hypothesis was initially based on the striking similarities between plant chloroplasts and cyanobacteria: both undergo oxygenic photosynthesis (Ho and Krogmann, 1982), both assimilate ammonia via the glutamine synthetase / glutamine: 2-oxoglutarate aminotransferase pathway (Kerby *et al.*, 1989) and both synthesise carotenoids via phytoene (Chamovitz *et al.*, 1991). Studies comparing 16S rRNA sequences from higher plants, algae and cyanobacteria have been documented. These nucleic acid molecules have been used to establish evolutionary relationships between different organisms because of their ubiquity in nature. This research suggests that plant and algal plastids originate from a common cyanobacterium-like progenitor (Giovannoni *et al.*, 1988; Douglas and Turner, 1991). Further analysis has also indicated that the cyanobacteria had diversified into a variety of different species prior to endosymbiosis with an ancient eukaryotic cell. Nelissen *et al.* (1995) argue that plastid development occurred during the early stages of cyanobacterial diversification. Once endosymbiosis had been established, it is believed that a significant amount of genetic material was transferred from the prokaryotic genome to the eukaryotic genome since the majority of chloroplast proteins need to be imported from the cytosol of the plant cell. Also, Martin *et al.* (1993) showed that plant nuclear genes encoding different isoforms of glyceraldehyde-3-phosphate dehydrogenase could have originated from an endosymbiont of cyanobacterial origin because of their similarity to corresponding sequences in *A. variabilis*.

The enzymes involved in aromatic amino acid biosynthesis are located in the plastids of higher plant and algal cells. The organisation and regulation of the enzymes

of this pathway were analysed in higher plants, *Euglena gracilis*, a representative of the green algae and cyanobacteria (Bonner *et al.*, 1995). In this context, the cyanobacteria are very different from these other eukaryotic organisms. On the basis of this evidence, the plastids of higher plants, green algae and euglenoids did not arise from the endosymbiosis of a cyanobacterium-like progenitor. This study contradicts the massive weight of evidence indicating the contrary. The characteristics of the aromatic amino acid biosynthetic pathway in other algal representatives have yet to be ascertained but once determined would contribute to the debate on the origin of plastid organelles.

The amino acid sequences of EPSP synthases from various plants and microorganisms have been analysed and compared. The similarity of the Gram negative and higher plant sequences supports the hypothesis that the plastids of higher plants are descendants of endosymbiotic prokaryotes that invaded an ancient eukaryotic cell (Griffin and Griffin, 1991; Gasser *et al.*, 1988; Margulis, 1970). The plant *aro*A sequence may, therefore, represent a gene that was transferred from the endosymbiont to the nucleus. As such, these genes would be expected to retain some homology to the corresponding sequences in prokaryotes (Weedon, 1981). Interestingly, the EPSP synthase sequence from the unicellular cyanobacterium, *Synechocystis* sp. PCC 6803, shows more homology to the Gram positive bacterium, *Bacillus subtilis* than to the available plant sequences (Mayes *et al.*, 1993). So *Synechocystis* may not be the modern-day representative of the cyanobacterial sub-line that became the progenitor of plant and algal plastids. Determining the relationship between the *A. variabilis* and higher plant EPSP synthase sequences may have given a new perspective to the theory of endosymbiosis and perhaps created more discussion.

Chapter 8

References

Ahmed, S.I. and Giles, N.H. (1969) *Journal of Bacteriology*, **99**, (1), 231-237.

Ajlani, G., Kirilovsky, D., Picaud, M. and Atier, C. (1989) *Plant Molecular Biology*, **13**, 469-479.

Alberg, D.G. and Bartlett, P.A. (1989) *Journal of the American Chemical Society*, **111**, 2337-2338.

Alberg, D.G. and Bartlett, P.A. (1992) *Journal of the American Chemical Society*, **114**, 3535-3546.

Alm, R.A., Dalrymple, B.P. and Mattick, J.S. (1994) *Gene*, **145**, 97-101.

Amrhein, N., Deus, B., Gehrke, P. and Steinrucken, H.C. (1980) *Plant Physiology*, **66**, 830-834.

Amrhein, N., Hollander-Czytko, H., Johanning, D., Schulz, A., Smart, C.C. and Steinrucken, H.C. (1987) *Plant Biology*, **3**, 119-133.

Amrhein, N., Johanning, D., Schab, J. and Schulz, A. (1983) *FEBS Letters*, **157**, (1), 191-196.

Anderson, K.S. and Johnson, K.A. (1990) *Chemical Reviews*, **90**, 1131-1149.

Anderson, K.S., Sikorski, J.A., Johnson, K.A. (1988a) *Biochemistry*, **27**, 1604-1610.

Anderson, K.S., and Sikorski, J.A. (1988b) *Journal of the American Chemical Society*, **110**, 6577-6579.

Anderson, M.L.M. and Young, B.D. (1985) Quantitative Filter Hybridisation. In: *Nucleic Acid Hybridisation - a practical approach* (eds. Hames, B.D. and Higgins, S.J.), pp. 73-111, IRL Press, Oxford.

Arnaud, L., Nurit, F., Ravanel, P. and Tissut, M. (1994) *Pesticide Science*, **40**, 217-223.

Arnheim, N. and Erlich, H. (1992) *Annual Review of Biochemistry*, **61**, 131-156.

Banerji, S, Wakefield, A.E., Allen, A.G., Maskell, D.J., Peters, S.E. and Hopkin, J.M. (1993) *Journal of General Microbiology*, **139**, 2901-2914.

Barinaga, M. (1990) *Science*, **250**, 1512.

Bej, A. K., Mahbubani, M.H., Atlas, R.M. (1991) *CRC Critical Reviews Biochemistry and Molecular Biology*, **26**, (3/4), 301-334.

Bentley, R. (1990) *Critical Reviews in Biochemistry and Molecular Biology*, **25**, (5), 307-383.

Berlyn, M.B. and Giles, N.H. (1969) *Journal of Bacteriology*, **99**, 220-230.

Berlyn, M.B., Ahmed, S.I. and Giles N.H. (1970) *Journal of Bacteriology*, **104**, 768-774.

Bickle, T.A. (1993) The ATP-dependent Restriction Enzymes. In: *Nucleases* (eds. Linn, S.M., Lloyd, R.S. and Roberts, R.G), pp. 89-109, CSH, New York, U.S.A.

Birnboim, H.C. and Doly, J. (1979) *Nucleic Acids Research*, **7**, (6), 1513-1523.

Black, T.A. and Wolk, C.P. (1994) *Journal of Bacteriology*, **176**, (8), 2282-2292.

Bondinell, W.E., Vnek, J., Knowles, P.F., Sprecher, M. and Sprinson, D.B. (1971) *Journal of Biological Chemistry*, **246**, 6191-6196.

Bonner, C.A., Fischer, R.S., Schmidt, R.R., Miller, P.W. and Jensen, R.A. (1995) *Plant Cell Physiology*, **36**, (6), 1013-1022.

Boocock, M.R. and Coggins, J.R. (1983) *FEBS Letters*, **154**, (1), 127-133.

Boujoub, A., Vernotte, C. and Atier, C. (1993) *Plant Molecular Biology*, **21**, 249-258.

Bryant, D.A. and Tandeau de Marsac, N. (1988) *Methods in Enzymology*, **167**, 755-765.

Buikema, W.J. and Haselkorn, R. (1991a) *Genes and Development*, **5**, 321-330.

Buikema, W.J. and Haselkorn, R. (1991b) *Journal of Bacteriology*, **173**, 1879-1885.

Byng, G.S., Kane, J.F. and Jensen, R.A. (1982) *CRC Critical Reviews in Microbiology*, **9**, 227-252.

Campbell, W.H. and Gowri, G. (1990) *Plant Physiology*, **92**, 1-11.

Carr, N.G. and Whitton, B.A. (1982) (eds.) *The Biology of Cyanobacteria*, Blackwell Scientific, Oxford.

Castellino, S., Leo, G.C., Sammons, R.D. and Sikorski, J.A. (1989) *Biochemistry*, **28**, 3856-3868.

Chamovitz, D., Pecker, I. and Hirschberg, J. (1991) *Plant Molecular Biology*, **16**, 967-974.

Chamovitz, D., Sandmann, G. and Hirschberg, J. (1993) *Journal of Biological Chemistry*, **268**, (23), 17348-17353.

Charles, I.G., Keyte, J.W., Brammar, W.J., Smith, M. and Hawkins, A.R. (1986) *Nucleic Acids Research*, **14**, (5), 2201-2213.

Christensen, A.M. and Schaefer, J. (1993) *Biochemistry*, **32**, 2868-2873.

Clarke, L. and Carbon, J. (1976) *Cell*, **9**, 91-99.

Coggins, J.R., Boocock, M.R., Campbell, M.S., Chaudhuri, S., Lambert, J.M., Lewendon, A., Mousdale, D.M. and Smith, D.D.S. (1985) *Biochemistry Society Transactions*, **13**, 299-303.

Comai, L., Facciotti, D., Hiatt, W.R., Thompson, G., Rose, R.E. and Stalker, D.M. (1985) *Nature*, **317**, 741-744.

Comai, L., Sen, L.C. and Stalker, D.M. (1983) *Science*, **221**, 370-371.

Curtis, S.E. and Martin, J.A. (1994) The Transcription Apparatus and the Regulation of Transcription Initiation. In: *The Molecular Biology of Cyanobacteria* (ed. Bryant, D.A.), pp. 613-639, Kluwer Academic Publishers, The Netherlands.

dalla Chiesa, M., Mayes, S.R., Maskell, D.J., Nixon, P.J. and Barber, J. (1994) *Gene*, **144**, 145-146.

Davis, B.D. (1953) *Science*, **118**, 251-252.

Davis, B.D. (1955) *Advances in Enzymology*, **16**, 247-312.

Deka, R.K., Anton, I.A., Dunbar, B and Coggins, J.R. (1994) *FEBS Letters*, **349**, 397-402.

Delannay, X., Bauman, T.T., Beighley, D.H., Buettner, M.J., Coble, H.D., DeFelice, M.S., Derting, C.W., Diedrick, T.J., Griffin, J.L.,

Hagood, E.S., Hancock, F.G., Hart, S.E., LaVallee, B.J., Loux, M.M., Lueschen, W.E., Matson, K.W., Moots, C.K., Murdock, E., Nickell, A.D., Owen, M.D.K., Paschal II, E.H., Prochaska, L.M., Raymond, P.J., Reynolds, D.B., Rhodes, W.K., Roeth, F.W., Sprankle, P.L., Tarochione, L.J., Tinius, C.N., Walker, R.H., Wax, L.M., Weigelt, H.D. and Padgette, S.R. (1995) *Crop Science*, **35**, (5), 1461-1467.

della Cioppa, G., Bauer, S.C., Klein, B.K., Shah, D.M., Fraley, R.T. and Kishore, G.M. (1986) *Proc. Natl. Acad. Sci. U.S.A.*, **83**, 6873-6877.

della Cioppa, G., Bauer, S.C., Taylor, M.L., Rochester, D.E., Klein, B.K., Shah, D.M., Fraley, R.T., Kishore, G.M. (1987) *Bio/technology*, **5**, 579-584.

della Cioppa, G and Kishore, G.M. (1988) *The EMBO Journal*, **7**, (5), 1299-1305.

Dieffenbach, C.W., Lowe, T.M.J. and Dveksler, G.S. (1993) *PCR Methods and Applications*, **3**, S30-S37.

DiMagno, L. and Haselkorn, R. (1993) *Plant Molecular Biology*, **21**, 835-845.

Doolittle, W.F. (1982) Molecular Evolution. In: *The Biology of Cyanobacteria* (eds. Carr, N.G. and Whitton, B.A.), pp. 307-331, Blackwell Scientific, Oxford.

Douglas, S. (1994) Chloroplast Origins and Evolution. In: *The Molecular Biology of Cyanobacteria* (ed. Bryant, D.A.), pp. 91-118, Kluwer Academic Publishers, The Netherlands.

Douglas, S.E. and Turner, S. (1991) *Journal of Molecular Evolution*, **33**, (3), 267-273.

Duncan, K. and Coggins, J.R. (1984) *Biochemical Society Transactions*, **12**, 274-275.

Duncan, K. and Coggins, J.R. (1986) *Biochemical Journal*, **234**, 49-57.

Duncan, K., Edwards, R.M. and Coggins, J.R. (1987) *Biochemistry Journal*, **246**, 375-386.

Duncan, K., Lewendon, A. and Coggins, J.R. (1984a) *FEBS Letters*, **165**, (1), 121-127.

Duncan, K., Lewendon, A. and Coggins, J.R. (1984b) *FEBS Letters*, **170**, (1), 59-63.

Dyer, W.E., Weller, S.C., Bressan, R.A. and Herrmann, K.M. (1988) *Plant Physiology*, **88**, 661-666.

Elgin, E., Mackman, C., Pabst., M., Kretz, P.L. and Greener, A., (1991) *Strategies*, **4**, 8-9.

Elhai, J. (1994) *Journal of Applied Phycology*, **6**, 177-186.

Feinberg, A.P. and Vogelstein, B. (1983) *Analytical Biochemistry*, **132**, 6-13.

Feinberg, A.P. and Vogelstein, B. (1984) *Analytical Biochemistry*, **137**, 266-267.

Fillatti, J.J., Haissig, B., McCown, B., Comai, L. and Riemenschneider, D. (1988) Development of Glyphosate Tolerant Populus Plants through Expression of a Mutant *aro*A Gene from *Salmonella typhimurium.* In: *Genetic Manipulation of Woody Plants* (eds. Hanover, J.W. and Keathley, D.E.), pp. 243-249, Plenum Publishing Corporation.

Fillatti, J.J., Kiser, J., Rose, R. and Comai, L. (1987) *Bio/technology*, **5**, 726-730.

Fisher, R., Tuli, R. and Haselkorn, R. (1981) *Proc. Natl. Acad. Sci. U.S.A.*, **78**, (6), 3393-3397.

Floriano, B., Herrero, A., and Flores, E. (1992) *Molecular Microbiology*, **6**, (15), 2085-2094.

Forlani, G., Parisi, B. and Nielson, E. (1994) *Plant Physiology*, **105**, 1107-1114.

Franz, J.E. (1985) Discovery, Development and Chemistry of Glyphosate. In: *The Herbicide Glyphosate* (eds. Grossbard, E. and Atkinson, D.), pp. 3-17, Butterworths, Woburn, MA.

Friedberg, D. and Seijffers, J. (1988) *Archives of Microbiology*, **150**, 278-281.

Friedman, E. and Jensen, R.A. (1978) *Biochemical Genetics* **16**, 867-881.

Ganson, R.J. and Jensen,R.A. (1988) *Archives of Biochemistry and Biophysics*, **260**, (1), 85-93.

Garbe, T., Jones, C., Charles, I., Dougan, G. and Young, D. (1990) *Journal of Bacteriology*, **172**, (12), 6774-6782.

Gasser, C.S. and Klee, H.J. (1990) *Nucleic Acids Research*, **18**, 2821.

Gasser, C.S., Winter, J.A., Hironaka, C.M. and Shah, D.M. (1988) *Journal of Biological Chemistry*, **263**, (9), 4280-4289.

Gibson, F. and Pittard, J. (1968) *Bacteriological Reviews*, **32**, 465-492.

Gibson, T.J. (1984) Studies on the Epstein-Barr Virus Genome, *Ph.D. Thesis*, University of Cambridge, England, U.K.

Giovannoni, S.J., Turner, S., Olsen, G.J., Barns, S., Lane, D.J. and Pace, N.R. (1988) *Journal of Bacteriology*, **170**, (8), 3584-3592.

Goldsbrough, P.B., Hatch, E.M., Huang, B., Kosinski, W.G., Dyer, W.E., Herrmann, K.M. and Weller, S.C. (1990) *Plant Science*, **72**, 53-62.

Gollub, E., Zalkin, H. and Sprinson, D.B. (1967) *Journal of Biological Chemistry*, **242**, 5323-5328.

Granger, S.P. (1989) Studies on the Genes for the Enzymes of the Shikimate Pathway from *Pisum sativum*, *Ph.D. Thesis*, University of Glasgow, Scotland, U.K.

Greener, A. (1990) *Strategies*, **3**, 5-6.

Greener, A. and Jerspeth, B. (1993) *Strategies*, **6**, (2), 57.

Griffin, H.G. and Gasson, M. J. (1995) *Molecular and General Genetics*, **246**, 119-127.

Griffin, H.G. and Griffin, A.M. (1991) *Journal of General Microbiology*, **137**, 113-121.

Grossman, A.R., Schaefer, M.R., Chiang, G.G., Collier, J.L. (1993) *Microbiological Reviews*, **57**, (3), 725-749.

Gruys, K.J., Marzabadi, M.R., Pansegrau, P.D. and Sikorski, J.A. (1993) *Archives of Biochemistry and Biophysics*, **304**, (2), 345-351.

Gruys, K.J., Walker, M.C. and Sikorski, J.A. (1992) *Biochemistry*, **31**, 5534-5544.

Hall, G.C., Flick, M.B., Gherna, R.L. and Jensen, R.A. (1982) *Journal of Bacteriology*, **149**, (1), 65-78.

Hall, G.C. and Jensen, R.A. (1980) *Journal of Bacteriology*, **144**, 1034-1042.

Hall, G.C. and Jensen, R.A. (1981a) *Journal of Bacteriology*, **148**, (1), 361-364.

Hall, G.C. and Jensen, R.A. (1981b) *Current Microbiology*, **6**, 189-194.

Hanahan, D. (1985) Techniques for Transformation of *E. coli*. In: *DNA Cloning - a practicle approach*, Volume 1 (ed. Glover, D.M.), pp. 109-135, IRL Press, Oxford.

Haselkorn, R., Golden, J.W., Lammers, P.J. and Mulligan, M.E. (1986) *Trends in Genetics*, **2**, 255-259.

Haslam, E. (1993) *Shikimic Acid: Metabolism and Metabolites*, John Wiley & Sons Ltd., West Sussex.

Hauptmann, R.M., della Cioppa, G., Smith, A.G., Kishore, G.M. and Widholm, J.M. (1988) *Molecular and General Genetics*, **211**, 357-363.

Hawkins, A.R. (1985) *Nucleic Acids Research*, **13**, (22), 8119-8128.

Henner, D.J., Band, L., Flaggs, G. and Chen, E. (1986) *Gene*, **49**, 147-152.

Henner, D and Yanofsky, C. (1993) Biosynthesis of Aromatic Amino Acids. In: Bacillus subtilis *and Other Gram Positive Bacteria* (eds. Sonenshein, A., Hoch, J.A. and Losick, R.) pp. 269-280, American Society for Microbiology, Washington D.C.

Herdman, M., Janvier, M., Waterbury, J.B., Rippka, R. and Stanier, R.Y. (1979) *Journal of General Microbiology*, **111**, 63-71.

Herrero, A., Elhai, J., Hohn, B. and Wolk, C.P. (1984) *Journal of Bacteriology*, **160**, (2), 781-784.

Ho, K.K. and Krogmann, D.W. (1982) Photosynthesis. In *The Biology of Cyanobacteria* (eds. Carr, N.G. and Whitton, B.A.), pp. 191-214, Blackwell Scientific Publications, Oxford.

Hoiseth, S.K. and Stocker, B.A.D. (1981) *Nature*, **291**, 238-239.

Hoiseth, S.K. and Stocker, B.A.D. (1985) *Journal of Bacteriology*, **163**, (1), 355-361.

Hollander, H. and Amrhein,N. (1980) *Plant Physiology*, **66**, 823-829.

Hollander-Czytko, H., Sommer, I. and Amrhein, N. (1992) *Plant Molecular Biology*, **20**, 1029-1036.

Holmes, D.S. and Quigley, M. (1981) *Analytical Biochemistry*, **114**, 193-197.

Holtzclaw, W.D., Chapman, L.F. and Gowans, C.S. (1972) *Biochimica et Biophysica Acta*, **268**, 562-572.

Horsburgh, M.J. (1995) Chorismate Synthase from *Staphylococcus aureus* , *Ph.D. Thesis*, University of Glasgow, Scotland, U.K.

Huynh, Q.R. (1987) *Archives of Biochemistry and Biophysics*, **258**, (1) 233-239.

Huynh, Q.R. (1988) *Journal of Biological Chemistry*, **263**, (24), 11631-11635.

Huynh, Q.K. (1990) *Journal of Biological Chemistry*, **265**, (12), 6700-6704.

Huynh, Q.K. (1991) *Archives of Biochemistry and Biophysics*, **284**, (2), 407-412.

Huynh, Q.K. (1992) *Biochemical and Biophysical Research Communications*, **185**, (1), 317-322.

Huynh, Q.K. (1993) *Biochemical Journal*, **290**, 525-530.

Huynh, Q.K., Kishore, G.M. and Bild, G.S. (1988a) *Journal of Biological Chemistry*, **263**, (2), 735-739.

Huynh, Q.K., Bauer, S.C., Bild, G.S., Kishore, G.M. and Borgmeyer, J.R. (1988b) *Journal of Biological Chemistry*, **263**, (24), 11636-11639.

Jaworski, E.G. (1972) *Journal of Agricultural and Food Chemistry*, **20**, (6), 1195-1198.

Jensen, R.A. (1986) Tyrosine and Phenylalanine Biosynthesis: Relationship Between Alternative Pathways, Regulation and Sub-cellular Location. In: *The Shikimic Acid Pathway*, Recent Advances in Phytochemistry, Volume 20, (ed. Conn, E.E.), pp. 57-81, Plenum Press, New York.

Jensen, R.A. and Hall, G.C. (1982) *TIBS*, **7**, 177-180.

Jensen, R.A. and Stenmark, S.L. (1975) *Journal of Molecular Evolution*, **4**, 249-259.

Jerspeth, B., Greener, A., Short, J.M., Viola, J. and Kretz, P.L. (1992) *Strategies*, **5**, (3), 81-83.

Johnstone, A. and Thorpe, R. (1982) Production of Antibodies. In: *Immunochemistry in Practice*, pp. 27-40, Blackwell Scientific Publications, London, England.

Kaiser, K. and Murray, N.E. (1985) The Use of Phage Lambda Replacement Vectors in the Construction of Representative Genomic DNA Libraries. In: *DNA Cloning - a practicle approach*, Volume 1 (ed. Glover, D.M.), pp. 1-47, IRL Press, Oxford.

Katayama, M., Wada, Y., and Ohmori, M. (1995) *Journal of Bacteriology*, **177**, (13), 3873-3878.

Kerby, N.W., Rowell, P. and Stewart, W.D.P. (1989) The Transport, Assimilation and Production of Nitrogenous Compounds by Cyanobacteria and Microalgae. In: *Algal and Cyanobacterial Biotechnology* (eds. Cresswell, R.C., Rees, T.A.V., Shah, N.), pp.50-76.

Kishore, G.M., Brundage, L., Kolk, K., Padgette, S.R., Rochester, D., Huynh, Q.K., della Cioppa, G. (1986) *Federation Proceedings* **45**, 1506.

Kishore, G.M., Padgette, S.R. and Fraley, R.T. (1992) *Weed Technology*, **6**, 626-634.

Kishore, G.M. and Shah, D.M. (1988) *Annual Review of Biochemistry*, **57**, 627-663.

Klee, H.J., Muskopf, Y.M. and Gasser, C.S. (1987) *Molecular and General Genetics*, **210**, 437-442.

Knoth, K., Roberds, S., Poteet, C., Tamkun, M. (1988) *Nucleic Acids Research*, **16**, (22), 10932.

Koshiba, T. (1978) *Biochimica et Biophysica Acta*, **522**, 10-18.

Krishnaswamy, S. and Schanmugasundraram, S. (1995) *Current Science*, **69**, (2), 182-185.

Lea, P.J., Joy, K.W., Ramos, J.L. and Guerrero, M.G. (1984) *Phytochemistry*, **23**, 1-6.

Leonhardt, K.G. and Straus, N.A. (1989) *Nucleic Acids Research*, **17**, (11), 4384.

Levin, J.G. and Sprinson, D.B. (1964) *Journal of Biological Chemistry*, **239**, 1142-1150.

Lewendon, A. and Coggins, J.R. (1983) *Biochemical Journal*, **213**, 187-191.

Lewendon, A. and Coggins, J.R. (1987) *Methods in Enzymology*, **142**, 342-348.

Likidlilid, A. (1989) Antibodies Against the Shikimate Pathway Enzymes, *M.Sc. Thesis*, University of Glasgow, Scotland, U.K.

Luinenburg, I. and Coleman, J.R. (1992) *Journal of General Microbiology*, **138**, 685-691.

Lumsden, J and Coggins, J.R. (1977) *Biochemistry Journal*, **161**, 599-607.

MacKinney, G. (1941) *Journal of Biological Chemistry*, **140**, 315-322.

Majumder, K., Selvapandiyan, A., Fattah, F.A., Arora, N., Ahmad, S. and Bhatnagar, R.K. (1995) *European Journal of Biochemistry*, **229**, 99-106.

Margulis, L. (1970) Eukaryotic Plants. In: *Origin of Eukaryotic Cells* (ed. Margulis, L.), pp. 276-293, Yale University Press, New Haven, CT.

Marmur, J.G. and Doty, P. (1961) *Journal of Molecular Biology*, **3**, 585-594.

Martin, W., Brinkmann, H., Savonna, C. and Cerff, R. (1993) *Proc. Natl. Acad. Sci. U.S.A.*, **90**, 8692-8696.

Maskell, D. (1993) *Gene*, **129**, 155-156.

Maskell, D.J., Morrissey, P. and Dougan, G. (1988) *Journal of Bacteriology*, **170**, (6), 2467-2471.

Matsudaira, P. (1990) *Methods in Enzymology*, **182**, 602-613.

Mayes, S.R., alla Chiesa, M., Zhang, Z. and Barber, J. (1993) *FEBS*, **325**, (3), 255-261.

Mazur, B.J., Rice, D. and Haselkorn, R. (1980) *Proc. Natl. Acad. Sci. U.S.A.*, **77**, (1), 186-190.

McConaughy, B.L., Laird, C.L. and McCarthy, B.J. (1969) *Biochemistry*, **8**, 3289-3295.

Meeks, J.C., Wolk, C.P., Lockau, W., Schilling, N., Shaffer, P.W. and Chien, W.S. (1978) *Journal of Bacteriology*, **134**, 125-130.

Meinkoth, J. and Wahl, G. (1984) *Analytical Biochemistry*, **138**, 267-284.

Messing, J. (1983) *Methods in Enzymology*, **101**, 20-78.

Mousdale, D.M. and Coggins, J.R. (1984) *Planta*, **160**, 78-83.

Mousdale, D.M. and Coggins, J.R. (1985) *Planta*, **163**, 241-249.

Mousdale, D.M. and Coggins, J.R. (1986) *FEBS Letters*, **205**, (2), 328-332.

Muro-Pastor, M.I. and Florencio, F.J. (1994) *Journal of Bacteriology*, **176**, (9), 2718-2726.

Nafziger, E.D., Widholm, J.M., Steinrucken, H.C. and Killmer, J.L. (1984) *Plant Physiology*, **76**, 571-574.

Nakatsukasa, W.M. and Nester, E.W. (1972) *Journal of Biological Chemistry*, **247**, 5972-5979.

Nelissen, B., Van de Peer, Y., Wilmotte, A. and De Wachter, R. (1995) *Molecular Biology and Evolution*, **12**, (6), 1166-1173.

Nichols, J.M. and Adams, D.G. (1982) Akinetes. In: *The Biology of Cyanobacteria* (eds. Carr, N.G. and Whitton, B.A.), pp. 387-412, Blackwell Scientific, Oxford.

Nierzwicki-Bauer, S.A., Curtis, S.E. and Haselkorn, R. (1984) *Proc. Natl. Acad. Sci. U.S.A.*, **81**, 5961-5965.

Nimmo, G.A. and Coggins, J.R. (1981) *Biochemical Journal*, **197**, 427-436.

Niven, G.W., Kerby, N.W., Rowell, P. and Stewart, W.D.P. (1988) *Archives of Microbiology*, **150**, 272-277.

Nyhus, K.J., Sonoike, K. and Pakrasi, H.B. (1994) *Biochimica et Biophysica Acta*, **1185**, 247-251.

O'Connell, C., Pattee, P.A. and Foster, T.J. (1993) *Journal of General Microbiology*, **139**, 1449-1460.

O'Gaora, P., Maskell, D., Coleman, D., Cafferkey, M. and Dougan, G. (1989) *Gene*, **84**, 23-30.

Owttrim, G.W. and Coleman, J.R. (1987) *Journal of Bacteriology*, **169**, (5), 1824-1829.

Packer, L. and Glazer, A.N. (1988) (eds.) Cyanobacteria, *Methods in Enzymology*, **167.**

Padgette, S.R., Biest Re, D., Gasser, C.S., Eichholtz, D.A., Frazier, R.B., Hironaka, C.M., Levine, E.B., Shah, D.M., Fraley, R.T. and Kishore, G.M. (1991) *Journal of Biological Chemistry*, **266**, (33), 22364-22369.

Padgette, S.R., della Cioppa, G., Shah, D.M., Fraley, R.T. and Kishore, G.M. (1989) Selective Herbicide Tolerance through Protein Engineering. In: *Cell Culture and Somatic Cell Genetics of Plants*, Volume 6, pp. 441-476, Academic Press Inc.

Padgette, S.R., Kolacz, K.H., Delannay, X., Re, D.B., LaVallee, B.J., Tinius, C.N., Rhodes, W.K., Otero, Y.I., Barry, G.F., Eichholtz, D.A., Peschke, V.M., Nida, D.L., Taylor, N.B. and Kishore, G.M. (1995) *Crop Science*, **35**, (5), 1451-1461.

Padgette, S.R., Smith, C.E., Huynh, Q.K. and Kishore, G.M. (1988a) *Archives of Biochemistry and Biophysics*, **266**, (1), 254-262.

Padgette, S.R., Huynh, Q.K., Aykent, S., Sammons, R.D., Sikorski, J.A. and Kishore, G.M. (1988b) *Journal of Biological Chemistry*, **263**, (4), 1798-1802.

Padhy, R.N., Hottat, F.G., Coenc, M.M., Hoet, P.P. (1988) *Journal of Bacteriology*, **170** (4) 1934-1939.

Patel, V.B. and Giles, N.H. (1979) *Biochimica et Biophysica Acta*, **567**, 24-34.

Patil, R.V. and Dekker, E.E. (1990) *Nucleic Acids Research*, **18**, (10), 3080.

Peterson, H.G., Boutin, C., Martin, P.A., Freemark, K.E., Ruecker, N.J. and Moody, M.J. (1994) *Aquatic Toxicology*, **28**, 275-292.

Pinto, J.E.B.P., Dyer, W.E., Weller, S.C. and Herrmann, K.M. (1988) *Plant Physiology*, **87**, 891-893.

Pittard, J. and Wallace, B.J. (1966) *Journal of Bacteriology*, **91**, (4), 1494-1508.

Polley, L.D. (1978) *Biochimica et Biophysica Acta*, **526**, 259-266.

Porter, R.D., Buzby, J.S., Pilon, A., Fields, P.I., Dubbs, J. and Stevens, S.E. Jr. (1985) Gene Cloning and Merodiploid analysis in *Agmenellum quadruplicatum*, PR-6. In: *Vth International Symposium on Photosynthetic Prokaryotes*, Grindelwald, Switzerland. Abstract, pp. 331.

Powell, H.A. (1991) The Effects of the Herbicide Glyphosate on Cyanobacteria, *Ph.D. Thesis*, University of Dundee, Scotland, U.K.

Powell, H.A., Kerby, N.W. and Rowell, P. (1991) *New Phytology*, **119**, 421-426.

Powell, H.A., Kerby, N.W., Rowell, P., Mousdale, D.M. and Coggins, J.R. (1992) *Planta*, **188**, 484-490.

Raleigh, E.A., Benner, J., Bloom, F., Braymer, H.D., DeCruz, E., Dharmalingam, K., Heitman, J., Noyer-Weidner, M., Piekarowicz, A., Kretz, P.L., Short, J.M., Woodcock, D. (1991) *Journal of Bacteriology*, **173**, (8), 2707-2709.

Ream, J.E., Steinrucken, H.C., Porter, C.A. and Sikorski, J.A. (1988) *Plant Physiology*, **87**, 232-238.

Ream, J.E., Yuen, H.K., Frazier, R.B. and Sikorski, J.A. (1992) *Biochemistry*, **31**, 5528-5534.

Rippka, R., Deruelles, J., Waterbury, J.B., Herdman, M. and Stanier, R.Y. (1979) *Journal of General Microbiology*, **111**, 1-61.

Rippka, R. and Waterbury,J.B. (1977) *FEMS Microbiology Letters*, **2**, 83-86.

Rogers, S.G., Brand, L.A., Holder, S.B., Sharps, E.S. and Brackin, M.J. (1983) *Applied and Environmental Microbiology*, **46**, (1), 37-43.

Rubin, J.L., Gaines, C.G. and Jensen, R.A. (1984) *Plant Physiology*, **75**, 839-845.

Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A., Arnheim, N. (1985) *Science*, **230**, 1350-1354.

Sakamoto, T, Wada, H., Nishida, I., Ohmori, M. and Murata, N. (1994) *Plant Molecular Biology*, **24**, 643-650.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manuel*, Cold Spring Harbor Press, New York.

Sammons, R.D., Gruys, K.J., Anderson, K.S., Johnson, K.A. and Sikorski, J.A. (1995) *Biochemistry*, **34**, 6433-6440.

Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H. and Roe, B.A. (1980) *Journal of Molecular Biology*, **143**, 161-178.

Schmid, J. and Amrhein, N. (1995) *Phytochemistry*, **39**, (4), 737-749.

Schmidt, J., Bubunenko, M. and Subramanian, A.R. (1993) *Journal of Biological Chemistry*, **268**, (36), 27447-27457.

Schmitz, O., Kentemich, T., Zimmer, W., Hundeshagen, B. and Bothe, H. (1993) *Archives of Microbiology*, **160**, (1), 62-67.

Schulz, A., Kruper, A. and Amrhein, N. (1985) *FEMS Microbiology Letters*, **28**, 297-301.

Selvapandiyan, A., Majumder, K., Fattah, F.A., Ahmad, S., Arora, N. and Bhatnagar, R.K. (1995) *FEBS Letters*, **374**, 253-256.

Shah, D.M., Horsch, R.B., Klee, H.J., Kishore, G.M., Winter, J.A., Tumer, N.E., Hironaka, C.M., Sanders, P.R., Gasser, C.S., Aykent, S., Siegel, N.R., Rogers, S.G. and Fraley, R.T. (1986) *Science*, **233**, 478-481.

Short, J.M., Fernandez, J.M., Sorge, J.A. and Huse, W.D. (1988) *Nucleic Acids Research*, **16**, (15), 7583-7600.

Shuttleworth, W.A. and Evans, J.N.S. (1994) *Biochemistry*, **33**, 7062-7068.

Shyr, Y.Y.J., Carctto, S. and Widholm, J.M. (1993) *Plant Science*, **88**, 219-228.

Smart, C.C. and Amrhein, N. (1987) *Planta*, **170**, 1-6.

Smart, C.C., Johanning, D., Muller, G. and Amrhein, N. (1985) *Journal of Biological Chemistry*, **260**, (30), 16338-16346.

Smart, L.B. and McIntosh, L. (1991) *Plant Molecular Biology*, **17**, 959-971.

Sost, D. and Amrhein, N. (1990) *Archives of Biochemistry and Biophysics*, **282**, (2), 433-436.

Stalker, D.M., Hiatt, W.R. and Comai, L. (1985) *Journal of Biological Chemistry*, **260**, (8), 4724-4728.

Stallings, W.C., Abdel-Meguid, S.S., Lim, L.W., Shieh, H.S., Dayringer, H.E., Leimgruber, N.K., Stegeman, R.A., Anderson, K.S., Sikorski, J.A., Padgette, S.R. and Kishore, G.M. (1991) *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 5046-5050.

Stanier, R.Y. and Cohen-Bazire, G. (1977) *Ann. Rev. Biochem*, **31**, 225-274.

Steinrucken, H.C. and Amrhein, N. (1980) *Biochemical and Biophysical Research Communications*, **94**, (4), 1207-1212.

Steinrucken, H.C. and Amrhein, N. (1984a) *European Journal of Biochemistry*, **143**, 341-349.

Steinrucken, H.C. and Amrhein, N. (1984b) *European Journal of Biochemistry*, **143**, 351-357.

Steinrucken, H.C., Schulz, A., Amrhein, N., Porter, C.A. and Fraley, R.T. (1986) *Archives of Biochemistry and Biophysics*, **244**, (1), 169-178.

Stenmark, S.L., Pierson, D.L., Jensen, R.A. and Glover, G.I. (1974) *Nature*, **247**, 290-292.

Stewart, W. D. P., Haystead, A., Pearson, H.W. (1969) *Nature*, **224**, 226-228.

Tandeau de Marsac, N. and Houmard, J. (1987) Advances in Cyanobacterial Molecular Genetics. In: *The Cyanobacteria*, (eds. Fay, P. and Van Baalen, C.), pp. 251-302, Elsevier Science Publishers.

Tatum, F.M., Briggs, R.E. and Halling, S.M. (1994) *Applied and Environmental Microbiology*, **60**, (6), 2011-2016.

Thiel, T. (1993) *Journal of Bacteriology*, **175**, (19), 6276-6286.

Thiel, T. (1994) Genetic Analysis of Cyanobacteria. In: *The Molecular Biology of Cyanobacteria* (ed. Bryant, D.A.), pp. 581-611, Kluwer Academic Publishers, The Netherlands.

Tumer, N.E., Robinson, S.J., and Haselkorn, R. (1983) *Nature*, **306**, 337-342.

Valverde, F., Losada, M., Serrano, A. (1995) Cloning by Functional Complementation in *E.coli* of the *gap2* gene of *Synechocystis* sp. PCC 6803 Supports an Amphibolic Role for Cyanobacterial NAD(P)-Dependent Glyceraldehyde-3-Phosphate Dehydrogenase. In: *Proceedings of the Xth International Photosynthesis Congress* 1-1 (ed. Mathis, P.), Kluwer Academic Publishing, The Netherlands.

Vaughan, L.M., Smith, P.R. and Foster, T.J. (1993) *Infection and Immunity*, **61**, (5), 2172-2181.

Wada, H., Gombos, Z. and Murata, N. (1990) *Nature*, **347**, 200-203.

Wang, Y., Jones, J.D., Weller, S.C. and Goldsbrough, P.B. (1991) *Plant Molecular Biology*, **17**, 1127-1138.

Weber, H.L. and Bock, A. (1968) *Archives of Microbiology*, **61**, 159-168.

Weedon, N.F. (1981) *Journal of Molecular Evolution*, **17**, 133-139.

Weiss, B., Jacquemin-Sablon, A., Live, T.R., Fareed, G.C. and Richardson, C.C. (1968) *Journal of Biological Chemistry*, **243**, (17), 4543-4555.

Wetmur, J.G. and Davidson, N. (1968) *Journal of Molecular Biology*, **31**, 349-370.

White, P.J., Miller, G. and Coggins, J.R. (1988) *Biochemical Journal*, **251**, 313-322.

Whitton, B.A. and Carr,N.G. (1982) Cyanobacteria: Current Perspectives. In: *The Biology of Cyanobacteria* (eds. Carr, N.G. and Whitton, B.A.), pp. 1-8, Blackwell Scientific, Oxford.

Wilmotte, A. (1994) Molecular Evolution and Taxonomy of the Cyanobacteria. In: *The Molecular Biology of Cyanobacteria* (ed. Bryant, D.A.), pp. 1-25, Kluwer Academic Publishers, The Netherlands.

Wolk, C.P. (1982) Heterocysts. In: *The Biology of Cyanobacteria* (eds. Carr, N.G. and Whitton, B.A.), pp. 359-386, Blackwell Scientific, Oxford.

Wolk, C.P., Cai, Y., Cardemil, L., Flores, E., Hohn, B., Murray, M., Schmetterer, G., Schrautemeier, B. and Wilson, B. (1988) *Journal of Bacteriology*, **170**, 1239-1244.

Yanisch-Perron,C., Vieira, J. and Messing, J. (1985) *Gene*, **33**, 103-119.