THE UNIVERSITY OF
WARWICK

University of Warwick institutional repository: http://go.warwick.ac.uk/wrap

A Thesis Submitted for the Degree of PhD at the University of Warwick

http://go.warwick.ac.uk/wrap/3999

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it. Our policy information is available from the repository home page.

A Study of the regulation of undecylprodigiosin biosynthesis in *Streptomyces coelicolor* A3(2).

Christine Susan Flaxman BSc. (Hons.)

Thesis submitted for the Degree of Doctor of Philosophy

University of Warwick

Department of Biological Sciences

September 1995

A Study of the regulation of undecylprodigiosin biosynthesis in *Streptomyces coelicolor* A3(2).

Christine Susan Flaxman BSc. (Hons.)

Thesis submitted for the Degree of Doctor of Philosophy

University of Warwick

Department of Biological Sciences

September 1995


A Study of the
regulation of
undecylprodigiosin
biosynthesis in
Streptomyces
coelicolor A3(2)

Table of Contents

List of Figures vii

List of Tables x

Acknowledgements xi

Declaration xiii

Summary xv

1 Introduction:

1.1 Taxonomy 1

1.2 Life cycle 1

1.3 Genetic techniques 2

1.4 Molecular Genetic techniques 4

1.4.1 Plasmid vectors 4

1.4.2 Bacteriophage vectors 5

1.5 Genetic analysis of antibiotic biosynthesis in *S. coelicolor* A3(2) 6

1.5.1 Actinorhodin 7

1.5.2 Methylenomycin 8

1.5.3 Undecylprodigiosin 9

1.5.4 Calcium dependent antibiotic 16

1.5.5 The *whiE* cluster encoded spore associated pigment 16

1.6 A-factor 17

1.6.1 Regulation of streptomycin biosynthesis in *Streptomyces griseus* by A-factor 17

1.6.2 Other Autoregulator molecules in *Streptomyces* 18

1.6.3 Autoregulator molecules in other organisms 19

1.7 Genes with pleiotropic effects on the production of antibiotics 20

1.7.1 *afsB*, *afsR* and *afsK* 21
1.7.2 *afsQ1* and *afsQ2* 22
1.7.3 *afsR2* 23
1.7.4 *bld* mutants 23
1.7.4.1 *bldA* mutants 24
1.7.5 Antibiotic activator (*aba*) mutants 27
1.7.6 Antibiotic synthesis (*abs*) mutants 27
1.8 Physiological factors that effect antibiotic production 30
1.8.1 Stringent response 30
1.8.2 Catabolite repression 31
1.9 Interaction between primary and secondary metabolism 33
1.9.1 Involvement of proline in osmoregulation 34
1.9.2 Proline transport 36
1.9.3 Proline tranport in enteric bacteria 39
1.9.3.1 Proline porter I encoded by *putP* 39
1.9.3.2 Proline porter II encoded by *proU* 40
1.9.3.3 Proline porter III encoded by *proU* 40
1.9.4 Proline transport in Gram positive organisms 41
1.9.5 Proline utilisation 42
1.9.6 Proline synthesis 44
1.9.7 Proline synthesis in other organisms 46
1.9.8 Proline metabolism in *S. coelicolor* 47
1.10 Aims of this project 48

2 Methods and Materials:

2.1 Growth and Maintainance of bacterial strains 49

2.1.1 Growth of *Streptomyces* 54

2.2 Media 55

2.2.1 *E.coli* media 55

2.2.2 *Streptomyces* media 56

2.3 Proline permease assays 58

2.3.1 Growth on proline as sole nitrogen source 58

2.3.2 Resistance to proline analogues 58

2.3.3 ^{14}C proline transport 59

2.3.4 Determination of the counting efficiency of ^{14}Cproline 60

2.3.5 Calculation to determine the amount of ^{14}C proline transported into the cell 60

2.4 Transformation of *E.coli* and *Streptomyces* 61

2.4.1 Calcium chloride transformation of *E.coli* 61

2.4.2 DMSO transformation of *E.coli* 62

2.4.3 Electroporation of *E.coli* 63

2.4.4 Transformation of *Streptomyces* 63

2.4.5 Preparation of lysogens 63

2.5 Molecular techniques 64

2.5.1 Restriction endonuclease digestion of DNA 64

2.5.2 Agarose gel electrophoresis 64

2.5.3 Dephosphorylation of vector DNA 64

2.5.4 DNA Ligation 65

2.5.5 Fragment end repair 65

2.6 DNA Preparation 65

2.6.1 *E.coli* plasmid preparation 65

2.6.2 *Streptomyces* plasmid preparation 65

2.7 Di-deoxy DNA sequencing 66

2.7.1 Single stranded template preparations 66

2.7.2 Sequencing reactions 66

2.8 S1 nuclease mapping 66

3 Study of proline transport mutants:

3.1 Introduction 68

3.1.1 Isolation of proline transport mutants 68

3.1.2 Phenotype of proline transport mutants 71

3.2 Interaction between Put phenotype and Red biosynthesis 72

3.2.1 Strategy for isolating a Red$^-$ Put$^-$ mutant 74

3.2.2 Isolation of Put$^-$ Red$^-$ lysogens 77

3.3 Characterisation of Put$^-$ Red$^-$ lysogens 77

3.3.1 Growth with proline as sole nitrogen source 77

3.3.2 Resistance to proline analogues 83

3.3.3 Transport of ^{14}C proline 83

3.3.4 Inactivation of *redD* in a Put$^-$ strain 88

3.3.5 S1 nuclease mapping of *redD* and *redX* 93

3.4 Discussion 100

4 DNA sequence and DNA sequence analysis of the Pwb region

4.1 Introduction 102

Isolation of Pwb mutants 102

4.2 Cloning of the 800bp EcoRI regions containing the Pwb mutations 106

4.2.1 Cloning of the 800bp *Eco*RI fragment of pIJ2530 106

4.2.2 Cloning of the 800bp *Eco*RI fragment of pIJ2520 and pIJ2540 108

4.2.3 Cloning of the wild-type 800bp *Eco*RI fragment of the *red* cluster 108

4.3 Sequencing Strategy 110

4.3.1 Construction of deletion clones of ØCF1 and ØCF2 110

4.3.2 Analysis of the sequence data obtained of the 800bp *Eco*RI fragment of *pwb6* 113

4.3.3 Cloning of DNA upstream of ORF1 117

4.3.4 Strategy used to sequence ØCF3, ØCF4, ØCF5, ØCF6, ØCF7, ØCF8 120

4.4 Sequencing Data 120

4.5 Analysis of sequence data 123

4.6 Similarity studies 136

4.7 Discussion 142

5 Characterisation of RedZ

5.1 Introduction 145

5.2 Isolation of a *redD* lysogen in the Pwb-9 mutant 146

5.3 Isolation of a Pwb lysogen in a *redD* mutant 152

5.4 Interaction of *redZ* with other regulatory genes 155

5.4.1 Isolation of Pwb lysogens of AbsA and AbsB mutants 157

5.5 Discussion 159

6 Concluding remarks

6.1 Concluding remarks and future work 161

Appendix 1

Table of restriction sites within the *Bgl*II *Eco*RI *redZ* region 170

Sequence data of *pwb* + 185

Sequence data of *pwb-6* 188

References 194

List of Figures

Figure 1.1	Genetic and physical map of the *S. coelicolor* A3(2) chromosome	3
Figure 1.2	Chemical structure of prodigiosin and undecylprodigiosin	10
Figure 1.3	Pathway of Prodigiosin Biosynthesis	11
Figure 1.4	Organisation of the Red cluster	13
Figure 1.5	Map of the right hand end of the Red cluster	15
Figure 1.6	Biosynthesis of tylosin in *S. fradiae*	32
Figure 1.7	Chemical structure of proline, proline betaine and glycine betaine	35
Figure 1.8	Chemical structure of proline and proline analogues	38
Figure 1.9	The pathway of proline utilisation	43
Figure 1.10	The proline biosynthetic pathway	45
Figure 3.1	Isolation of the Put$^-$ mutants	70
Figure 3.2	Insertional Mutagenesis	75
Figure 3.3	Photograph of Act$^-$ Red$^-$ US218 Lysogens	78
Figure 3.4	Photograph of J802 and US224 growth on Syton silica medium with proline as sole nitrogen source	82
Figure 3.5	Resistance of US224 and J802 to proline analogues	84
Figure 3.6	Resistance of US218 and US224 lysogens to the proline analogue DHP	85
Figure 3.7	^{14}C Proline transport of US218, US224 and J802	86
Figure 3.8	^{14}C Proline transport of US218 lysogens and J802	87
Figure 3.9	^{14}C Proline transport of US224 lysogens and J802	89
Figure 3.10	Proline analogue resistance of *redD* US224 lysogen	91
Figure 3.11	^{14}C Proline transport of *redD* US224 lysogen	92

Figure 3.12 Growth curve of J802 and US224 from which RNA was isolated 94
Figure 3.13 Representation of the *redD* DNA used as a probe in S1 nuclease analysis 95
Figure 3.14 Representation of *redX* DNA used as a probe in S1 nuclease analysis 96
Figure 3.15 Autoradiograph showing S1 nuclease protected probes 97
Figure 3.16 ^{14}C Proline transport of cells used to isolate RNA, for S1 nuclease analysis 99
Figure 4.1 Location of the DNA inserts of plasmids pIJ2520, pIJ2530 and pIJ2540 105
Figure 4.2 Restriction analysis of pPWB6 109
Figure 4.3 Restriction map to the location of Pwb^+ clones 111
Figure 4.4 Restriction map to show clones and deletion clones used in DNA sequencing 112
Figure 4.5 Restriction map to show the location of oligonucleotides used in DNA sequencing 115
Figure 4.6 FRAME Analysis of 800bp *Eco*RI fragment 116
Figure 4.7 Cloning strategy used to obtain 1.7kb *Bam*HI fragment of the *red* cluster 118
Figure 4.8 Photograph of restriction digests of ØCF9 and ØCF10 119
Figure 4.9 Orientation of the DNA inserts of ØCF9 and ØCF10 121
Figure 4.10 DNA Sequence obtained for Pwb-9, Pwb-16 and Pwb^+ 122
Figure 4.11 Nucleotide and predicted amino acid sequence of *redZ* 124

Figure 4.12 FRAME analysis of *redZ* 126
Figure 4.13 Restriction map of *redZ* 128
Figure 4.14 DNA Sequence of Pwb$^+$ showing the in-frame TTA codon. 131
Figure 4.15 DNA Sequence of Pwb-6 showing the in-frame TTA codon 132
Figure 4.16 DNA Sequence of Pwb-9 showing the in-frame TTA codon 133
Figure 4.17 DNA Sequence of Pwb-16 showing the in-frame TTA codon 134
Figure 4.18 DNA Sequence showing the Pwb-6 mutation 135
Figure 4.19 Alignment of RedZ to other regulators. 138
Figure 4.20 Alignment of RedZ with response regulators 139
Figure 5.1 Pwb-9 lysogens of ØKC899 (!RedDi) 148
Figure 5.2 Pwb-9 ØKC899 lysogen and progenitor strains Pwb-9 and J1700 149
Figure 5.3 Pwb-9 ØKC899 lysogens produce sectored and white colonies 150
Figure 5.4 Location of DNA inserts of Ø2 and Ø3 153
Figure 5.5 Lysogens of AbsA and AbsB with Ø2 158

List of Tables

Table 2.1	Bacterial strains	49
Table 2.2	Plasmids and bacteriophage	52
Table 3.1	A Summary of the origin and phenotype of Put⁻ mutants	73
Table 3.2	Number of lysogens isolated	79
Table 3.3	Summary of proline transport tests performed on lysogens	81
Table 4.1	Isolation and cloning of Pwb mutants	103
Table 4.2	A Summary of clones and deletion clones used in DNA sequencing	107
Table 4.3	Sequence of oligonucleotides used for DNA sequencing	114
Table 4.4	Similarity of RedZ to response regulators	137

Acknowledgements

First and foremost I would like to thank Dr. Dave Hodgson for his enthusiasm, his patience in explaining the same concept to me time and time again, for his help in checking this thesis and for all the Uncle Joes mintballs.

I also want to thank Dr. M. Bibb for allowing me to visit his lab to learn the technique of S1 nuclease mapping and for performing the FRAME analyses. Thanks also for allowing me to use *red* cluster maps and *redZ* gene alignments and for many useful discussions. Thanks are also due to Janet White for her help with S1 mapping and for allowing me to discuss unpublished results in this thesis. I would like to thank the members of SL1.48 especially Leony, Rekha, Wendy, Mark and Janet who helped make my time at the John Innes so enjoyable.

A special thank you is due to Susie Wharam for her company as one of my bay-mates, for proof reading and for her sound advice through the course of this project. I would also like to thank Vince 'fountain-of-all-knowledge' Mulholland for his advice and company, the lab isn't the same without you. Thanks also to the members of Micro II and Environmental past and present with whom I have shared the past four years: to Vicky, Woody, Paul Wan, the newly married Mrs Jo Thomas, Terry-la for the impressive oakridge tube and eppendorf bombs and car-surfing, Phil Reeves also of oakridge bomb fame, Cathy-la, Oona, the sequencing king Mo, Charlotte, Derek, Uthaya, Denise, Nick Housby, Matt, the lab nightingale Nick Thompson, Tony, Sue, Fahim, Heather, Dawn, Daniel, Andy, Doug, Willie, Kev, Pete, Liesa, Neil, Jo, Jane Green and

others. I also want to thank Ian Brock for the loan of a printer, it made a huge difference.

Thanks also to friends (Julie, Liesa, Ruth, Charlie and Dan) and family who have been a constant support.

Last but by no means least thanks to Jonathan for keeping me sane through the many rough patches with my work and for your love and support, I couldn't have done it without you.

Declaration

The work described in this thesis is the result of research conducted by myself under the supervision of Dr. D. A. Hodgson. In the instances where others have contributed to the work, specific acknowledgments have been made.

None of the information contained in this thesis has been used in any previous application for a degree.

Christine Flaxman

Christine Flaxman

Summary

Undecylprodigiosin is one of four secondary metabolites with antibacterial activity produced by *S. coelicolor* A3(2). The overall aim of this study was to further investigate the control of biosynthesis of the secondary metabolite undecylprodigiosin (Red) in *Streptomyces coelicolor* A3(2).

Proline transport mutants (Put^-) were isolated, and the over-production of Red was observed in these strains. It was hypothesised that Red biosynthesis is essential as a shunt for excess proline in the Put^- mutants. Red biosynthesis was abolished by disrupting the *redX* structural gene in a Put^- mutant. The Put^- $RedX^-$ mutants were viable, demonstrating that Red is not essential in Put^- mutants. S1 nuclease mapping of *redD* and *redX* genes in a Put^- mutant revealed that *red* genes are transcribed earlier in the growth phase of Put^- mutants compared to the progenitor strain J802.

Pwb (pigmented whilst *bld*) mutants had been isolated due to their ability to produce Red in a *bldA* background. The regions believed to contain the mutations of Pwb-6, Pwb-9, Pwb-16 and Pwb^+ were sub-cloned and sequence data obtained. An open reading frame was identified which is predicted to encode a protein showing homology to the UhpA-LuxR family of regulators. The open reading frame, contains an in-frame TTA codon. It is proposed that this gene, named *redZ*, mediates the *bldA* dependence of Red biosynthesis. The Pwb-6 mutation was located to the putative $^-35$ promoter region. The mutation makes the promoter more similar to the enteric bacterium major sigma factor promoter $^-35$ consensus sequence. It is anticipated that greater transcription from the promoter causes the Pwb phenotype. Introduction of the Pwb-9 *redZ* gene into antibiotic biosynthesis mutants, *absA* and *absB*, did not result in Red biosynthesis.

List of Abbreviations

Act	Actinorhodin
AMP	adenosine monophosphate
ATP	adenosine 5' triphosphate
AZC	Azetidine-2-carboxylate
bla	β-lactamase gene
bp	base pairs
Cda	Calcium-dependent antibiotic
DHP	3,4-dehydroproline
DMSO	Di methylsulphoxide
DNA	deoxyribonucleic acid
DNR	Daunorubicin
IPTG	Isopropyl β-D thiogalactopyranoside
J.I.I	John Innes Institute
kb	kilo base pairs
Mmy	Methylenomycin
OD_{600}	Absorbance at 600nm
OHHL	N-(3 oxohexanoyl)-L-homoserine lactone
P5C	Pyrrolidine-5-carboxylic acid
PEG	Polyethylene glycol
ppGpp	Guanosine 3', 5'-Bispyrophosphate
PU	used to describe proline utilising mutants
PUM	a class of proline transport mutants
Put	used to describe the phenotype of proline transport mutants
Pwb	used to describe the phenotype of mutants of *bldA* that have aquired the ability to produce Red
Red	Undecylprodigiosin
redZ^{P}	used to describe the *redZ* gene from a Pwb mutant
RHO	Rhodomycinone
RNA	ribonucleic acid
X-Gal	5-bromo-4-chloro-3 indolyl-galactosidase

CHAPTER 1

INTRODUCTION

1.1 Taxonomy

Streptomyces is a genus of Gram positive procaryotes in the order Actinomycetales. The actinomycetes are soil dwelling organisms that have a characteristic high G+C DNA content. Within the actinomycetes there is a group of organisms, the sporoactinomycetes, that are capable of producing exospores and display a complex life cycle. The genus *Streptomyces* is a member of the sporoactinomyces.

1.2 Life cycle

Streptomycetes have a life-cycle surprisingly complex for a bacterium, in part existing as a multicellular organism with differentiated cell types. The life cycle begins with germination and outgrowth of a primary germ tube from the spore. Growth and branching of the primary hyphae leads to the formation of a mat of interconnecting tube-like vegetative cells, called the substrate mycelium. Growth of the substrate mycelium occurs by cell wall extension at the hyphal tips. The substrate mycelium grows over and into the surface of the growth medium. Exoenzymes are secreted which degrade macromolecules, the soluble products released are transported into the cell to support growth. At some unknown stimulus, vegetative growth slows and aerial hyphae grow vertically from the substrate mycelium. There is an association with the initiation of aerial mycelium formation and the production of secondary metabolites. In contrast to the substrate mycelium that contains few oblique cellular cross walls, aerial hyphae produce regular, transverse septa which include one nucleoid per cell compartment. A spore is formed from each cellular compartment of the septated aerial hyphae. Spores are the dispersal phase of the life-cycle. It is believed that the hydrophobicity of the

spores aids their dispersal on the surface of water droplets. Streptomycete spores are resistant to desiccation, but not to heat or ultraviolet irradiation. For a more detailed description of streptomycete differentiation see Chater (1993) and Hodgson (1992).

Streptomyces spp. are unusual bacterium not only because of the complexity of their life-cycle but also because they produce vast numbers of secondary metabolites, many of which are commercially important compounds; such as the anti-helminitic agent avermectin, the herbicide bialaphos and the antibiotics streptomycin, tetracyclins and the β-lactamase inhibitor, clavulanic acid. This novel characteristic makes the members of the *Streptomyces*, not only interesting organisms to research, but also commercially important.

1.3 Genetic techniques

Streptomyces coelicolor A3(2) is genetically the best characterised strain within the genus. *S. coelicolor* A3(2) shows more similarity to *Streptomyces violaceoruber* in the taxonomic group 21 of Williams *et al.*, (1983) than to true *S. coelicolor* (Müller). In this thesis the name *S. coelicolor* A3(2) will be used. A genetic and physical map of the *S. coelicolor* chromosome has been constructed (Kieser *et al.*, 1992). Findings of Lin *et al.* (1993) provides evidence that the *S. lividans* chromosome is linear. Re-examination of genetic cross and restriction mapping data revealed that the *S. coelicolor* chromosome is also linear (Figure 1.1). There are four secondary metabolites produced by *S. coelicolor* that have antibacterial activity; these are actinorhodin (Act), undecylprodigiosin (abbreviated to Red due to its colour), methylenomycin (Mmy) and a calcium-dependent-antibiotic (Cda). The other secondary

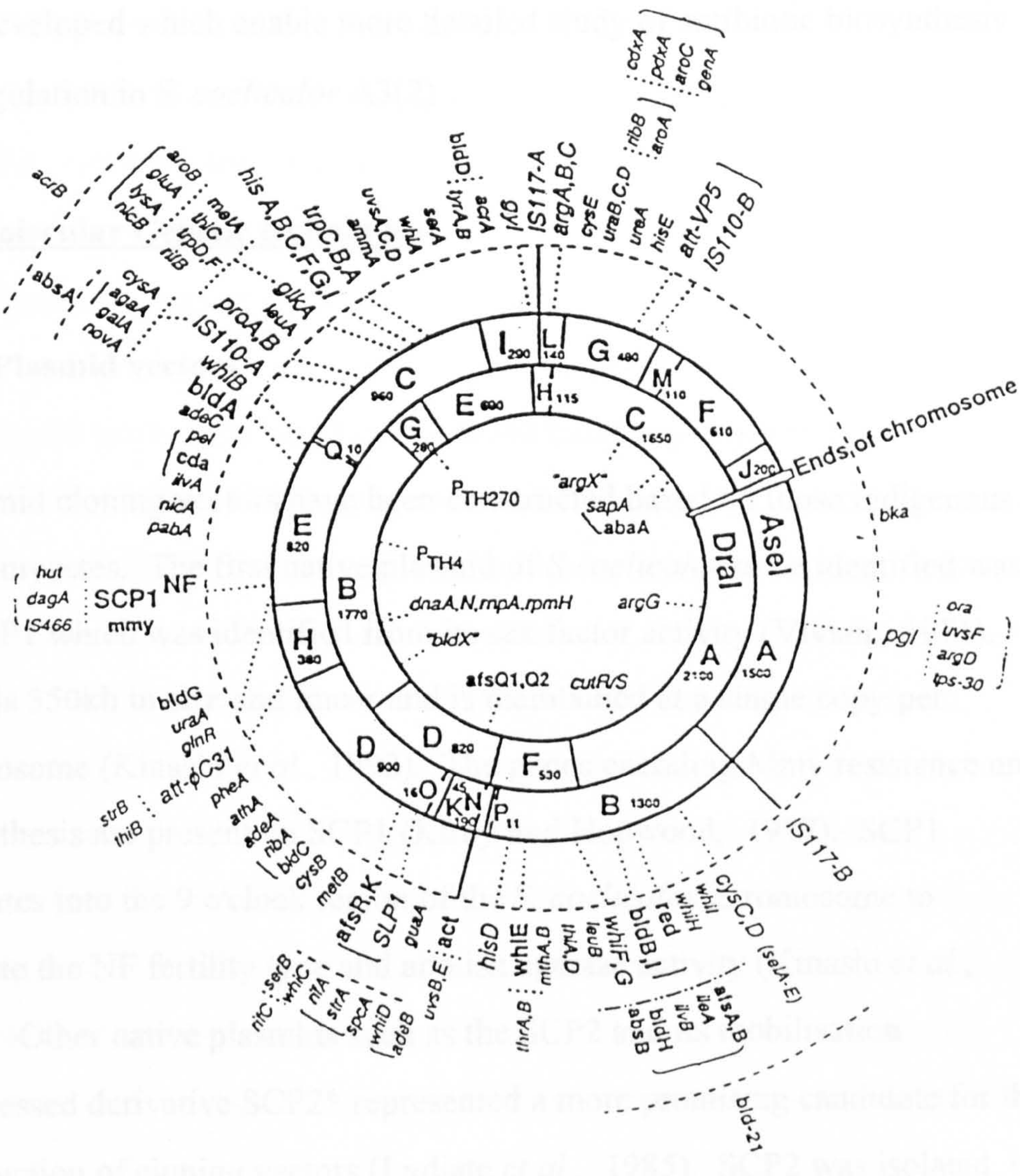
Figure 1.1

Genetic and Physical map of the *S. coelicolor* A3(2) chromosome

The physical map of the *S. coelicolor* A3(2) derivative M145 for *AseI* and *DraI* is shown overleaf in bold circles, restriction fragments are lettered and sizes are in kb. The outermost circle shown in hatched lines represents the genetic map. Groups of loci of undetermined order are bracketed.

Figure 1.1

Genetic and Physical map of the *S. coelicolor* A3(2) chromosome



Reproduced from Hopwood *et al.*, 1994

metabolites include a brown spore-associated pigment and an A-factor-like butyrolactone. The A-factor-like molecule is necessary for both sporulation and streptomycin production in *S.griseus*, for a review see Horinouchi and Beppu (1994). The genetic map locations of the gene clusters involved in secondary metabolism are shown on Figure 1.1. Molecular techniques have been developed which enable more detailed study of antibiotic biosynthesis and regulation in *S. coelicolor* A3(2).

1.4 Molecular Genetic techniques

1.4.1 Plasmid vectors

Plasmid cloning vectors have been constructed based on those indigenous to streptomycetes. The first native plasmid of *S.coelicolor* to be identified was the SCP1 which was identified from its sex factor activity (Vivian, 1971). SCP1 is 350kb in size and linear and is maintained at a single copy per chromosome (Kinashi *et al.*, 1992). The genes encoding Mmy resistance and biosynthesis are present on SCP1 (Kirby and Hopwood, 1977). SCP1 integrates into the 9 o'clock region of the *S. coelicolor* chromosome to generate the NF fertility type and abolish agarase activity (Kinashi *et al.*, 1992). Other native plasmids such as the SCP2 and its mobilisation derepressed derivative SCP2* represented a more promising candidate for the construction of cloning vectors (Lydiate *et al.*, 1985). SCP2 was isolated from *S. coelicolor* by Schrempf and Goebel (1977). SCP2* is stably maintained at a copy number of one or two per chromosome and has a broad host range within the genus. Plasmid constructs pIJ2520, pIJ2530 and pIJ2540 used in this study (Chapter 4) were constructed from an SCP2* derived vector. Other native plasmids such as the *S.lividans* 66 SLP1 which

can only be used in *S. coelicolor* A3(2) and the multicopy plasmid pIJ101 have been used as the basis for several cloning vectors (Hopwood *et al.*, 1985).

In addition to simple cloning vectors, more sophisticated *E.coli/Streptomyces* shuttle vectors have been constructed that allow DNA to be manipulated in *E.coli* then transformed into *Streptomyces* spp. (Mazodier *et al.*, 1989; Gormley and Davies, 1991).

1.4.2 Bacteriophage vectors

Bacteriophage vectors based on the broad host range temperate *Streptomyces* phage ØC31 have proved useful. Many ØC31 derived vectors have been constructed that vary in the availability of restriction sites and in their capacity for inserted DNA (Hopwood *et al.*, 1985). ØC31 has a genome of 41.5kb, and much of its function has been characterised (Lomoskaya, 1980; Smith *et al.*, 1992). The left half of the genome encodes functions used late in the life cycle. The *c* repressor gene, which represses the lytic cycle and therefore maintains lysogeny, is located centrally. Right of the *c* gene there is 10kb of DNA believed to be involved in early functions. There is 8kb which appears to be non-essential in lytic growth. Included in this region is the *attP* site which is required for prophage integration into the host genome.

One important consideration when using ØC31 based vectors in *S. coelicolor* is the Pgl (Phage growth limiting) status of the host strain. When a Pgl^+ strain of *S.coelicolor* is infected with ØC31-derived phage the first cycle of infection occurs normally. The progeny phage released after the first

cycle of infection are modified in some way by the Pgl system, which reduces the capacity of these phage to infect a Pgl^+ host. Laity *et al.* (1993) identified at least two classes of Pgl mutants, although the exact nature of the modification performed by the Pgl system is unknown. *S. lividans* is naturally Pgl^-.

Bruton *et al.* (1991) reported a ØC31 derived phage that can be used as a promoter probe as it contains a promoterless *xylE* reporter gene. The *xylE* gene encodes the catechol 2,3-dioxygenase enzyme that catalyses the conversion of colourless catechol to the yellow hydroxymuconic semialdehyde. By insertion of DNA upstream of the promoterless *xylE* gene the promoter activity can be assessed by the XylE activity.

Another valuable use of the ØC31 based vectors is the insert directed gene disruption technique called insertional mutagenesis (Chater and Bruton, 1983). Cloning of a DNA fragment internal to a gene into an $attP^-$ derivative of ØC31, directs insertion of the phage into the target gene on the chromosome via homologous recombination and hence allows the lytic mutant to lysogenise the host. The addition of antibiotic resistance genes allows selection for lysogens. This system was first exploited to clone DNA fragments containing Mmy biosynthesis and resistance determinants (Chater and Bruton, 1983).

1.5 Genetic analysis of secondary metabolite biosynthesis in *S.coelicolor* A3(2)

Use of the above vector has assisted the cloning of the Act, Red, Mmy and the brown spore associated pigment (*whiE*) gene clusters of *S.coelicolor* A3(2). As a general rule genes involved in biosynthesis and resistance to

secondary metabolites are clustered. Many other secondary metabolite gene clusters have also been cloned for example: the streptomycin genes from *S.griseus* (Distler *et al.*, 1987), the tetracenomycin genes of *S.glaucescens* (Decker and Hutchinson, 1993) and the bialaphos genes from *S.hygroscopicus* (Murakami *et al.*, 1986).

1.5.1 Actinorhodin

Act is a polyketide antibiotic derived from acetate and malonate precursors. Act is blue pigmented in very alkaline conditions (pH11), however, recent studies have shown the blue pigment sometimes seen on agar plates at physiological pH is an actinorhodin derivative γ actinorhodin (Hopwood *et al.*, 1994). The *act* cluster was isolated by complementation of a blocked mutant with large DNA fragments cloned in a low copy number vector. Three clones that restored blue pigmentation after ammonia fuming were obtained and contained overlapping DNA fragments (Malpartida and Hopwood, 1984). The entire cluster was cloned onto a single vector by re-ligation of the DNA fragments of two clones to create pIJ2303. Analysis of this clone has determined that a 21kb segment is required for Act biosynthesis in *Streptomyces parvulus* (Malpartida and Hopwood, 1984). The DNA sequence of the 21kb fragment has been obtained and 23 open reading frames identified. The function of genes involved in export, biosynthesis and regulation of Act have been determined, largely by open reading frame product sequence comparison (Caballero *et al.*, 1991; Hallam *et al.*, 1988; Fernández-Moreno *et al.*, 1991, 1992a).

The regulation of Act biosynthesis is of particular interest to this study, for a detailed description of Act export and biosynthesis see Hopwood *et al.* (1994).

Original cosynthesis experiments of Act⁻ mutants identified a class of regulatory mutants (ActII) that did not co-synthesise with any other class (Rudd and Hopwood, 1979). The *actII*ORF4 gene, when present in multicopy, caused over-production of Act (Fernández-Moreno *et al.*, 1991). The *actII*ORF4 gene product shows end to end similarity with the Red activator *redD* (see below). When the *actII*ORF4 gene was disrupted the Act$^+$ strain J1501 became Act⁻ (Fernández-Moreno *et al.*, 1991). These results led to the conclusion that *actII*ORF4 is an Act pathway activator

1.5.2 Methylenomycin

The methylenomycin (Mmy) genes were the first secondary metabolite gene cluster to be cloned from *S. coelicolor*. The SCP1 plasmid, on which the Mmy genes are located, could not be purified in sufficient quantities to allow cloning of the Mmy genes directly (Chater and Bruton, 1985). The Mmy genes were cloned using ØC31-derived phage vectors. A number of ØC31-derived phage (ØKC400) containing SCP1 DNA were isolated. Analysis of the DNA of these clones showed that at least 17kb of contiguous DNA was required for Mmy biosynthesis and resistance (Chater and Bruton, 1985). Insert directed integration into the Mmy cluster identified the function of four regions. Integration into the left hand end of the cluster resulted in over-production of Mmy (Chater and Bruton, 1985). This suggested the inactivation of a repressor or the duplication of an activator gene. Deletion of the left hand end of the cluster resulted in Mmy over-production which provides evidence that this region encodes a repressor (*mmyR*). The DNA sequence of the repressor has been determined and similarity to the repressor of Act export (*act*II-ORF1) was identified (reviewed by Hopwood *et al.*, 1994).

The genetic basis of Mmy resistance has been identified from DNA sequence analysis of the *mmr* resistance gene. The *mmr* predicted gene product is a hydrophobic protein containing several transmembrane spanning domains and has homology with the TetL and TetK proteins that confer tetracyline resistance in *Bacillus* and *Staphlococcus* spp. (Hopwood *et al.*, 1994).

1.5.3 Undecylprodigiosin

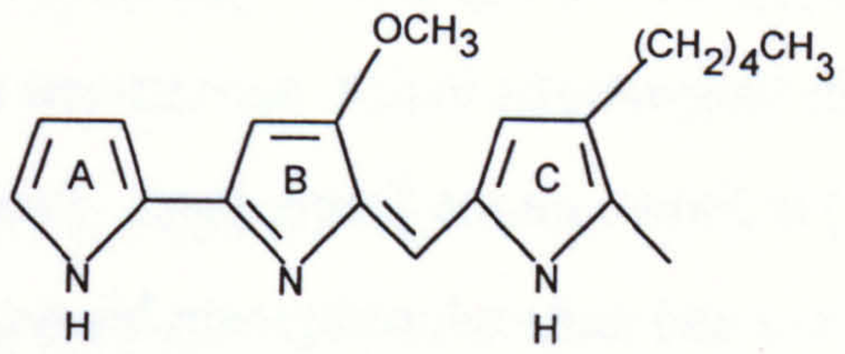
Undecylprodigiosin (Red) is a non-diffusible red mycelial pigment with limited antibacterial activity and some immunosuppressive properties (Tsao *et al.*, 1985). The red pigment produced by *S. coelicolor* has been shown to be a mixture of prodigionines of which Red is the major constituent. The chemical structure of Red is similar to the antibiotic prodigiosin of *Serratia marcescens* (Figure 1.2). Biosynthesis of prodigiosin by *Ser. marcescens* proceeds via a bifurcated pathway (Figure 1.3). One branch of the pathway produces a bipyrrole aldehyde, the other branch produces a substituted monopyrrole. Condensation of the monopyrrole and bipyrrole results in the formation of prodigiosin (Wasserman, 1974). The biosynthesis of undecylprodigiosin in *S. longisporus ruber* occurs by a similar pathway to that in *Ser. marcescens*.

Labelling studies using ^{13}C demonstrated that acetate, glycine/serine and proline are precursors incorporated into Red as shown in Figure 1.2 (Wasserman *et al.*, 1974). This led to the investigation of the origin of the proline molecule incorporated into Red as discussed later. Mutants of *S. coelicolor* that were unable to produce Red were isolated (Rudd and Hopwood, 1980). The Red$^-$ mutants were grouped on the basis of

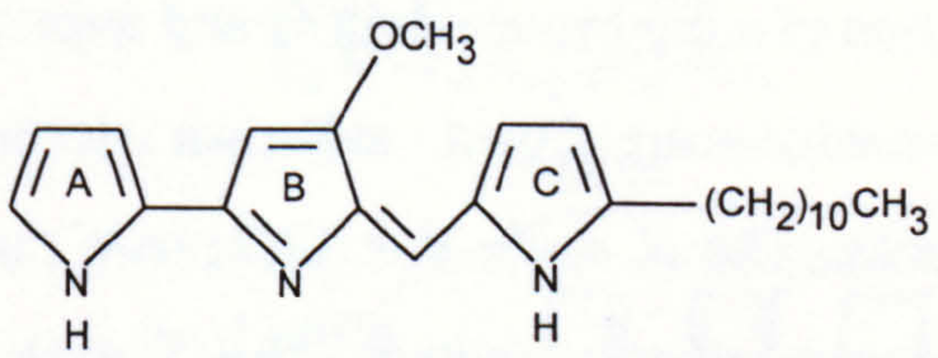
Figure 1.2

Chemical structure of Prodigiosin and Undecylprodigiosin

A



Prodigiosin



Undecylprodigiosin

B

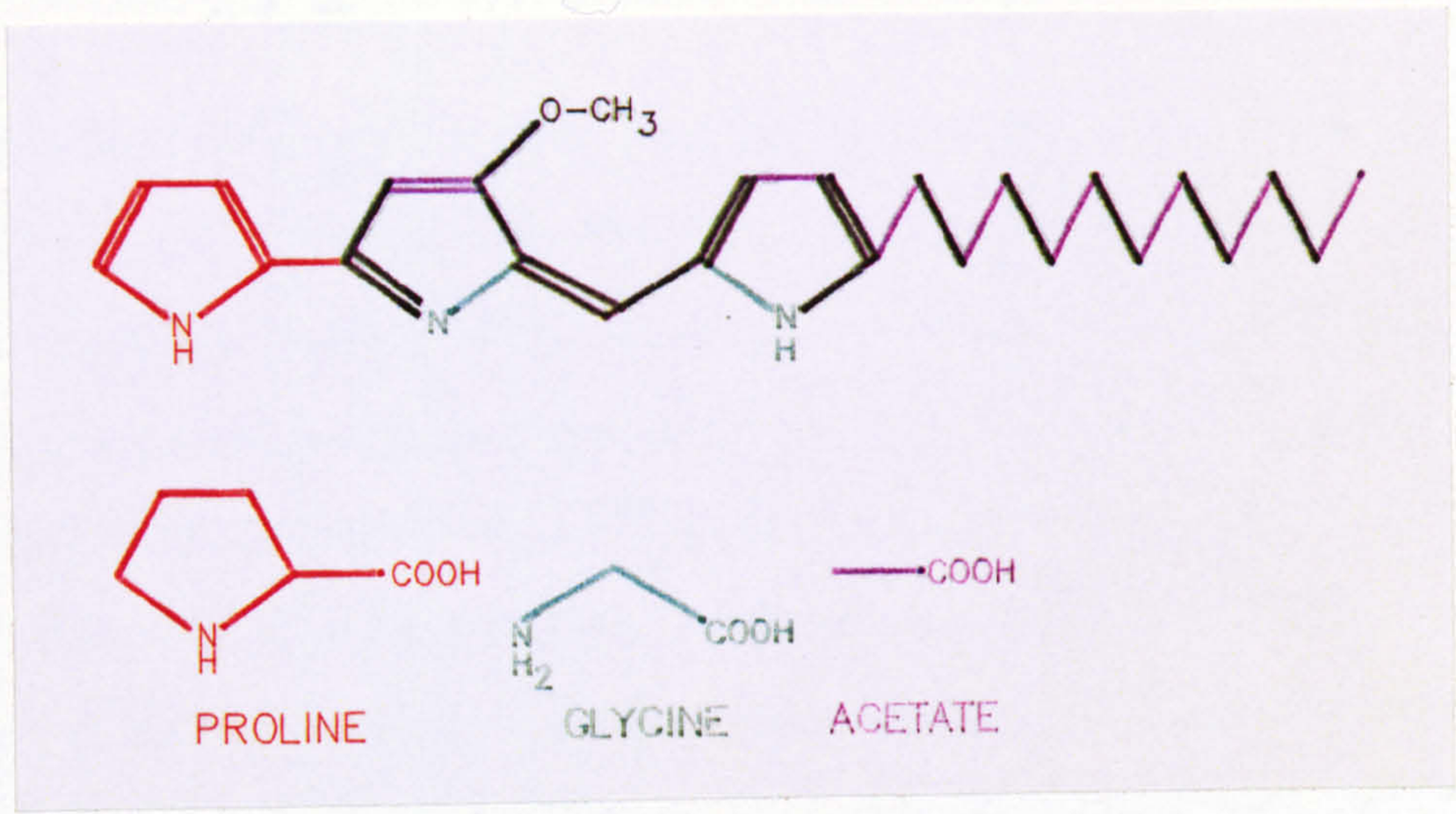
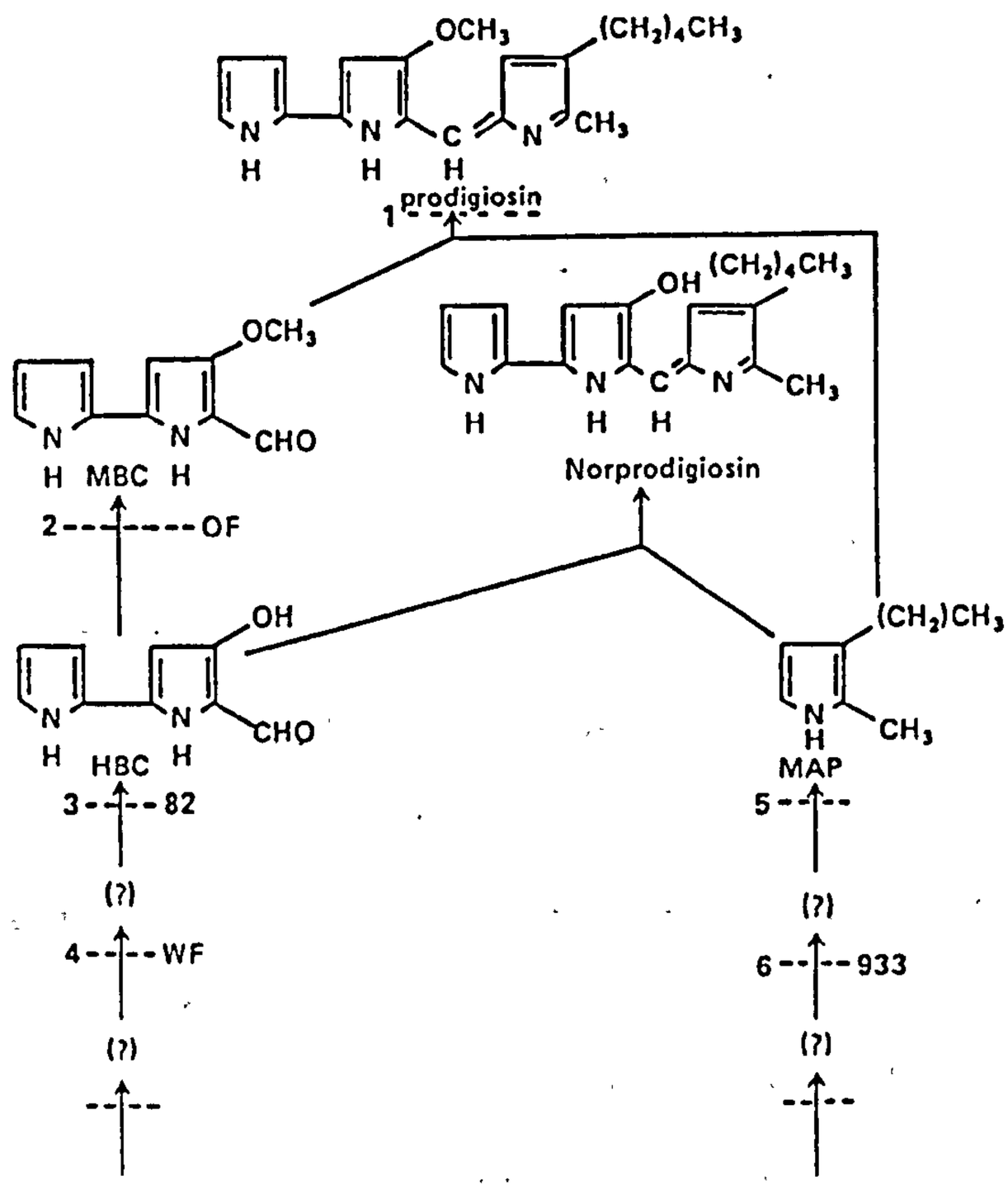


Figure 1.3

Pathway of Prodigiosin Biosynthesis



Reproduced from Williams and Quadri, (1980)

co-synthesis experiments into five classes *redA-E*. Cross feeding experiments between *Ser. marcescens* prodigiosin non-producing mutants and *S. coelicolor* Red non-producing mutants were performed. This identified similarity between WT mutants of *Ser. marcescens* and *redA* mutants of *S. coelicolor* and also between OF mutants of *Ser. marcescens* and *redE* mutants of *S.coelicolor*. This provides further similarity of the biosynthetic pathways of prodigiosin in *Ser. marcescens* and *S. coelicolor*.

The *S. coelicolor redA-E* mutants were shown to be defective in the biosynthesis of the bipyrrole molecule. Coco *et al.* (1991) set out to isolate mutants defective in the production of the monopyrrole. The strategy used was to supplement growth of an existing *redA* mutant with bipyrrole and to select for Red non-producing mutants. This work resulted in the isolation of Red mutants *redG* through to *redR* (Coco *et al.*, 1991). A further mutant *redX* was identified by insertional inactivation of Red production by a DNA fragment to the right hand end of the *red* cluster.

The Red gene cluster was cloned by Malpartida *et al.*, (1990) as shown in (Figure 1.4). A 21kb clone pIJ759, that complemented Red mutants in classes A, B, E and F was isolated. A subsequent clone pIJ2340, with a larger DNA insert also failed to complement mutants *redC* and *redD*. In a random cloning experiment a clone (pIJ2356) was isolated that caused overproduction of Red in *S.lividans* 66. By Southern hybridisation it was shown that pIJ2356 and pIJ2340 contained overlapping DNA fragments. The entire Red cluster was obtained on a clone by cloning fragments from opposite ends of the cluster from pIJ2356 and pIJ2340 adjacent to each other in a low-copy number plasmid. This construct was transformed into a Red^+ *S. coelicolor* A3(2) strain. Clones that had undergone recombination with the

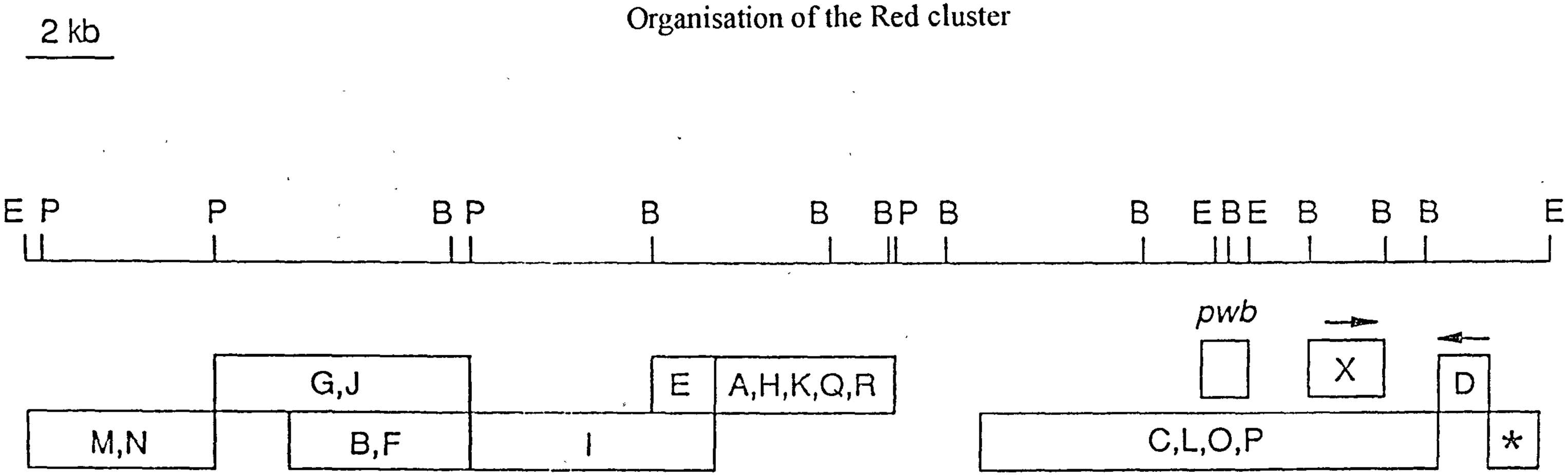
Figure 1.4

Organisation of the Red cluster

The approximate location of mutations representing the Red non-producing mutant classes are shown in boxes below the restriction map. The *redL*, *redO* and *redP* mutants were complemented by a fragment that also contained the region marked *.

Figure 1.4

Organisation of the Red cluster



Reproduced from Hopwood *et al.*, 1994

host chromosome to transfer the entire *red* cluster to the vector were selected as those that could restore Red production to a *redE* mutant (Malpartida *et al.*, 1990).

One class of mutants (*redD*) did not co-synthesise with any other class, this suggested it has a regulatory role. When cloned in multicopy *redD* elicits Red over-production. In *redD* mutants Red production is abolished and there is a lack of expression of *redE* and *redF* genes. This evidence implies that *redD* is a positive activator (Narva and Feitelson, 1990). The *redD* gene has been sequenced, a DNA binding domain was not identified. However, *redD* does show end to end homology to *act*IIORF4 (Narva and Feitelson, 1990).

S. coelicolor A3(2) *redE* mutants showed a similar pattern of co-synthesis to the *Serratia marcescens* OF mutants of Williams and Quadri (1980). This led to the suggestion that both *redE* mutants and OF mutants are deficient in an O-methyltransferase (Feitelson and Hopwood, 1983). Later it was determined that both *redE* and *redF* are required for a functional O-methyltransferase enzyme (Feitelson *et al.*, 1985). The O-methyltransferase catalyses the penultimate step in Red biosynthesis which is the methylation of the yellow intermediate, undecylnorprodigiosin at ring B (Figure 1.2).

The right hand end of the red cluster has been sequenced and five open reading frames identified as shown in Figure 1.5 (M.Bibb personal communication). The predicted amino acid sequence of the open reading frame corresponding to *redX* showed similarity to a β ketosynthetase of type I polyketide synthases. The predicted amino acid sequence of the open reading frame designated *redW* is predicted to encode an acyl CoA dehydrogenase.

Figure 1.5

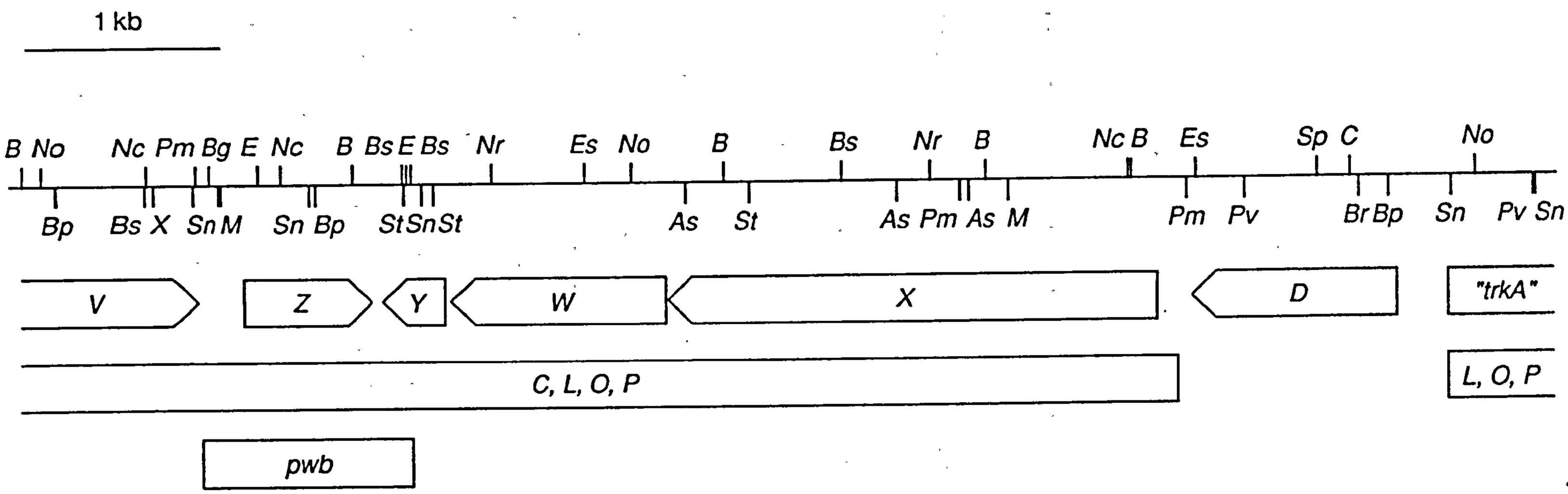
Map of the right hand end of the Red cluster

Sequence data has been obtained of the right hand end of the Red cluster. The open reading frames identified *redV*, *redZ*, *redY*, *redW*, *redX* and *redD* are shown in boxed arrows below the restriction map. Previous studies of Rudd and Hopwood (1980) and Coco *et al.*, (1991) classed Red non-producing mutants into classes *redA* through to *redR*. During the cloning of the entire Red cluster regions of DNA were identified that complemented different mutants. Mutants C, L, O and P were complemented by the region marked in a box overleaf.

The plasmids pIJ2530, pIJ2520 and pIJ2540 contain DNA of the Red cluster from differing Pwb mutants (E. Guthrie unpublished results). The regions D, C, E and F marked above the restriction map indicate the DNA fragments which the plasmids contained in common.

Figure 1.5

Map of the right hand end of the Red cluster



As, AscI; B, BamHI; Bg, BglII; Bp, BspEI; Br, BsrGI; Bs, BstEII; C, ClaI; E, EcoRI; Es, EspI; M, MluI; Nc, NcoI; No, NotI; Nr, NruI; Pm, PmlI; Pv, PvuII; Sn, SnoI; Sp, SphI; St, StuI; X, XhoI

Reproduced with permission of M.Bibb

The predicted protein sequences of *redY* and *redV* show no similarity to anything in the database.

1.5.4 Calcium-dependent-antibiotic

The map position of gene cluster involved in the biosynthesis of the calcium dependent antibiotic has been determined (Figure 1.1). It has been shown to be a lipopeptide that forms a membrane channel. Producing and non-producing strains have been identified. Cda is effective against a wide range of Gram positive organisms in the presence of calcium (Lakey *et al.*, 1983)

1.5.5 The *whiE* cluster encoded spore-associated pigment

The *whi* class of developmental mutants were identified as those that could form aerial mycelium but failed to produce the brown pigment associated with mature spores. The aerial mycelium of *whi* mutants remains white even after prolonged incubation. The brown pigment is thought to be a polyketide, although it is not known to have antibacterial activity. Nine *whi* loci have been identified *whiA* through to *whiI* by the analysis of 52 *whi* mutants (Chater, 1972). Some of these loci have been studied in greater detail. Some of the *whi* mutants, notably those effecting *whiA*, were severely defective in sporulation. The *whi* mutants *whiC193*, *whi-53* and *whi-77* produced reduced amounts of spores. The predicted gene product of *whiG* shows almost 40% identity with the sigma factors σ^D of *Bacillus subtilis* σ^F of *Salmonella typhimurium* and *Pseudomonas aeruginosa* (Chater *et al.*, 1989). The *whiB* gene has also been sequenced (Davies and Chater, 1992). The predicted gene product is a small protein containing a high proportion of charged amino acids. The *whiB70* mutant has been shown to lack septa involved in

sporulation (Davies and Chater, 1992). Potentially WhiB may represent a transcriptional activator as similarity of WhiB to eucaryotic transcription factors has been shown.

The *whiE* gene cluster has been sequenced and seven open reading frames identified (Davies and Chater, 1990). The open reading frames of the *whiE* gene cluster show identity to genes involved in polyketide biosynthesis; notably *whiE*ORFIII to the condensing enzymes of polyketide synthases (Davies and Chater, 1990).

1.6 A-Factor

A-factor or 2-isocapryloyl-3-*R*-hydroxymethyl γ butyrolactone is produced by *S. griseus* and other streptomycetes (Hara *et al.*, 1983). In *S. griseus* A-factor is an autoregulatory factor involved in sporulation and streptomycin biosynthesis. A-factor is effective at concentrations of 10^{-9}M. In *S. coelicolor* A-factor mutants (*afsA*) lack A-factor but produce antibiotics and sporulate normally. It appears that A-factor does not play the same role in activating sporulation and antibiotic biosynthesis in *S. coelicolor* as it does in *S. griseus*.

1.6.1 Regulation of Streptomycin biosynthesis in *S. griseus* by A-Factor

In *S. griseus*, A-factor is required for the transcription of the positive streptomycin activator *strR* (Vujaklija *et al.*, 1993). In the course of their studies three proteins that bound to the region of DNA upstream of the *strR* promoter region were detected. One of these proteins, named protein X in the review of Horinouchi and Beppu (1992), is present in wild-type strains but

not in A-factor-deficient strains. The other two proteins are present in wild-type and A-factor-deficient strains. There is evidence to suggest that protein X is involved in activation of *strR* (Horinouchi and Beppu, 1994). The A-factor signal is believed to be transmitted via a cytoplasmic protein that binds A-factor and was identified in *S. griseus* by Miyake *et al.* (1990). No evidence of an A-factor binding protein was detected in *S. coelicolor* (Miyake *et al.*, 1989). The A-factor binding protein represses sporulation and streptomycin biosynthesis in the absence of A-factor. When A-factor is bound to the protein the repression is alleviated. The A-factor binding protein is a repressor, which is distinct from protein X that acts as an activator of *strR*.

1.6.2 Other autoregulator molecules in *Streptomyces*

Autoregulators such as A-factor, are low molecular weight molecules that are effective at low concentrations. A number of butryolactones similar to A-factor have been reported; such as the *S. virginae* VB compounds which are inducers of virginiamycin production (reviewed by Horinouchi and Beppu, 1992). The production of a low molecular weight molecule may permit signalling to other cells within a colony by diffusion of this chemical signal. It has been shown that the A-factor binding protein does not bind the *S. virginiae* VB-C butyrolactone. Although the *S. virginiae* VB binding protein was less specific with respect to differences in the acyl chain of the VB family of compounds. It appears that the chemical signals are specifically recognised by the producing species and not by other related streptomycetes.

1.6.3 Autoregulator molecules in other organisms

A-factor is structurally similar to the autoinducer of bioluminescence in *Vibrio fischeri*, *N*-(3-oxohexanoyl)-L-homoserine lactone (OHHL). The concentration of autoinducer is low at low cell densities, at high cell densities the autoinducer accumulates until a critical concentration required for induction of luminescence is reached. The autoregulator (OHHL) is believed to bind to the LuxR regulator, which induces the LuxR protein to activate transcription of the *luxICDABE* operon (reviewed by Choi and Greenberg, 1992). The *luxA* and *luxB* genes comprise sub-units of the luciferase enzyme. The *luxC*, *luxD* and *luxE* gene products are involved in the synthesis of the aldehyde substrate of luciferase. The *luxI* gene is as yet the only gene identified required for the synthesis of OHHL. The LuxR protein also represses the transcription of the *luxR* gene.

The accumulation of autoinducer in a cell-density dependent manner allows the bacteria to sense their own population density. A number of cell-density dependent physiological processes that use these diffusible chemical signals to sense population density have been identified (reviewed by Salmond *et al.*, 1995). One system is involved in the regulation of the production of the antibiotic, carbepenem by *Erwinia carotovora.* The production of carbepenem is induced by OHHL in the presence of an active LuxR homologue, CarR (McGowan *et al.*, 1995). Other processes regulated by a similar system include; elastase production in *Pseudomonas aeruginosa*, Ti plasmid conjugal transfer in *Agrobacterium* spp. and exoenzyme virulence factor synthesis by *Erwinia carotovora* (Passador *et al.*, 1993; Zhang *et al.*, 1993; Piper *et al.*, 1993; Jones *et al.*, 1993).

LuxR shows similarity at the C-terminal DNA binding domain to a family of regulators that, in general, form one component of a two component regulatory system. There is usually a sensor kinase that is often membrane bound and a response regulator that is cytoplasmic. The sensor kinase phosphorylates the regulator in response to an environmental signal. The regulator component is activated by phosphorylation in the N-terminal region (Stock *et al.*, 1989). The family of two component regulatory proteins is defined by similarities of a helix-turn-helix DNA binding domain at the C-terminus. This family has been divided into sub-families on the basis of N-terminal region similarities, one of which is the UhpA sub-family. This thesis reports the identification of a *S. coelicolor* gene with homology to the UhpA sub-family.

A-factor does not appear to regulate antibiotic biosynthesis in *S. coelicolor* A3(2). As previously mentioned *afsA* mutants are deficient in A-factor production but sporulate and produce antibiotics as normal. As also mentioned *S. coelicolor* does not have an A-factor binding protein. Another class of *afs* mutants, *afsB* mutants are deficient in A-factor production and are also unable to produce Act and Red and produce reduced amounts of Mmy and Cda. The *afsB* gene represents a growing number of genes that have a pleiotropic effect on antibiotic production.

1.7 Genes with pleiotropic effects on the production of antibiotics

In order to further our understanding of the regulation of secondary metabolism, mutants that were altered in the production of antibiotics were sought. It is worth noting that not all these mutants represent genes involved directly in the regulation of antibiotic gene clusters, but by virtue of a lesion

in a metabolic or cellular function they influence secondary metabolite production.

1.7.1 *afsB, afsR and afsK*

AfsB mutants have a pleiotropic phenotype, as mentioned above (Horinouchi *et al.*, 1983). Supplementing *afsB* mutants with A-factor does not restore antibiotic production. Attempts to clone the *afsB* gene resulted in the cloning of a DNA fragment that restored the pigment deficiency of *afsB* mutants but did not map to the *afsB* chromosomal locus (Horinouchi *et al.*, 1986; Stein and Cohen 1989). The open reading frame identified from sequence analysis was named *afsR*. It is interesting that *afsR* cloned in multicopy restores pigment production in *S. coelicolor* when a chromosomal copy of *afsR* is presumably present in *afsB* mutants. The *afsR* gene encodes a protein of 993 amino acids (Horinouchi *et al.*, 1990). Both the C- or N-terminal halves of the protein were able to partially confer pigment production on *S. lividans* (Hong *et al.*, 1991). The AfsR protein has two domains a C-terminal DNA binding region and an N-terminal ATP binding region. The N-terminal portion shows similarity to *redD* and *actII*ORF4 (Fernandez-Moreno *et al.*, 1991). Disruption of the *afsR* gene resulted in reduction but not loss of Act production in *S. coelicolor* and *S. lividans* (Horinouchi *et al.*, 1990). This demonstrates that *afsR* is not essential for Act biosynthesis. Although some influence on the activation of biosynthesis of Act and Red is exerted by *afsR*. The presence of the ATP binding region led to investigation of the interaction of AfsR and ATP. On purification of the AfsR protein, phosphorylation in the presence of a *S. coelicolor* cell extract and ATP was reported (Hong *et al.*, 1991). A protein kinase, encoded by *afsK* located downstream of *afsR*, was shown to be responsible for the phosphorylation of

afsR (Hopwood *et al.*, 1994). AfsK is a membrane bound serine/threonine kinase.

1.7.2 *afsQ1* and *afsQ2*

The same cloning strategy used to isolate *afsR* was repeated and led to the cloning of the *afsQ1* and *afsQ2* genes (Ishizuka *et al.*, 1992). When present in multicopy, *afsQ1* but not *afsQ2*, stimulated Act production. Similarity of *afsQ1* and *afsQ2* to the response regulator and sensor kinase components, respectively, of two component regulatory systems has been reported (Ishizuka *et al.*, 1992). The two components, the response regulator and sensor kinase, of signal transduction systems each contain conserved characteristic domains. The response regulator contains a C-terminal helix-turn-helix DNA binding domain. DNA binding activity of the C-terminal H-T-H domain is often activated by phosphorylation of the N-terminal 'receiver' region (Stock *et al.*, 1989). The sensor kinase responsible for the phosphorylation of the response regulator is often membrane bound and the gene encoding it is often located proximally on the chromosome to the gene encoding the response regulator. AfsQ2 represents a histidine protein kinase which contains N-terminal hydrophobic regions that potentially span the cytoplasmic membrane. AfsQ1 contains the conserved residues involved in phosphorylation; in particular Asp52, which is the predicted site of phosphorylation. Replacement of Asp52 with a glutamate residue resulted in the inability of the altered *afsQ1* to stimulate pigmented antibiotic production in *S. lividans* (Ishizuka *et al.*, 1992). Disruption of *afsQ1* and *afsQ2* resulted in no detectable change in antibiotic production or morphological differentiation. It appears that *afsQ1* or *afsQ2* are not required for antibiotic production in *S. coelicolor* A3(2). Considering that both the *afsR-afsK* and

afsQ1-afsQ2 systems bear some similarity to two-component sensor-activator systems, it seems likely that the pleiotropic effects on antibiotic biosynthesis may be due to cross-talk between signal transduction components.

1.7.3 *afsR2*

Another small open-reading frame *afsR2* is located adjacent to *afsR* and stimulates Act and Red biosynthesis in multicopy (Vögtli *et al.*, 1994). The stimulatory effect of *afsR2* on Act production, is mediated by activation of transcription of *actII*ORF4. Sequence similarity of *afsR2* was found to a domain found in prokaryotic sigma factors (Vögtli *et al.*, 1994). However, the AfsR2 protein is only 63 amino acids long which is much smaller than any known sigma factor. AfsR2 also lacks the most highly conserved domain within sigma factors (protein region 2), involved in promoter recognition and core binding (Helmann and Chamberlin, 1988). It is therefore unlikely that AfsR2 functions as a sigma factor.

1.7.4 *bld* mutants

S. coelicolor mutants were isolated that were defective in the formation of normal aerial hyphae and spores; these were called bald (*bld*) mutants (Merrick, 1976). The *bld* mutants in general are deficient in antibiotic production, in addition to being unable to form aerial mycelium and sporulate normally. The production of secondary metabolites correlates with the onset of aerial mycelium formation so it was predicted that the *bld* mutants would represent mutations in elements common to both processes. The *bld* mutants *bldA*, *bldD* and *bldG* all produce aerial mycelium, but no antibiotics, when grown on minimal medium with mannitol replacing glucose as sole carbon

source (Champness, 1988). The *bldB* mutants remained unable to produce aerial mycelium irrespective of the carbon source used for growth. Champness (1988) proposed the hypothesis that there are two mechanisms of control of morphological differentiation, one of which is subject to glucose repression.

Some of the *bld* mutants, such as those affecting the *bldH* gene, show a conditional phenotype with respect to antibiotic production and differentiation; when grown on mannitol as sole carbon source in place of glucose *bldH* mutants produce Red and aerial mycelium (Champness, 1988). According to the hypothesis of Champness *bldA*, *bldD*, *bldG* and *bldH* are mutations which prevent the use of the glucose independent pathway of cellular differentiation. Champness (1988) predicts the *bldH* gene is not required for antibiotic biosynthesis whereas the *bldA*, *bldD* and *bldG* mutations effect antibiotic biosynthesis. It is clear that the production of antibiotics and morphological differentiation share common elements.

1.7.4.1 *bldA*

bldA is the best characterised of the *bld* genes. Sequence analysis has revealed that *bldA* encodes a tRNA that recognises the leucine codon UUA (Lawlor *et al.*, 1987). The leucine codon TTA is rarely used in streptomycete genes due to the high G+C content of the DNA. Most of the *Streptomyces* spp. genes that contain TTA codons are involved in secondary metabolism or differentiation (Leskiw *et al.*, 1991a). Since the study of Leskiw *et al.* (1991a) two new genes involved in aerial mycelium formation containing TTA codons have been identified: namely the *S. lividans ramR* gene (Ma and Kendall, 1994) and the *S. griseus amfR* gene (Ueda *et al.*, 1993). It has been

suggested that the *bldA* encoded tRNA represents a mechanism of temporal control of the translation of genes that contain TTA codons (Leskiw *et al.*, 1991a). The study of the abundance of *bldA* encoded tRNA throughout the growth cycle has resulted in two different observations. Gramajo *et al.* (1993) reported that the translation of a TTA containing fusion protein was equally efficient at all stages in growth in liquid culture. Leskiw *et al.* (1991b) reported that mature processed *bldA* tRNA was accumulated in older mycelium of colonies growing on solid media. The difference in growth conditions might be responsible for the apparently conflicting results.

Work of Leskiw *et al.* (1991b) showed that TTA containing genes were not translated in *bldA* mutants. By converting the TTA codons to CTC codons in the *carB* gene of *S.thermotolerans*, *bldA* independent expression was observed (Leskiw et al., 1991b). A similar study of Fernandez-Moreno *et al.* (1991) converted the TTA codon of *actII*-ORF4 to the leucine codon TTG. This resulted in *bldA* independent expression of *actII*-ORF4. The results of Gramajo *et al.* (1993) suggested transcription of *actII*-ORF4 rather than the efficiency of translation of TTA codons was responsible for the temporal regulation of actinorhodin production. Functional *bldA* was shown to be present in exponentially grown cultures of J1501 by the translation of an actIIORF4::*ermE* fusion protein containing a TTA codon (Gramajo *et al.*, 1993). The fusion proteins that contained TTA codons were not translated in *bldA* mutants.

Red biosynthesis is observed in *bldA* mutants grown on media with a concentration of less than 0.5mM phosphate (Guthrie and Chater, 1990). Restoration of production of the other antibiotics; methlylenomycin, actinorhodin and the calcium dependent antibiotic, is not seen under these

conditions. This implies that there is a *bldA* independent mechanism of activating Red biosynthesis that is used in low phosphate conditions. It can also be concluded from this result that there are no TTA codons present in the Red structural genes. The Red pathway-specific activator *redD* has been sequenced and does not contain a TTA codon. This suggests that there is a regulatory gene involved in the induction of Red biosynthesis that contains a TTA codon, in order for Red biosynthesis to be *bldA* dependent.

Guthrie *et al.* (1990) screened mutants of *bldA39* for those that had acquired the ability to produce pigmented antibiotics. This phenotype was designated Pwb, **p**igmented **w**hilst **b**ald. There were no mutants isolated that had restored ability to produce Act. Mapping of the Pwb mutations was inconclusive. Three mutants were studied further these were named Pwb-6, Pwb-9 and Pwb-16. The Pwb mutants showed differing amounts of pigment production on different media (E. Guthrie unpublished results). There were also differences in the amount of pigment produced by the different mutants relative to each other. Libraries of chromosomal DNA from each of the mutants were prepared using the streptomycete vector pIJ698. The DNA libraries were conjugated into a *bldA39* host (J1700) and a number of Red transconjugates isolated (E. Guthrie unpublished results). The production of Red by the transconjugants indicates that the Pwb mutation is dominant. A plasmid clone from each Pwb mutant that caused Red production in J1700 was studied further. The three plasmids pIJ2530 (Pwb-6), pIJ2520 (Pwb-9) and pIJ2540 (Pwb-16) were shown to have several restriction fragments in common (Figure 1.5). Southern hybridisation analysis proved that the region of DNA present in pIJ2530, pIJ2520 and pIJ2540 was part of the Red cluster. The region of DNA from the three plasmids containing the Pwb mutations was partially restricted with *Sau3A*I and randomly cloned into ØKC861. The Pwb

containing ØKC861 clones were used to lysogenise a *bldA* strain, the subsequent lysogens were analysed for Red production. An 800bp *EcoR*I restriction fragment was common to all Red producing clones. This suggested that the mutations causing the Pwb phenotype lay within or close to this DNA fragment. Further characterisation of the *EcoR*I fragment from pIJ2530, pIJ2520, pIJ2540 and from Pwb^+ DNA is one of the major aims of this work.

1.7.5 Antibiotic activator (*aba*) mutants

The *abaA* locus was isolated by its ability to stimulate Act production when present in multicopy (Fernández-Moreno *et al.*, 1992b). This cloned locus is distinct from previously characterised genes involved in antibiotic biosynthesis. The *abaA* locus effects antibiotic production but sporulates normally. Five open reading frames were identified by sequence analysis. The ORF responsible for the Act activating activity is ORFB. Although disruption of a second open reading frame, ORFA, was unsuccessful so the possibility remains that this ORF may also be involved in the activation of Act. Inactivation of *abaA* results in the abolition of Act production, reduction in Red and Cda but Mmy remains unaffected. No significant homology to ORFB to anything in the databases was identified (Fernández-Moreno *et al.*, 1992b).

1.7.6 Antibiotic synthesis (*abs*) mutants

The *bld* mutants demonstrate that antibiotic production and sporulation share common elements. A new class of developmental mutants were isolated, which were defective for production of all four antibiotics, but were unaffected in morphological differentiation. The mutants were mapped to two

specific loci *absA* and *absB* (Adamidis *et al.*, 1990; Adamidis and Champness, 1992). The isolation of the *abs* mutants suggests that antibiotic biosynthesis of all four antibiotics are subject to a common control that is separate from the control of morphological differentiation.

Sequence analysis of *absA* revealed similarity of the predicted AbsA product to the sensor kinase element of two-component regulatory systems (W.Champness, personal communication). In close proximity to *absA* an open reading frame was identified whose product showed similarity to the response regulator partner of signal transduction systems. This must be a strong candidate for the response regulator partner of *absA*.

Act production is restored to the *absB* mutants by introducing multiple copies of *actII*ORF4 (Adamidis and Champness, 1992). Both *absA* and *absB* mutations were bypassed when the C-terminal domain of *afsR* was present in multicopy (Horinouchi *et al.*, 1990). The *afsQ1* gene present in multicopy causes activation of Act in *absA* mutants (Ishizuka *et al.*, 1992).

The interaction between these genes suggests that a complex network of signals that activate antibiotic biosynthesis. Considering the similarity of these genes to elements of the two-component regulatory systems it raises the possibility that cross-talk is occurring between genes not involved in secondary metabolism and regulators involved in antibiotic biosynthesis.

Cross-talk between systems occurs when a sensor kinase component activates the response regulator of a separate system. This has been demonstrated in *E. coli* where the sensor kinase module involved in regulation of chemotaxis (CheA) is able to phosphorylate the response regulator involved

in nitrogen regulation (NtrC) (Ninfa *et al.*, 1988). The converse interaction of activation of the chemotaxis response regulator (CheY) by the Ntr system kinase (NtrB) was also shown. In these studies more of the non-partner response regulator was needed and phosphotransfer was slower between non-partner response regulators and the sensor kinase. Wanner (1992) raises the issue of whether cross-talk is a mechanism of co-ordinating cellular functions or whether it is the result of chance interaction between regulators that have a common mechanism of function. The interaction of two systems to co-ordinate regulation is defined as cross-regulation, in order to make a distinction to that of the chance interaction of crosstalk. It is suggested that, because excess of the non-partner activator molecule is required in order for phosphotransfer to occur between the Ntr and chemotaxis genes, then cross talk and not cross regulation occurred (Wanner, 1992).

Cross regulation may be an important process to link regulation of cellular functions. Is cross talk or cross regulation occurring between the genes that exert an effect on the regulation of antibiotic biosynthesis? Disruption of *afsQ1* and *afsQ2* does not affect antibiotic biosynthesis which could imply cross talk occurred to activate antibiotic biosynthesis i.e. that *afsQ1* and *afsQ2* are not normally a part of the regulatory cascade to activate antibiotic biosynthesis.

It seems clear that in general the activation of Act is via the pathway-specific gene actIIORF4, this has been shown to be the case for *afsR*. The *act*IIORF4 gene does not show significant homology to response regulator transcriptional activators and as such the activation of this gene by cross talk of sensor kinases of unrelated systems is unfeasible. This may suggest that

there is a response regulator that activates *act*IIORF4 that is as yet unidentified.

1.8 Physiological factors that effect antibiotic production

1.8.1 Stringent response

In liquid medium antibiotic production occurs in stationary phase or in cultures with a low growth rate (Takano *et al.*, 1992). The onset of antibiotic production has been correlated with an increase in the intracellular levels of a phosphorylated guanosine residue ppGpp (Ochi, 1986). In *E. coli* ppGpp plays a role in the stringent response, which occurs following amino acid limitation of cells (Cashell and Rudd, 1987). During the stringent response there are changes in gene expression. Transcription of genes encoding stable RNA i.e. tRNA and rRNA, is reduced. There is also evidence to relate the presence of ppGpp with the control of growth rate in *E. coli*. Bascarán et al. (1991) did not find a relationship between the levels of ppGpp present in the cells and the amount of cephalosporins produced in *S. clavuligerus*. In *S. coelicolor* the induction of a peak in ppGpp synthesis was induced by nutritional shiftdown or by the addition of serine hydroxamate. However, this did not have a stimulatory effect on the transcription of the antibiotic regulatory gene *redD* (Takano *et al.*, 1992). Takano *et al.* (1992) also noted that the production of Red correlated more closely with the slowing of the growth rate of cultures than with the appearance of ppGpp. A peak of ppGpp synthesis and the transcription of *redD* were observed at the transition into stationary phase of an unmanipulated culture. It was suggested that an increase in ppGpp concentrations was not sufficient to activate transcription of antibiotic biosynthesis genes.

1.8.2 Catabolite repression

Antibiotic production can be repressed or stimulated depending on elements present in the growth media. Repression of Act production is seen in *S. coelicolor* A3(2) in conditions of excess phosphate or where there are easily assimilated sources of nitrogen (Hobbs *et al.*, 1990). Secondary metabolites show different sensitivities to the same catabolite. For instance the exclusive production Mmy is observed in medium of high phosphate defined by Hobbs *et al.* (1992). Act synthesis is more sensitive to elevated levels of ammonium ions than is Red synthesis but if phosphate concentration is reduced this relieves the effect of ammonium inhibition. The effect of phosphate inhibition is not reversed by ammonium ions.

Ammonium repression of secondary metabolite synthesis is also seen in other *Streptomyces* spp. and closely related organisms. Flores and Sanchez (1985) report ammonium repression of erythromycin biosynthesis in *Saccharopolyspora erythrae*. High ammonium concentrations also suppress the synthesis of avermectins in *Streptomyces avermitilis* (Novák *et al.*, 1992). The basis of ammonium repression of tylosin in *Streptomyces fradiae* has been characterised. Tylosin is a macrolide antibiotic synthesised from a protylonolide precursor to which three sugars (mycaminose, mycinose and mycarose) are attached (Figure 1.6). Omura *et al.* (1984) reported that the protylonolide precursor biosynthesis is the process that is repressed by high ammonium concentrations. The protylonolide is synthesised from acetate, proprionate and n-butyrate molecules, that are supplied by amino acid degradation. ^{14}C labelling studies demonstrated that valine, threonine and succinate (or more correctly their degradation products) were the most efficiently incorporated into protylonolide (Omura and Tanaka, 1986).

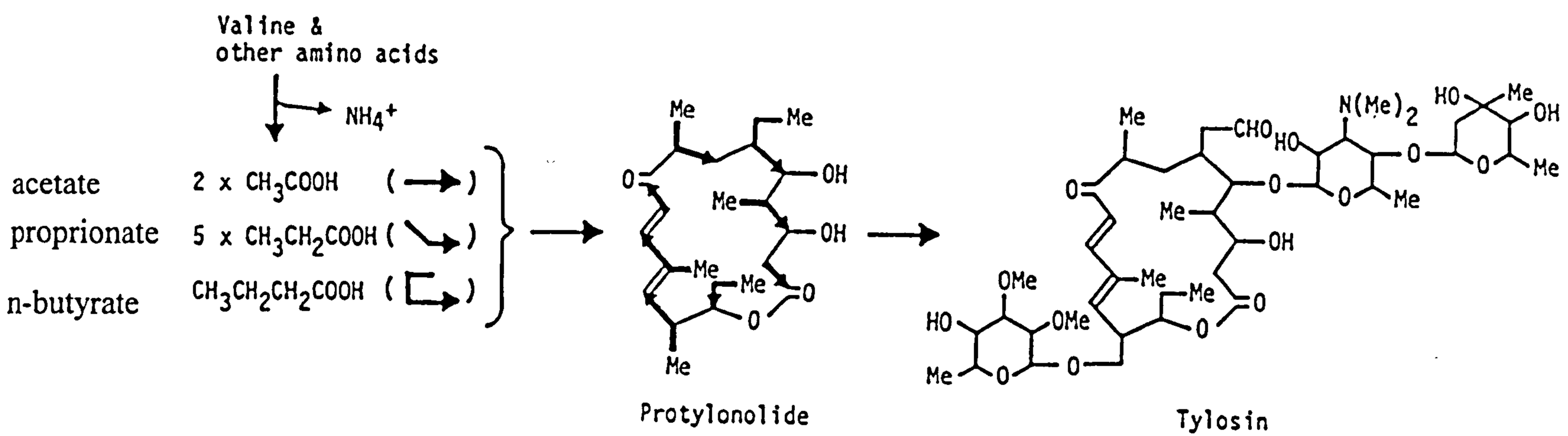
Figure 1.6

Biosynthesis of tylosin in *S. fradiae*

Acetate, proprionate and n-butyrate derived from amino acid catabolism are precursor molecules for the synthesis of protylonolide. Tylosin is synthesised from the precursor protylonolide to which three sugar molecules are attached

Figure 1.6

Biosynthesis of tylosin in *S. fradiae*



Reproduced from Omura and Tanaka, (1986)

The addition of acetate, proprionate and n-butyrate reversed the effect of ammonium repression. As did isobutyrate and/or α-keto-isovalerate, both degradation products of valine. This suggested that valine catabolism is the source of the precursor molecules; acetate, proprionate and n-butyrate for the formation of protylonolide. This led to investigation of valine catabolic enzymes, in particular valine dehydrogenase, the first enzyme of branched chain amino acid catabolism. Omura and Tanaka (1986) report that the effect of phosphate and ammonium repression of tylosin biosynthesis is mediated by the repression of valine dehydrogenase.

1.9 Interaction between primary and secondary metabolism

The correlation between the level of valine dehydrogenase activity and the amount of production the macrolide antibiotic tylosin, also demonstrates that the supply of precursor molecules is also a source of control of antibiotic synthesis. The catabolism of branched chain amino acids (in particular valine and threonine) as a source of proprionyl CoA and methylmalonyl CoA for macrolide biosynthesis in *S. fradiae* and *S. ambofaciens* has been suggested (Hutchinson *et al.*, 1993). Hutchinson and co-workers (1993) cloned the valine dehydrogenase (*vdh*) gene from *S. coelicolor* and were able to determine the effect of disrupting this gene on antibiotic production. Disruption of *vdh* in *S. fradiae* and *S. ambofaciens* resulted in a two to ten fold reduction in antibiotic production. Addition of proprionate to the culture or, addition of the *vdh* gene to the *vdh* disrupted mutants, restored antibiotic production to near wild-type levels (Hutchinson *et al.*, 1993).

An *S. lividans* class of arginine mutants (*argC*) are deficient in clavulanic acid biosynthesis (Ludovice *et al.*, 1992). The clavulanic acid deficiency of

these mutants is restored by the introduction of the cloned *S. lividans argC* gene. This highlights the use of primary metabolites such as amino acids as precursors for secondary metabolites.

1.9.1 Involvement of proline in osmoregulation

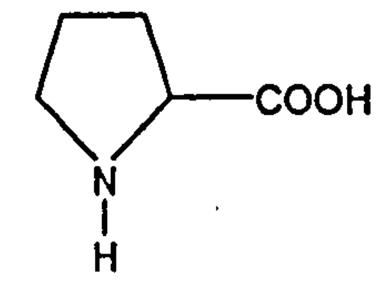
This study focuses on the regulation of Red in *S. coelicolor*. An intact proline molecule is incorporated into Red (Wassermann *et al.*, 1974) as part of ring A (Figure 1.2). An additional cellular function of proline is as an osmoprotectant. Osmoprotectants or compatible solutes are low molecular mass solutes which can be accumulated to relatively high concentrations without becoming toxic to the cell. This prevents dehydration of the cell in conditions of increased osmolarity. Compatible solutes are also believed to be involved in maintaining the conformation of proteins and membranes in conditions of osmotic stress. In addition to proline other compatible solutes used by procaryotes include glycine betaine, glutamate, trehalose and proline betaine (Figure 1.7).

The adaption of *E. coli* to hyperosmotic shock occurs in two phases (reviewed by Csonka and Hanson, 1991). The initial response is rapid uptake of potassium ions and synthesis of glutamate. The potassium and glutamate concentrations fall and are replaced by other compatible solutes. In *E. coli* trehalose is synthesised to functionally replace the potassium ions and glutamate. In streptomycetes proline is a major compatible solute (Killham and Firestone, 1984a).

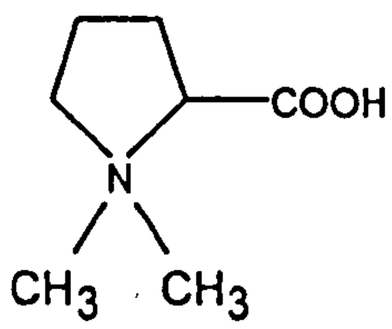
In Gram positive bacteria glutamate is often the major compatible solute. this was not found to be the case in salt-stressed *S. griseus* and *S. californicus*

Figure 1.7

Chemical structure of proline, proline betaine and glycine betaine



Proline



Proline betaine

$(CH_3)_3 N^+CH_2COO^-$ Glycine betaine

(Killham and Firestone, 1984a). Killham and Firestone (1984a) showed that proline concentration increased from less than 6% in *S. griseus* cells grown in basal medium, to 50% of the free amino acid pool in cells grown in medium containing 1M NaCl. The concentration of the intracellular free amino acid pool also increased ten fold in response to 1M salt-stress. Intracellular concentrations of glutamine and alanine also increased in response to osmotic stress, but to a lesser extent than the increase in proline concentration. Addition of proline to the growth medium increased the specific growth yield of *S. griseus* at salt concentrations of up to 1M NaCl (Killham and Firestone 1984b). Proline uptake was shown to increase with increased osmotic stress. Of the proline transported into the cell 94% remained as free proline.

Work has been conducted in our laboratory studying the biosynthesis, catabolism and transport of proline and tryptophan. One of the aims of the study was to compare the metabolism of proline and tryptophan and identify differences between an amino acid that is a precursor for a secondary metabolite and one that is not (to our knowledge). Another purpose for this study was to elucidate the origin of the proline molecule incorporated into Red. Was it from the extracellular environment or from *de novo* synthesis? A review of the findings is presented with comparisons to metabolism of proline in other organisms.

1.9.2 Proline transport

Proline transport mutants (Put^-) were isolated to investigate if the origin of proline used in the biosynthesis of Red from the extracellular environment. The Put^- mutants were isolated by selecting for cells resistant to proline analogues azetidine-2-carboxylate (AZC) and 3,4-dehydroproline (DHP),

their chemical structures are presented in Figure 1.8. This strategy was successfully used by Maloy *et al.* (1987) to isolate mutants of the major proline permease (*putP*) in *E. coli*.

One proline transport mutant PTM44f was isolated directly from the strain J802 on screening for proline analogue resistance. In addition to resistance to proline analogues PTM44f was unable to utilise proline as sole carbon or nitrogen source. This is the phenotype expected of a proline permease mutant. During these experiments it was discovered that J802 (NF *agaA7*) could use proline as a nitrogen source but not as a carbon source. The isogenic Aga^+ strain J801 (NF) could utilise proline as sole nitrogen or carbon source.

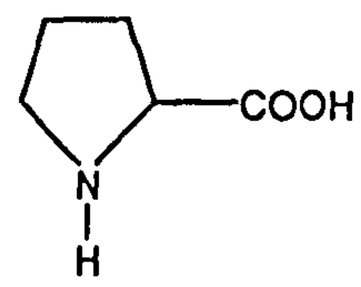
Mutants that were able to utilise proline as a carbon source (PU mutants) were selected. The PU mutants were used to isolate proline transport mutants by selecting for resistance to proline analogues. This second round of mutants were named PUM mutants (U. Swoboda, personal communication). The PUM mutants and PTM44f are collectively called proline transport mutants (Put^-).

The proline transport mutants (Put^-) over-produced Red in addition to the expected phenotype; resistance to the proline analogues, inability to grow on proline as sole nitrogen and carbon source and inability to transport ^{14}C proline. The Put^- mutants were also unable to grow in liquid or on complex medium and did not produce aerial mycelium or spores.

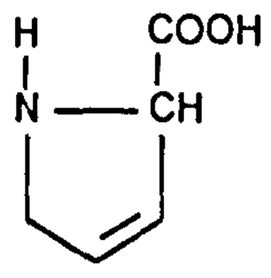
In *S. clavuligerus* proline transport is mediated by two transport systems, neither of which are induced by proline (Bascaran *et al.*, 1992). Two proline transport systems have also been reported for *S. antibioticus* (May and

Figure 1.8

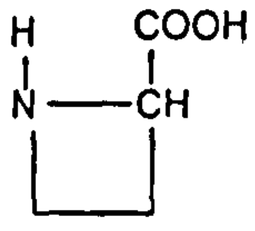
Chemical structure of proline and proline analogues



Proline



DHP



AZC

Formica, 1978). In *S. antibioticus* one system is induced by glutamate and is substrate non-specific. The second system is constitutively expressed and specific for transport of proline. Work of N.Wood has failed to identify a second proline transport system in *S. coelicolor*, even under conditions of osmotic stress (personal communication).

1.9.3 Proline transport in enteric bacteria

Proline transport systems of enteric bacteria have been well documented. Three transport systems have been identified in *E. coli* and *Salmonella typimurium*, proline porters I, II and III (Maloy *et al.* 1987, Cairney *et al.* 1985 a and b). Proline porter I (PPI) is the major proline permease. It is an integral membrane protein encoded by *putP*. Active transport of proline is driven by a sodium-proline symport. This transport system has a high affinity for proline with a Km of 2μM. A second transport system (PPII) is primarily a glycine betaine permease encoded by *proP*, that has a low affinity for proline. The third proline transport system (PPIII) is encoded by *proU* and is induced in conditions of high osmolarity.

1.9.3.1 Proline porter I encoded by *putP*

The *putP* gene product has protein interactions with PutA. A bi-functional enzyme with both proline oxidase and P5C dehydrogenase activities is encoded by *putA*. The *putA* gene product binds to the cytoplasmic face of the inner membrane. Proline oxidase activity requires direct coupling to the electron transport chain. The enzyme is only active when bound to the inner membrane. Both *putA* and *putP* map to the 22min region of the *E. coli* chromosome. Sequence analysis has shown that *putA* and *putP* genes are

adjacent and are divergently transcribed. A regulatory function has been ascribed to the overlapping region of the two genes. In the absence of exogenous proline, PutA binds to the *put* regulatory region repressing transcription of both *putA* and *putP*. In the presence of high proline concentrations PutA binds to the membrane where it becomes enzymatically active (Maloy *et al.*, 1987).

1.9.3.2 Proline porter II encoded by *proP*

Proline porter II is induced in conditions of high osmolarity. Betaine rather than proline is the primary substrate for this transport system (Cairney *et al.*, 1985a). Transport of proline and betaine increased three-fold in cells grown at high osmolarity (defined as 0.3M NaCl). Transcription of *proP* correspondingly increased three-fold. This implies that increased transcription of *proP* and subsequent synthesis of ProP, accounts for the increase of transport of proline and betaine by PPII at high osmolarity. At high salt concentrations the affinity of ProP for proline is reduced and the affinity for betaine is increased. Significant transport of betaine by PPII only occurs at high salt concentrations. Cairney *et al.* (1985a) suggested that a conformational change in the permease protein, in response to osmotic stress would account for the increase in transport of betaine. At very high salt concentrations, in excess of 0.8M NaCl, transport of betaine and proline by PPII is inhibited.

1.9.3.3 Proline porter III encoded by *proU*

Proline uptake by proline porter III is negligible by comparison to uptake of proline by PPI and PPII. The primary substrate for this transport system as for

the *proP* permease is betaine. Proline uptake is stimulated three-fold in growth at high osmolarity. In a *putP*$^-$ *proP*$^-$ strain when grown at 300μM proline concentrations at high osmolarity, significant proline transport was not detected (Cairney *et al.*, 1985b).

Transcription of *proU* is induced over one hundred-fold in response to hyperosmotic stress (Cairney *et al.*, 1985b). The PPIII has two membrane associated components, ProU and Pro W and a periplasmic binding protein ProX (Csonka et al., 1983). The *proU* operon is comprised of these three genes and is transcribed in the order *proU, proX, proW*. Mapping of the *proU* promoter showed that the -10 and -35 regions were separated by longer than usual spacing. From the results it is expected that *proU* is a weak promoter in the absence of an activator or of DNA supercoiling. In *E. coli* a positive transcriptional activator is required for stimulation of *proU*. In *Salmonella typhimurium*, mutants that caused over-expression of *proU* were located to the topoisomerase I gene (*topA*) involved in DNA supercoiling and *hns* a histone-like protein.

1.9.4 Proline transport in Gram-positive organisms

In *Staphlococcus aureus* two transport systems have been documented (Townsend and Wilkinson 1992). One system has a high affinity and is specific for L-proline. The second has a low affinity for proline and is less substrate specific. Proline and glycine betaine are the principle compatible solutes in *Staph. aureus*. Glycine betaine did not compete with proline for transport into the cell in either system unlike the ProU and ProP transporters of *E. coli*. Both systems are dependent on a sodium gradient. Proline betaine (N,N-dimethylproline) transport via the two proline transport systems has

been demonstrated (Utki *et al.* 1995). Proline uptake occurred more rapidly than glycine betaine uptake after hyperosmotic shock. In *Staph. aureus* proline uptake may be important in the initial response to osmotic stress.

1.9.5 Proline utilisation

Proline is catabolised to glutamate in a pathway that involves two enzymatic steps (Figure 1.9). Proline oxidase catalyses the conversion of proline to pyrrolidine-5-carboxylic acid P5C. P5C is spontaneously converted to glutamate semialdehyde. The second enzymatic step is catalysed by P5C dehydrogenase, where glutamate semialdehyde is converted to glutamate. The work of Hood *et al.* (1992) demonstrated that proline oxidase is membrane bound and P5C dehydrogenase resides in the cytoplasm in *S. coelicolor*. As previously mentioned, in *E. coli* both enzyme activities are present on a bi-functional enzyme encoded by *putA (*Maloy *et al.*, 1987). In *S. coelicolor* activities of both enzymes were detected in cells grown on medium lacking proline. Initial enzyme assays showed the specific activity of P5C dehydrogenase to be of the same order in the presence and absence of proline as either sole carbon or nitrogen source. Recent studies of Smith *et al.* (1995) revealed P5C dehydrogenase to be induced one hundred and thirty fold by the addition of proline to the growth medium of early exponential growth cells. This is in agreement with the findings of Bascaran *et al.* (1990) that catabolism of proline in *Streptomyces clavuligerus* is initiated by an inducible proline dehydrogenase. In *E. coli*, *Salmonella typhimurium* and *Klebsiella aerogenes* PutA enzyme activity is induced by proline. The P5C dehydrogenase of *S. coelicolor* has been purified to homogeneity (Smith *et al.*, 1995)

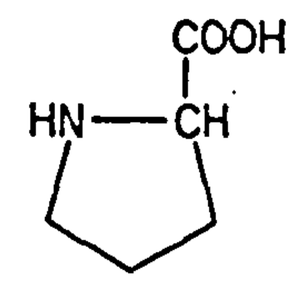
Figure 1.9

The pathway of proline utilisation

Proline is catabolised by a pathway that contains two enzymatic steps as shown on the facing page.

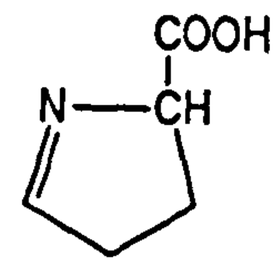
Figure 1.9

Pathway of proline utilisation



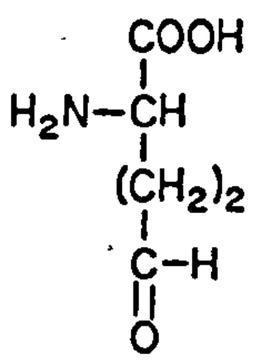
Proline

Proline oxidase



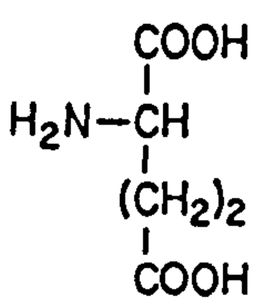
1-Pyrroline 5-carboxylate

spontaneous



Glutamate semialdehyde

P5C dehydrogenase



Glutamate

1.9.6 Proline synthesis

The synthesis of proline occurs by a four step pathway, three of the steps are catalysed enzymatically as shown in Figure 1.10. Glutamic acid is converted to glutamyl phosphate by the enzyme glutamate kinase, the *proB* gene product. Glutamyl phosphate is catalysed by glutamate semialdehyde dehydrogenase, the *proA* gene product, to form glutamate semialdehyde. Pyrroline-5-carboxylic acid (P5C) is formed in a spontaneous cyclisation reaction of glutamate semialdehyde. The *proC* gene product, P5C reductase catalyses the conversion of P5C to proline.

The *proBA* regions of *Salmonella typhimurium* and *E. coli* form an operon which is transcribed in the direction of *proB* to *proA* (Bachmann *et al.*, 1980; Mahan and Csonka, 1983). In *E. coli* the *proBA* region lies at min 6 on the genome, the *proC* gene is present at a separate location at min 9 on the chromosome (Bachmann *et al.*, 1980). In *Pseudomonas aeruginosa* the *proA*, *proB* and *proC* genes are unlinked. (Csonka and Baich, 1983). Hood *et al.* (1992) performed assays of P5C reductase in *S. coelicolor*. P5C reductase activity was detected in all strains tested including wild-type J802 and Put$^-$ mutant. Cells grown on proline and glutamate as sole nitrogen source showed a two fold increase in enzyme activity by comparison with cells grown on ammonium as sole nitrogen source. P5C reductase activity was much reduced in cells grown on glutamine, the mechanism of the repression is not known. P5C reductase has been purified for *S. coelicolor* and amino acid sequence obtained from peptide fragments (D. Smith personal communication).

A 4.5 kb region of DNA covering the *proA-proB* region of the *S. coelicolor* genome was sequenced. Three open reading frames were identified the first,

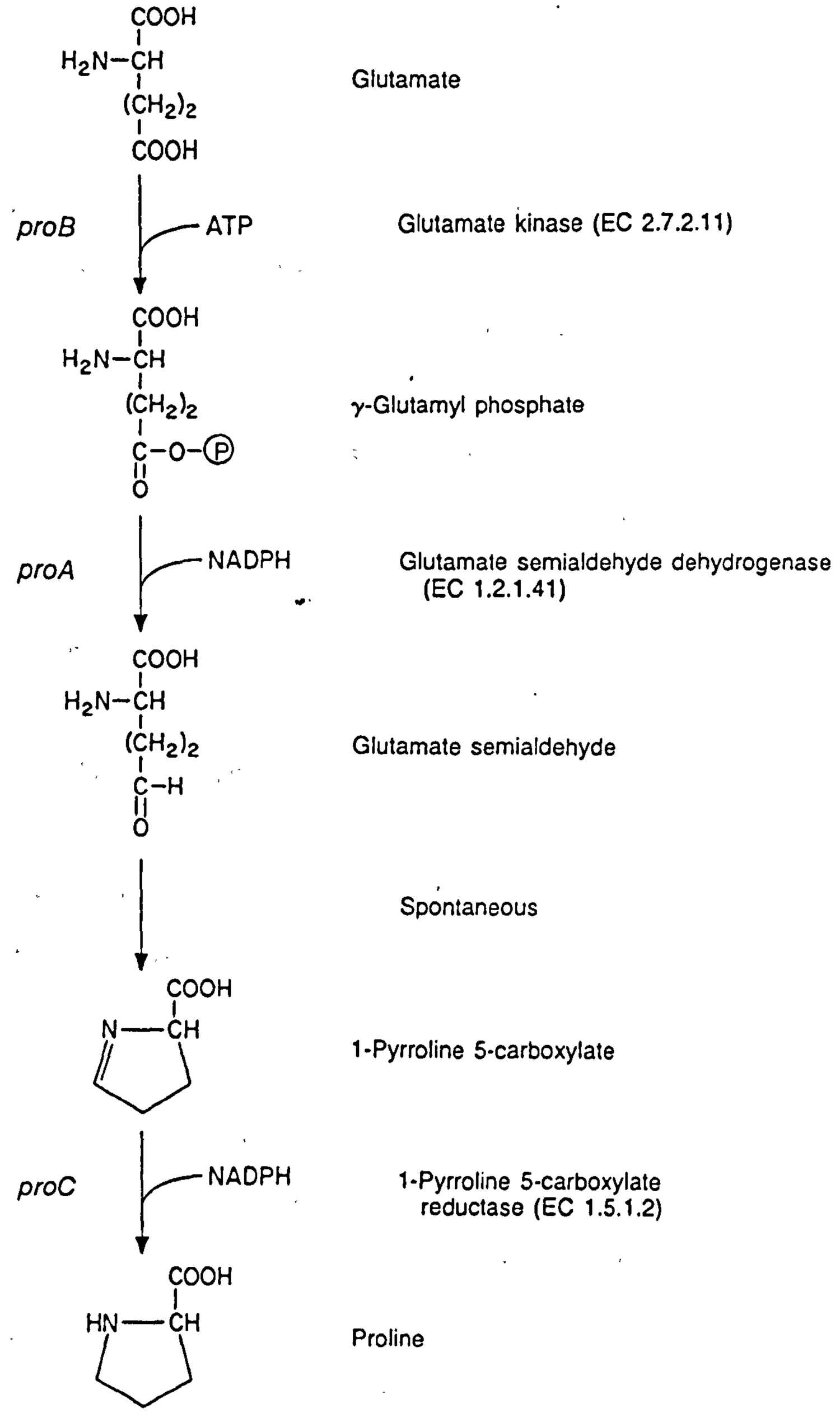
Figure 1.10

The Proline biosynthetic pathway

Proline is synthesised from glutamate in four steps. The first, second and fourth steps are catalysed enzymatically, the third is a spontaneous cyclisation reaction. The genes that encode the proline biosynthetic enzymes *proA*, *proB* and *proC* are shown.

Figure 1.10

The Proline biosynthetic pathway



Reproduced from Csonka and Baich (1983)

had a low but significant homology to the *E. coli proB* gene. The *E. coli proB* gene product is glutamate kinase (Figure 1.10). The second open reading frame designated ORFX has unknown function, it has unusual codon usage and it is speculated that it contains a DNA binding domain. The third open reading frame showed significant homology to the *E. coli proA* gene.

Promoter studies of the *proA* and *proB* genes were performed using the *whiE* based promoter probe vector pCLL38. Proline stimulated the promoter activity of *proA* and *proB* by approximately two-fold. The *proB* promoter gave consistently greater activity than that of the *proA* promoter. When proline was added to glycerol grown cells both promoter activities were enhanced, greater activity was found in *proA* than in *proB*. When glucose was used as a carbon source the activity of the promoters was repressed with the greatest effect being on *proB*.

1.9.7 Proline synthesis is other organisms

Proline over-producing mutants of *Salmonella typhimurium* were isolated. Higher intracellular levels of proline correlated with increased osmotolerance (Csonka *et al.*, 1981). The *pro-74* mutant which had the highest proline concentration of the over-producing mutants, produced thirty to four hundred fold more proline than the parental strain (Csonka *et al.*, 1981). The glutamate kinase (ProB) activity of the *pro-74* mutant is increased by comparison to the parental strain. The mutation has been located to the *proB* promoter-operator or structural gene (Mahan and Csonka, 1983). In *E. coli* glutamate kinase (ProB) is inhibited 80% by 1mM proline (Csonka and Baich, 1983). Fructose-6-phosphate, glucose-6-phosphate and AMP are also

inhibitors of glutamate kinase. The *P. aeruginosa* glutamate kinase is inhibited 50% by 5mM proline.

In *E. coli* P5C reductase (ProC) is inhibited by proline concentrations in excess of 15mM. Brady and Csonka (1983) reported the P5C reductase of *S. typhimurium* was not repressed by the presence of exogenously supplied proline. In fact *proC* was later shown to be constitutively expressed (Brady and Csonka, 1988).

1.9.8 Proline metabolism in *S. coelicolor*

A hypothesis was proposed by Hood *et al.* (1992) that proline synthesis and degradation are in dynamic equilibrium, if this equilibrium is disturbed, as in the Put$^-$ mutants, then Red biosynthesis acts as a sink for excess proline. The hypothesis is the subject of further investigation in chapter 3 of this study. Hood *et al.* (1992) demonstrated that proline synthesis is constitutive in *S.coelicolor* and this has an important implication for this study. If proline synthesis was not constitutive and was instead subject to feedback repression, accumulation of proline would not occur. The hypothesis proposed to explain over-production of Red in the proline transport mutants is dependent on proline accumulation within the cell. The hypothesis is supported by the finding of Hood and co-workers that increased Red production was shown by cells that had been transformed with the *proBXA* region on a multicopy plasmid. It is speculated that the introduction of extra copies of genes of *proA* and *proB* caused over-production of proline and that the excess proline was shunted into Red biosynthesis. If this were the case then it would provide a parallel for the over-production of Red in Put$^-$ mutants.

1.10 Project Aims

The overall aim of this project is to further understanding of the regulation of Red in *S. coelicolor* A3(3). Two main approaches were taken:

i To study the over-production of Red in the Put⁻ mutants by determining whether a Red⁻ Put⁻ lysogen is viable. If a Red⁻ Put⁻ lysogen is not viable this may suggest that Red is an essential shunt for proline in a Put⁻ background. The timing of the production of Red by the Put- mutants is also of interest. In order to determine whether Red is produced earlier in the growth phase in the Put⁻ mutants than in the progenitor strain J802 S1 nuclease mapping was performed.

ii To further characterise the Pwb mutations by cloning and sequencing the 800bp *EcoR*I Pwb containing fragment of pIJ2520 (Pwb-9), pIJ2530 (Pwb-6), pIJ2540 (Pwb-16) and from pCNB1003 (Pwb^+)

CHAPTER 2

MATERIALS AND METHODS

2.1 Growth and maintainance of bacterial strains

All the organisms used in this study are listed in Table 2.1 below. *Streptomyces* strains were maintained on R2YE after growth for 48-72 hours at 30°C. Streptomycete spore stocks were maintained in 20% (v/v) glycerol at -20°C. *Escherichia coli* strain TG1 was maintained on M9 medium supplemented with thiamine (10μg ml^{-1}), after growth for 48 hours in a 37°C incubator. *E.coli* strain GM48 was maintained on LB agar plates after growth for 24 hours incubated at 37°C.

Table 2.1 Bacterial strains

Strain	Phenotype	
Escherichia coli		
TG1	*supE, hsdD5, thi* Δ (*lac-proAB*) F' [*traD36, proAB*$^+$, *lacI*^{q}, *lacZ*(ΔM15)]	
GM48	*thr, leu, thi, lacY, galK, galT, ara, tonA31, tsx, dam, dcm, supE44*	
Streptomyces		Source/Reference
lividans		
TK24	*str-6* SLP1$^-$ SLP2$^-$	Hopwood *et al.*, 1985
Streptomyces		
***coelicolor* A3(2)**		
D132	SCP1$^-$ SCP2$^-$	John Innes Institute (J.I.I)
J1506	*hisA1 uraA1 strA1 pgl* SCP1$^+$	J.I.I

Strain	JF strain No.	Phenotype	Source/Reference
J1700		*bldA39 strA1 mthB2 pgl* NF	Leskiw *et al.*, 1993
J802		*agaA7* SCP2* NF	J.I.I
M124		*argA1 proA1 cysD18*	Hopwood *et al.*, 1985
M145		SCP1⁻ SCP2⁻ prototrophic	Hopwood *et al.*, 1985
Red⁻ strains			
JF1		*argA1 guaA1 act(II)77 redD42*	Feitelson and Hopwood, 1983
TK13		*proA1 argA1 cysD18 strA1 act(I)118 redA7*	K.F.Chater
TK16		*argA1 guaA1 actV117 redA59*	K.F.Chater
JF3		*redE60 act(IV)122 his strA1 pgl*	K.F.Chater
21	457	*red21 argA1 guaA1 actI*V117 *redA59*	J.Feitelson
G15	454	*redG15 argA1 guaA1 actI*V117 *redA59*	Coco *et al.*, 1988
H30	464	*redH30 argA1 guaA1 actI*V117 *redA59*	Coco *et al.*, 1988
I23	458	*redI23 argA1 guaA1 actI*V117 *redA59*	Coco *et al.*, 1988
J25	460	*redJ25 argA1 guaA1 actI*V117 *redA59*	Coco *et al.*, 1988
K2	452	*redK2 argA1 guaA1 actI*V117 *redA59*	Coco *et al.*, 1988
L27	461	*redL27 argA1 guaA1 actI*V117 *redA59*	Coco *et al.*, 1988
M13	453	*redM13 argA1 guaA1 actI*V117 *redA59*	Coco *et al.*, 1988
M31	465	*redM31 argA1 guaA1 actI*V117 *redA59*	Coco *et al.*, 1988
N1	451	*redN1 argA1 guaA1 actI*V117 *redA59*	Coco *et al.*, 1988
N28	462	*redN28 argA1 guaA1 actI*V117 *redA59*	Coco *et al.*, 1988
O17	456	*redO17 argA1 guaA1 actI*V117 *redA59*	Coco *et al.*, 1988
Q16	455	*redQ16 argA1 guaA1 actI*V117 *redA59*	Coco *et al.*, 1988

Strain		Phenotype	Source/Reference
R24	459	*redR24 argA1 guaA1 actI*V117 *redA59*	Coco *et al.*, 1988
C577		*hisA1 uraA1strA1abs-554* SCP1⁻ SCP2⁻ Pgl⁻	Adamidis *et al.*, 1990
C120		*hisA1 uraA1 strA1 abs-120* SCP1⁻ SCP2⁻ Pgl⁻	Adamidis and Champness, 1992
Proline transport mutants			
PTM44f		*put44 agaA7* NF	Hood *et al.*, 1992
PUM18		*put18 agaA7* NF	Hood *et al.*, 1992
PUM7r		*put7 agaA7* NF	Hood *et al.*, 1992
US218		*strA1 pgl put44* SCP2* NF	U.Swoboda
US224		*strA1 pgl put44* SCP2* NF	U.Swoboda
Pwb mutants			
EG111		*pwb9 bldA39 strA1 mthB2 pgl* NF	Guthrie and Chater, 1990
EG311		*pwb6 bldA39 strA1 mthB2 pgl* NF	Guthrie and Chater, 1990
EG411		*pwb16 bldA39 strA1 mthB2 pgl* NF	Guthrie and Chater, 1990

Plasmids and bacteriophages used in this study are listed in Table 2.2.

Table 2.2 Plasmids and Bacteriophages

Name	Classification (Origin)	Genotype	Reference/Source
pIJ702	pIJ101	*tsr mel*	Hopwood *et al.*, 1985
pIJ2520	pIJ698	*tsr hyg pwb-9*	Guthrie and Chater in press
pIJ2530	pIJ698	*tsr hyg pwb-6*	Guthrie and Chater in press
pIJ2540	pIJ698	*tsr hyg pwb-16*	Guthrie and Chater in press
pCNB1003	pIJ941	*tsr hyg red*$^{+}$	Malpartida *et al.*, 1990
pIC20H	pUC19	*bla*	Marsh *et al.*, 1984
pPWB6	pIC20H	*bla Eco*RI *pwb-6*	this study
pIJ6012	pIJ2926	*bla redX*	J.White, J.I.I
pIJ6013	pIJ2926	*bla redD*	J.White, J.I.I
Cosmid vectors			
pNRT104	pSF6	*str* prodigiosin gene cluster	N. Thompson

Bacteriophage vectors

M13 derived vectors

Name	Classification (Origin)	Genotype	Reference/Source
M13mp18		*lacZ*	Yanisch-Perron *et al.*, 1985
M13mp19		*lacZ*	Yanisch-Perron *et al.*, 1985
∅CF1	M13mp18	*pwb-6*	this study
∅CF2	M13mp18	*pwb-6*	this study
∅CF1*Bam*D	∅CF1	*pwb-6*	this study
∅CF2*Sma*D	∅CF2	*pwb-6*	this study
∅CF2*Bam*D	∅CF2	*pwb-6*	this study
∅CF2*NcoAsp*D	∅CF2	*pwb-6*	this study
∅CF3	M13mp18	*pwb-9*	this study
∅CF4	M13mp18	*pwb-9*	this study
∅CF5	M13mp18	*pwb-16*	this study
∅CF6	M13mp18	*pwb-16*	this study
∅CF7	M13mp18	pwb^+	this study
∅CF8	M13mp18	pwb^+	this study
∅CF9	M13mp19	*pwb-6*	this study
∅CF10	M13mp19	pwb^+	this study
∅CF9*Hinc*D	∅CF9	*pwb-6*	this study
∅CF10*Hinc*D	∅CF10	pwb^+	this study
∅CF11	M13mp19	pwb^+	this study
∅CF12	M13mp18	pwb^+	this study

ØC31 derived vectors

Name	Classification (Origin)	Genotype	Reference/Source
ØKC900	ØKC860	*actI vio tsr*	Bruton *et al.*, 1991
ØKC901	ØKC860	*actI vio tsr*	Bruton *et al.*, 1991
ØKC902	ØKC860	*redX vio tsr*	Guthrie and Chater, 1990
ØKC903	ØKC860	*redX vio tsr*	Guthrie and Chater, 1990
ØKC899	ØKC861	*redD*	J.White, J.I.I
Ø2	ØKC861	*pwb-9*	K.Chater, J.I.I
Ø3	ØKC861	*pwb-9*	K.Chater, J.I.I
ØL4(7)	ØC31	*redA-E*	J.Feitelson

2.1.1 Growth of *Streptomyces*

Spores were prepared from a sporulating lawn of *Streptomyces* according to the method of Hopwood *et al.*, (1985). To inoculate agar plates typically a 20µl aliquot of the spore preparation was used. All *Streptomyces* strains were grown on liquid and solid medium at 30°C.

Mycelial preparations were used in strains which did not sporulate. To prepare a mycelial inoculum, each agar plate was overlaid with a sterile dialysis tubing disc, onto which the inoculum was spread. Dialysis or Visking tubing (2½ inches in diameter) was obtained from Scientific Industries International inc. Discs were cut out from the tubing and were boiled in two changes of water before sterilisation. The dialysis tubing prevents the

mycelium from penetrating into the agar, which allows the removal of all the cellular material from the plate using an alcohol-flamed microspatula. The cellular material was resuspended in 20% glycerol. To inoculate an agar plate a 50µl aliquot is sufficient. Before inoculating into liquid cultures, the mycelia were sonicated for 30 sec to disperse clumps of cells. Sonication was performed using a Soniprep 150 with fine sonicator tip tuned to 50%. The sonicator tip was sterilised by sonicatng in ethanol for a few seconds.

Liquid cultures were grown in baffled flasks at 30°C and 180 rpm in a shaking incubator. Dispersed growth of *Streptomyces* in liquid culture was achieved using the supplemented minimal medium and method of Strauch *et al.* (1991).

2.2 Media

All solutions and media were made up using double distilled water from the Elgastat Spectrum (Elga) water purifying system. Items in parentheses were added after autoclaving. Difco Bacto agar was added to solid media at 15g per litre.

2.2.1 *E. coli* media

	Reference
Luria Bertani medium (LB)	Sambrook *et al.*, 1989
M9	Sambrook *et al.*, 1989
2YT	Hopwood *et al.*, 1985

	Constituents in g l^{-1}
Nutrient Broth	Difco nutrient broth 8g
Nutrient agar	Difco nutrient agar 23g
Top agar	Difco Bacto-tryptone 10g NaCl 8g Difco Bacto-agar 8g

2.2.2 *Streptomyces* media

Streptomyces Minimal medium (NMM)

All constituents are given in g l^{-1}

Ammonium sulphate	1g
diPotassium hydrogen phosphate	0.5g
Magnesium sulphate	0.2g
Iron(VI)sulphate	2 crystals
Difco agar	15g

After autoclaving glycerol was added to a final concentration of 10mM and 150ml l^{-1} 0.1M Phosphate buffer pH6.8.

0.1M Phosphate Buffer was made as follows:

Solution A 0.1M NaH_2PO_4 (13.8g l^{-1}) or 0.1M $NaH_2PO_4.H_2O$ (15.6g l^{-1})

Solution B 0.1M K_2HPO_4 (17.4g l^{-1})

To acheive pH 6.8 25.5ml of solution A was added to 24.5ml of solution B and the volume made up to 100ml with Elga water.

Melanin production medium	Paget *et al.*, 1994
R2YE	Hopwood *et al.*, 1985
YEME	Hopwood *et al.*, 1985
Supplemented Minimal Medium (SMM)	Strauch *et al.*, 1991

Silica based medium

Syton (from Du Pont, U.K.) was dialysed against two changes of 0.05M phosphate buffer at 30°C for 24 hours, then dialysed again against two changes of Elga water at room temperature for 24 hours. The syton silica was autoclaved with the constituents below:-

10 x NMM	2.5ml
Dialysed Syton silica	13.2ml
H_2O (Fisons)	5.5ml

After autoclaving 3.75ml of 0.1M phosphate buffer which had been autoclaved separately and glycerol as carbon source to a final concentration of

10mM were added. Proline, when used as sole nitrogen source, was added at this point to a final concentration of 37µg/ml. Where proline was used as sole nitrogen source the $(NH_4)_2SO_4$ was omitted from the NMM salt solution (U.Swoboda personal communication).

2.3 Proline permease assays

2.3.1 Growth on proline as sole nitrogen source

To determine if cells were able to transport proline, growth on silica based medium with proline as sole nitrogen source and ammonium as sole nitrogen source was assessed. Spores were washed with 20% glycerol made up in Fisons water before inoculation onto plates. This was a precaution to prevent nitrogen contamination being carried over in the inoculum. Where a dilution series of the inoculum was used, it was prepared using 20% glycerol, again made up with Fisons water. Nitrogen concentrations in the glycerol are negligible.

2.3.2 Resistance to proline analogues

A 5µl aliquot of the desired concentration of either azetidine-2-carboxylate (AZC) or 3,4-dehydroproline (DHP) was spotted onto a sterile filter disc placed on a minimal medium agar plate. AZC aliquots at concentrations of 0.5M, 0.25M or 0.1M were used. Dehydroproline was used at concentrations of 0.05M, 0.025M or 0.01M. The agar plates were incubated at 30°C for at least 48 hours before noting the zones of inhibition of growth.

2.3.3 ^{14}C proline transport

Cells were harvested from minimal medium agar plates that had been incubated for 72 hours. The agar plate was overlaid with a dialysis tubing disc before inoculation, to make harvesting the mycelium easier. The mycelium was scraped off the dialysis tubing asceptically using a flamed microspatula. The cells were washed twice in NMM, before resuspending in 400µl of NMM (minus carbon source).

To assay ^{14}C proline transport, the cells were first added to the constituents below:-

Assay mix

Cell Suspension	200µl
Chloramphenicol (3mg/ml)	1.5µl
Glucose (10%)	4.0µl
NMM	190.5µl

The last addition to the assay mix was 4µl of ^{14}C proline, the assay mixture was vortexed before removing 30µl aliquots at timed intervals. The aliquots were spotted onto Whatman 25mm glass fibre filters. The filters were held in a Millipore sampling manifold (1225) attached to a vacuum pump to assist filtering and washing of cells. The filters were washed with 2ml of NMM containing 10mM KCN. The filters were placed in scintillation vials and 3mls of scintillant (Pharmacia, Optiphase Safe) was added.

The amount of ^{14}C proline present in the cells was assessed by the counts per minute registered by the scintillation counter (Beckmann LS7000). A 4µl aliquot of ^{14}C proline was spotted onto a filter, scintillant added and counts

measured in the scintillation counter to assess the efficiency of counting. The results were given in ρmoles proline per mg protein using the calculations below.

2.3.4 Determination of the efficiency of counting ^{14}C proline

The expected disintegrations min^{-1} can be calculated knowing the level of radioactivity of the ^{14}C proline solution used. The ^{14}C solution used was 50μCi / ml, this is equivalent to 2.2×10^6 Bq/ml. The sample tested in the scintillation counter was 4μl. The expected disintegrations from 4μl was 4.4×10^5. This value was compared to the actual value as recorded by the scintillation counter. A typical result obtained was 3.91×10^5.

$(3.91 \times 10^5 / 4.4 \times 10^5) \times 100 = 88.8\%$ efficiency of counting

2.3.5 Determination of the amount of ^{14}C Proline transported into the cell

Measurements from the scintillation counter were received in the units of counts per minute. The results were calculated using the calculation below which converts the readings into ρ mol proline/mg protein. This is done knowing the disintegrations expected per ρmol proline of the solution of ^{14}C proline used which was 9.77 disintegrations/ρmol proline. The calculation also corrects the reading to take account of the counting efficiency recorded by the scintillation counter.

CPM represents the readings obtained from scintillation counter in counts per minute. The value 0.88 corresponds to the counting efficiency as calculated above. The reading of counts per minute were converted to counts per seconds, or Bq by dividing by 60.

$$\text{pmol proline/mg protein} = \frac{\textbf{CPM}}{\textbf{0.88 x 60 x 9.77 x protein concentration (in mg)}}$$

The concentration of protein was calculated by determining the concentration present in the remaining 200µl of harvested cells. The protocol used was based on the method of Lowry, modified as described by Herbert *et al.* (1971) for using suspensions of micro-organisms. The protein concentration of a 30µl aliquot of cells was calculated and used in the equation above

2.4 Transformation of *E.coli* and *Streptomyces*

2.4.1 Calcium chloride transformation of *E. coli*

A 10ml aliquot of LB was inoculated with *Escherichia coli* and grown overnight at 37°C with shaking. This overnight culture was used to inoculate another 5ml culture which was grown with shaking at 37°C until mid-logarithmic phase (OD_{600} of between 0.3-0.4). Cells were made competent according to the method of Sambrook *et al.* (1989) section 1.82.

Typically 100µl of competant cells were mixed with 10-50ng of DNA and left on ice for 30 minutes before being placed at 42°C for 2 minutes. Where expression of a plasmid encoded antibiotic resistance determinant was required, 1ml of LB was added to the cells and they were incubated for a further hour at 37°C. The cells were then mixed wth 3ml of top agar in a sterile bijoux bottle and used to overlay an LB agar plate. Where vectors were being transformed that carried the β-galactosidase gene (*lacZ*) to enable blue-white selection 25µl of IPTG (25mg/ml) and 25µl of X-gal (20mg/ml in DMF) were added to the top agar overlay.

2.4.2 DMSO transformation of *E.coli*

A 10ml logarithmic phase culture prepared as described in section 2.4.1 was treated according to the method of Chung *et al.* (1989). This is a one-step procedure for the preparation of competent *E. coli* cells using a storage and transformation buffer (TSS) described below:

TSS Buffer

PEG 6000	2g
DMSO	1ml
1M $MgSO_4$	600µl

LB was added to make the volume up to 20 ml.

The cells were centrifuged in a Wifug Labor-50M bench centrifuge at 3000rpm for 10 minutes. The pellet was resuspended in 1ml of ice-cold TSS. A 100µl aliquot of cells was added to the DNA to be transformed in an eppendorf tube and incubated on ice for 30 minutes. Where expression of a plasmid encoded antibiotic resistance determinant was required, 1ml of LB

was added to the cells and they were incubated for a further hour at 37°C. The cells were plated onto LB agar as described in section 2.4.1.

2.4.3 Electroporation of *E.coli*

A 10ml logarithmic phase culture of *E.coli* was prepared as described in section 2.4.1. The cells were pelleted in a bench centrifuge (Wifug Labor-50M) at 3000rpm for 10 minutes at 5°C. The cells were resuspended in 10ml ice cold sterile water and pelleted (3000rpm for 10 minutes). The cells were washed a further time with ice cold water as before and were then resuspended in 1ml of 10% glycerol. The cells were pelleted and resuspended in a final 1ml of 10% glycerol. For each electroporation 40µl of prepared cells were used, these were mixed with 10-50ng of DNA in a Bio-Rad electroporation cuvette (0.2cm electrode). Electroporation was acheived using a Bio-Rad Gene Pulser set at a resistance of 200Ω and a capacitance of 25µF. Cells were subjected to a potential difference of 2.5kV. Immediately 1ml of 2YT was dispensed into the cuvette and the cells were incubated in the cuvette at 37°C for 1 hour. The cells were pelleted and inoculated onto an LB agar plate as section 2.4.1

2.4.4 Transformation of *Streptomyces*

A liquid culture of *Streptomyces* was grown in YEME supplemented with a final concentration of 0.5% glycine. Protoplasts were prepared according to the protocol of Hopwood *et al.* (1985) page 12. Transformation was carried out using the Rapid small scale method as described in Hopwood *et al.*, 1985.

2.4.5 Preparation of lysogens

Lysogens were prepared using the method described by Hopwood *et al.*, (1985) procedure B on p23. Lysogens were grown in YEME supplemented

with 5μg/ml thiostrepton. An R2YE agar plate inoculated with *S.lividans* TK24 spores onto which a 5μl aliquot of the lysogen growth medium was spotted. A zone of lysis on the TK24 lawn where the growth medium was spotted was taken as verification that the cells grown were lysogens

2.5 Molecular techniques

2.5.1 Restriction endonuclease digestion of DNA

All restriction endonuclease digests of DNA were performed using enzymes purchased from Gibco BRL and were carried out using the buffers supplied, in accordance with the manufacturer's instructions.

2.5.2 Agarose gel electrophoresis

Agarose gel electrophoresis was used to separate and visualise DNA fragments. Gels were prepared using 0.7% agarose buffered with TBE and containing 0.5μg/ml ethidium bromide. Gels were routinely run overnight at a constant 35 Volts. DNA was visualised using a long wave transilluminator. Markers used to allow size estimation of DNA fragments were either a 1kb ladder purchased from Gibco BRL or λ DNA (Gibco BRL) restricted with *Hin*dIII. When required, DNA fragments were recovered from agarose gels using the Geneclean II kit from Stratech Ltd. according to the manufacturer's instructions.

2.5.3 Dephosphorylation of vector DNA

Alkaline phosphatase from calf intestine was purchased from Boehringer Mannheim and was used with the buffer supplied, in accordance with manufacturer's instructions.

2.5.4 DNA ligation

T4 DNA ligase from *E.coli* was purchased from Gibco BRL and used with the ligation buffer described by Sambrook *et al.* 1989 section 1.68.

2.5.5 Fragment end repair

End-filling of 5' overhangs was performed using the Klenow fragment of *E.coli* DNA polymerase I purchased from Gibco BRL, according to the method described by Sambrook *et al.* (1989).

2.6 DNA Preparation

2.6.1 Plasmid preparation from *E.coli*

Small scale plasmid preparation was performed by the alkaline lysis method as described by Sambrook *et al.* (1989) section 1.25.

Large scale plasmid DNA preparation was conducted using the lysis by alkali method described by Sambrook *et al.* (1989) section 1.38. Purification of plasmid DNA was achieved using caesium chloride density gradient centrifugation as described in section 1.42 of Sambrook *et al.* (1989).

2.6.2 *Streptomyces* plasmid preparation

Large scale plasmid preparation was performed using the technique of alkaline lysis as described by Hopwood *et al.* (1985) (procedure 2) from page 87-89. A scaled down version of the same procedure was used for small scale plasmid preparation also described by Hopwood *et al.* (1985) page 85-87.

2.7 Di-deoxy DNA sequencing

2.7.1 Single stranded DNA template preparation

Single stranded DNA was prepared from turbid plaques the day after transfection of *E.coli* strain TG1 with M13 DNA according to the method of Sambrook *et al.* (1989) section 13.42.

2.7.2 Sequencing reactions

DNA sequencing was based upon the method of Sanger *et al.* (1977). The Sequenase version 2.0 T7 DNA polymerase kit, obtained from USB Corporation was used according to the manufacturer's instructions. The DNA fragments produced by this method were separated using a 6% polyacrylamide buffer gradient gel as described in Sambrook *et al.* (1989) section 13.54.

The gel was run for approximately 3 hours after which it was immersed in 10% acetic acid for 15 minutes. The gel was transferred to 3MM Whatman paper, covered in cling film and dried at 80°C under vacuum in a gel drier. Once the cling film had been removed the gel was left over-night to expose a sheet of X-ray film

2.8 S1 nuclease mapping

RNA was prepared from cells grown in SMM liquid medium using the method of Strauch *et al.* (1991) for dispersed growth. RNA isolation was performed using the method of Hopwood *et al.* (1985)

The *redD* probe was made by uniquely labelling the 1.3kb *Cla*I *Nde*I fragment from pIJ6013 at the 5' end of the *Cla*I site with ^{32}P. The *redX* probe used was a 220bp *Bam*HI *Sma*I fragment from pIJ6012 which was uniquely labelled with ^{32}P at the 5' end of the *Sma*I site. Both *redD* and *redX* probes were labelled with ^{32}P according to the method of Maxam and Gilbert (1980).

For each S1 nuclease reaction 10-30μg of RNA was hybridised in NaTCA buffer (Strauch *et al.*, 1991) to approximately 0.02ρmol of labelled probe All subsequent steps were as described by Strauch *et al.* (1991).

CHAPTER 3

STUDY OF PROLINE TRANSPORT MUTANTS

3.1 Introduction

Proline transport mutants (*put*) were isolated as documented by Hood *et al.* (1992). A summary of their findings and further study of the interaction between the Put⁻ phenotype and undecylprodigiosin (Red) biosynthesis is presented.

Study of the Put⁻ phenotype in *Streptomyces coelicolor* A3(2) began as an investigation into the source of proline molecules that are incorporated into the bi-pyrrole sub-unit of Red (Wasserman *et al.* 1974). The hypothesis that the source of proline was from the extracellular environment was proposed. This was tested by selecting for mutants which were unable to transport proline into the cell, so that their ability to synthesise Red could be determined.

3.1.1 Isolation of proline transport mutants

The Put⁻ mutants were selected using proline analogues azetidine-2-carboxylate (AZC) and dehydroproline (DHP), an approach used by Maloy (1987) to isolate proline permease (*putP*) mutants in *E. coli*. The proline analogues are toxic to the cell when incorporated into proteins. Proline catabolic enzymes can oxidise DHP slowly to a non-toxic product, however, they are unable to metabolise AZC. The starting strain for the isolation of proline transport mutants in *S. coelicolor* was J802 (*agaA7* SCP2* NF). This strain was chosen because the *agaA7* lesion would abolish agar metabolism, which might obscure the results of proline catabolism as sole carbon source on solid media. Further investigation showed that J802 was unable to utilise the five carbon amino acids glutamate, glutamine, histidine, arginine and proline as sole carbon sources. The *aga* lesion was implicated in this debility as an

aga$^+$ isogenic strain, J801 was shown to be able to utilise the amino acids as sole carbon sources. In addition strain J802 in the presence of five carbon amino acids proline, glutamate, glutamine and histidine was unable to use acetate, pyruvate and Krebs cycle intermediates.

Mutants that utilise proline and would therefore have greater sensitivity than J802 to proline analogues were sought. The proline utilising (PU) mutants were selected from J802 after an NTG mutagenesis by the ability to utilise proline as sole carbon source on silica medium. The mutants of J802 generated by NTG mutagenesis were tested for their resistance to the proline analogue AZC. Proline utilising mutants are sensitive to proline analogues and in fact are sensitive to a wider range of proline analogues than J802 these include; atropine, DL-pyroglutamic acid and pyrrmole-2-carboxylic acid, in addition to DHP. The proline utilising mutants were used to isolate mutants defective in proline transport. This was achieved by selecting for PU mutants that acquired resistance to toxic proline analogues (Figure 3.1).

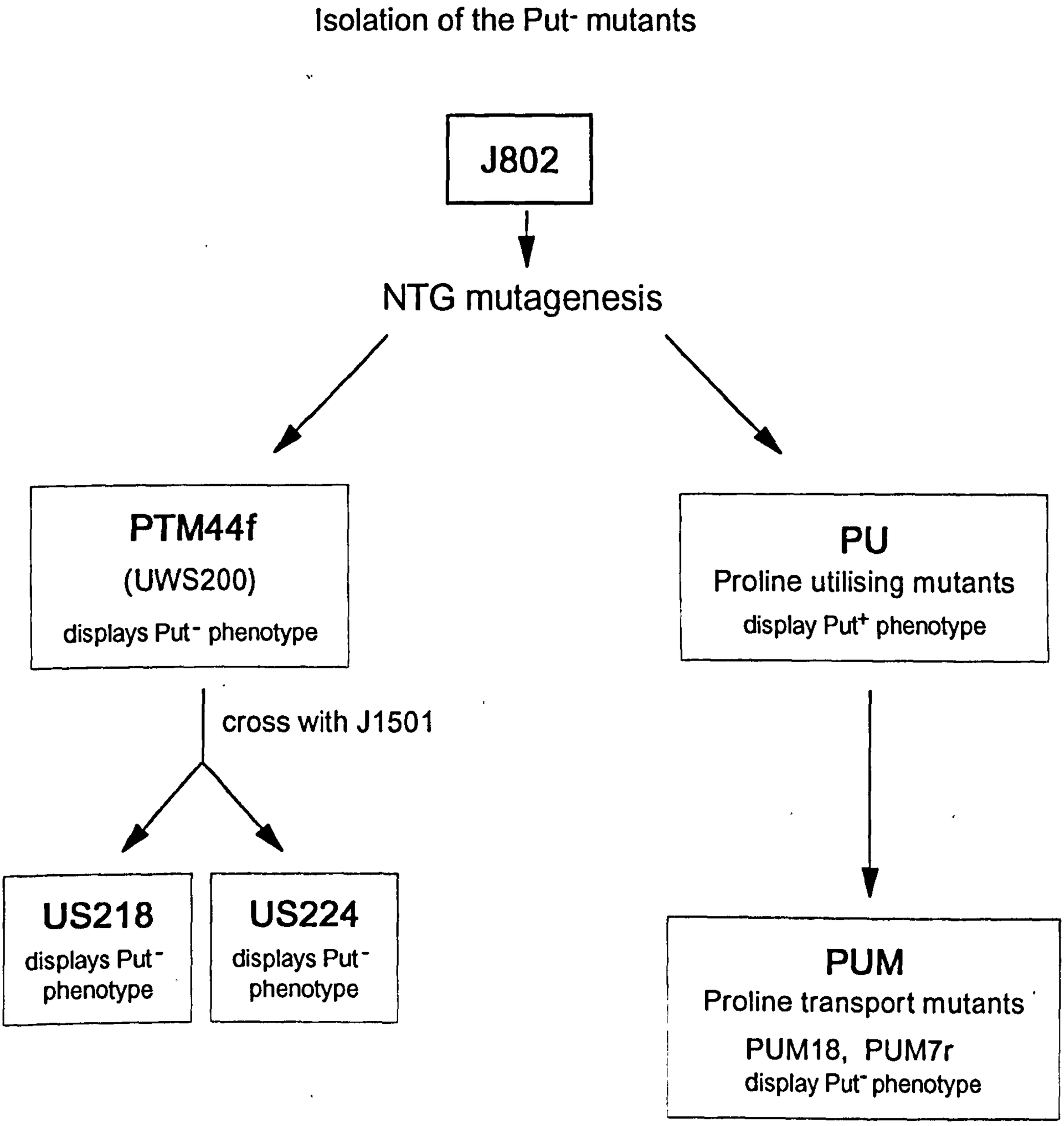
One mutant PTM44f (originally named UWS200) showed the phenotype expected of a proline permease mutant; resistance to DHP in addition to resistance to AZC and inability to use proline as a sole nitrogen source. The mutant PTM44f was isolated directly from J802 after NTG mutagenesis (Figure 3.1). In addition to the proline transport Put$^-$ phenotype, PTM44f did not sporulate and could not be used to make protoplasts. A cross between PTM44f and J1501 was performed in order to try to isolate recombinant Put$^-$ strains which were more amenable. The strains US218 and US224 were the progeny of the cross and are the Put$^-$Pgl$^-$ strains used in this study.

Figure 3.1

Isolationof the Put⁻ mutants

The starting strain for the isolation of the proline transport mutants in *S.coelicolor* was J802 (*agaA7* SCP2* NF). The *agaA7* lesion has been implicated in the debility of J802 to utilise the five amino acids glutamate, glutamine, histidine, arginine and proline. Proline utilising mutants were sought because the ability to utilise proline would cause increased sensitivity to the proline analogues. It was anticipated that this would increase the likelihood of isolating proline transport mutants. One proline transport mutant was isolated directly form J802, this was named PTM44f. A cross between PTM44f and J1501 was performed in order to isolate recombinant Put⁻ strains which were Pgl⁻.

Figure 3.1


Isolation of the Put- mutants
J802
NTG mutagenesis
PTM44f
(UWS200)
displays Put- phenotype
PU
Proline utilising mutants
display Put+ phenotype
cross with J1501
US218
displays Put- phenotype
US224
displays Put- phenotype
PUM
Proline transport mutants
PUM18, PUM7r
display Put- phenotype

The PU mutants were used to select for proline transport mutants which were resistant to the proline analogues. These second round mutants were designated PUM mutants (Figure 3.1).

3.1.2 Phenotype of the proline transport mutants

In addition to the expected phenotype of the proline transport mutants; resistance to proline analogues, abolition of growth on proline as sole nitrogen source and inability to import ^{14}C proline, the Put$^-$ mutants over-produced Red. This result shows that Red biosynthesis is not dependent on the supply of exogenous proline.

However, PTM44f and the PUM mutants (Put$^-$ mutants) exhibited a pleiotropic phenotype. None of the Put$^-$ mutants were capable of growth in any liquid or on complex media. The Put$^-$ mutants were all capable of growth in the presence of proline showing that they were not proline sensitive. The PUM mutants over-produced Red to a greater extent than PTM44f. The appearance of PTM44f is pink which is partly due to the absence of actinorhodin (Act) production, as only marginally more Red is produced as compared to J802. None of the PUM mutants or PTM44f, (as previously mentioned), were capable of sporulation. Both US224 and US218 were capable of sporulation, growth in liquid media and could be used to prepare protoplasts that were capable of regenerating hyphae.

The ability to catabolise proline in the Put$^-$ mutants was assessed initially by using the ability to use dipeptides leucyl-proline and glycyl-proline as carbon and nitrogen sources. The peptides were transported into the cell via the peptide transport system (Leigh, 1992). *S.coelicolor* is incapable of using

either glycine or leucine as sole nitrogen source. Therefore, grówth of cells on the peptides used was due to growth on the proline moiety of the peptide. The results presented on Table 3.1, show J802 could utilise the proline moiety of the peptide for growth whilst PUM7r, PUM18 and PTM44f could not. The enzymes involved in proline catabolism; proline oxidase and pyrroline-5-carboxylate dehydrogenase (P5C-dehydrogenase) were assayed individually. A summary of the results obtained by U.Swoboda is presented on Table 3.1. Proline oxidase was found to be membrane associated and was easily separated from the cytoplasmic P5C dehydrogenase. The proline catabolic enzymes of *S.coelicolor* (proline oxidase and P5C-dehydrogenase) are separate and are not present as a bi-functional enzyme as is the case in *Pseudomonas aeruginosa*, *Salmonella typhimurium* and *Escherichia coli* (Smith *et al.*, 1995). Proline biosynthesis has previously been shown to be constitutive and not subject to feedback repression in *S. coelicolor* A3(2) (Hood *et al.*, 1992).

It was proposed that proline accumulated within the cell in these mutants as proline biosynthesis was constitutive and proline catabolic enzymes were absent or reduced.

3.2 Interaction between Put phenotype and Red biosynthesis

A hypothesis was proposed that proline anabolism and catabolism are in dynamic equilibrium and if this equilibrium is disturbed, as in the Put$^-$ mutants, then Red biosynthesis acts as a sink for excess proline. It is the aim of this work to test this hypothesis. It was proposed to inactivate Red biosynthesis in a Put$^-$ mutant. If Red biosynthesis is required to reduce the proline concentration, then a Red$^-$Put$^-$ mutant might not be viable. In Put$^-$

Table 3.1

A Summary of the origin and phenotype of Put⁻ mutants

A number of tests were performed to characterise the Put⁻ mutants isolated. These included the ability to transport ^{14}C proline and the activity of proline catabolic enzymes proline oxidase and P5C dehydrogenase. The ability of the Put⁻ mutants to metabolise the proline moiety of glycl-proline or leucyl-proline was also tested. Proline transport into the cell is severely reduced in the Put⁻ mutants. To examine the activity of the proline catabolic enzymes, growth on proline containing dipeptides of glycine and leucine was determined. The dipeptides can enter the cell where they are broken down, *S.coelicolor* cannot utilise leucine or glycine as sole nitrogen source. Any growth detected is as a result of growth on proline as sole nitrogen source. This implies the proline catabolic enzymes are functional.

Table 3.1

Summary of the isolation and phenotype of Put$^-$ mutants.

Strain	Parent	*put* allele	Analogue used for isolation	Colour	Proline transport	Growth on dipeptide as N-source	Proline oxidase	P5C dehydro-genase
J802	n/a	Put$^+$	n/a	deep blue to purple	+	+	+	+
PTM44f	J802	*put*-44	AZC	pink	-	nt	reduced	reduced
PUM18	PU56	*put*-18	ATN	red	-	+	reduced	-
PUM7r	PU76	*put*-7	ATN	deep red	-	-	-	-
US224	PTM44f	*put*-44	AZC	red / purple	-	nt	nt	nt
US218	PTM44f	*put*-44	AZC	red / purple	-	nt	nt	nt

KEY

AZC	Azetidine-2-carboxylate
ATN	Atrophine
+	Activity detected
-	No activity detected
reduced	reduction of activity by comparison to J802
nt	not tested

cells, in the absence of proline catabolic enzymes, proline is accumulated which we propose is shunted into Red biosynthesis. If a Red$^-$Put$^-$ strain was isolated it would pose the questions; what effect does this have on the proline concentration? and how is excess proline removed?

3.2.1 Strategy for isolating a Red$^-$ Put$^-$ mutant

The approach used to inactivate Red biosynthesis in a Put$^-$ mutant was mutational integration, using modified ØC31 actinophage vectors (Bruton *et al.* 1991). The modified ØC31 phage have the normal phage DNA attachment site deleted (*attP*$^{\Delta}$). A DNA fragment internal to the gene to be disrupted is present on the phage, in this instance *redX*, a Red structural gene. Lysogens are obtained when homologus recombination occurs between the chromosome and the cloned DNA fragment (Figure 3.2). Disruption of the *redX* gene abolishes the biosynthesis of Red. The phage also carries viomycin and thiostrepton resistance genes so that lysogens can be easily selected. Disruption of Red production can be visually confirmed by the absence of Red pigment. The internal *redX* DNA fragment was located upstream of a promoterless *xylE* reporter gene. This was designed so that on recombination of the phage with chromosomal DNA the *xylE* gene was transcriptionally coupled to the *redX* promoter. The *xylE* gene product (catechol 2,3-dioxygenase) converts colourless catechol into yellow 2-hydroxymuconic semialdehyde which can be assayed spectrophotometrically to indicate the promoter activity of *redX*. A previous attempt to inactivate Red biosynthesis in a Put$^-$ background using ØL4(7), a ØC31 derivative containing a 2.2kb *Bam*HI *Sst*I fragment from the *redA-E* region of the *red* cluster, failed to generate Put$^-$Red$^-$ lysogens. Of the limited number of Red$^-$ lysogens isolated

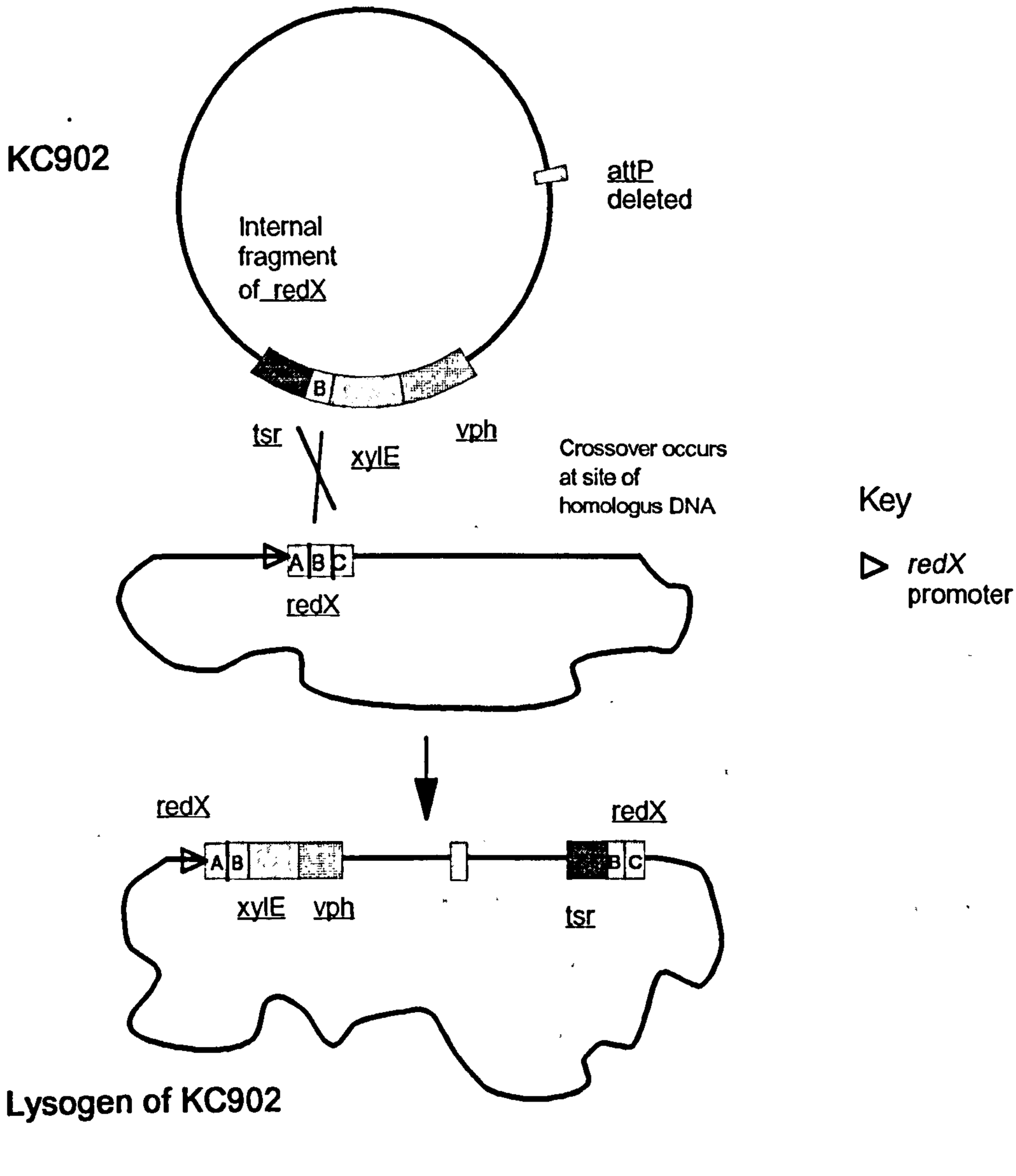
Figure 3.2

Insertional Mutagenesis

Insertional mutagenesis is a technique whereby a target gene can be inactivated. An internal DNA fragment of the gene to be disrupted is cloned onto a ØC31 altered phage. The attachment site *attP* is deleted, insertion of the phage DNA can only occur via homologus recombination of the cloned internal DNA fragment on the phage and the chromosomal copy of the gene. The phage DNA is inserted into the chromosome disrupting the target gene.

In this instance an internal fragment of *redX* has been cloned onto ØKC902, with the intention of inactivating Red biosynthesis. ØKC902 carries two selectable markers, viomycin (vph) and thiostrepton (tsr) resistance genes, so that lysogens can be selected. The internal fragment of the *redX* gene has been cloned upstream of a promoterless *xylE* gene. The construction of Ø KC902 is such that the *redX* promoter is fused to the promoterless *xylE* gene (Guthrie and Chater, 1990). Promoter activity of *redX* can be determined by assaying *xylE* activity of ØKC902 lysogens.

Figure 3.2


Insertional Mutagenesis
KC902
Internal fragment of redX
attP deleted
B
tsr
xylE
vph
Crossover occurs at site of homologus DNA
Key
redX promoter
A B C
redX
redX
redX
A B
xylE vph
B C
tsr
Lysogen of KC902

all were, on further study, shown to have reverted to a Put^+ phenotype (U. Swoboda personal communication).

When designing this experiment the previous work was considered. Phage containing either DNA internal to a *red* gene (*redX*) or an actinorhodin gene (*actI*) were used. The phage constructs were already available and were kindly supplied by K.Chater (ØKC900 and ØKC901; Bruton *et al.* 1991; ØKC902 and ØKC903 Guthrie and Chater 1990). The phage ØKC900 contains DNA internal to *actI* in the orientation such that the *actI* promoter is fused to the promoterless *xylE* gene, in ØKC901 the *actI* DNA is in the opposite orientation, with respect to the *xylE* gene. Similarly ØKC902 and Ø KC903 contain opposite orientations of the *redX* DNA with respect to the promoterless *xylE* gene. The phage ØKC902 contains an internal *redX* fragment such that the *redX* promoter will be fused to the promoterless *xylE* gene. Using phage to disrupt actinorhodin (Act) biosynthesis acts as a positive control to show that lysogeny and expression of foreign antibiotic resistance determinants are possible in a Put^- strain. The presence of viable Act^-Put^- mutants suggests that if Red^-Put^- mutants were not isolated, it would be because they were not viable due to the effect of disrupting Red biosynthesis and not because of problems with the lysogenisation. If Red^-Put^- mutants are isolated, the activity of the *xylE* product activity in lysogens of Put^- mutants and of the parent strain J802 could be compared. This would indicate the effect of the Put^- phenotype on *redX* expression.

The Put^- mutants used were strains US218 and US224. These and the parent strain J802 were treated with phage ØKC900, ØKC901, ØKC902 and ØKC903 to obtain lysogens (2.4.5).

3.2.2 Isolation of Put⁻Red⁻ lysogens

Put⁻Red⁻ lysogens were selected initially by screening for viomycin or thiostrepton resistance. This was confirmed by re-plating the resistant cells onto viomycin R2YE and thiostrepton R2YE media. Of the cells isolated that were resistant to both antibiotics, all were deficient in production of Act or Red, depending on the phage used, as can be seen in Figure 3.3. One explanation for the failure of previous work to isolate Put⁻Red⁻ lysogens using ØL*A*(7) (3.2.1) was that the thiostrepton resistance gene (*tsr*) was not being expressed in the Put⁻ strain. For this reason cells were selected on either viomycin or thiostrepton. The putative lysogens were then grown in liquid culture and the supernatant was tested on a lawn of *S.lividans* TK24 for the presence of free phage (2.4.5). The presence of free phage was taken as confirmation that the cells were lysogens. The numbers of lysogens isolated are indicated in Table 3.2. The relatively poor level of isolation of J802 lysogens reflects the fact this was a Pgl^+ strain whilst US218 and US224 are Pgl^- (1.4.2).

3.3 Characterisation of Put⁻Red⁻ lysogens

Three methods were used to determine whether proline was being transported into the cells.

3.3.1 Growth on proline as sole nitrogen source

The first method to determine proline transport into the cell involved growing cells on solid medium with proline as the sole nitrogen source (2.3.1). This was compared to media with ammonium as sole nitrogen source.

Figure 3.3

Photograph of Act⁻ Red⁻ US218 lysogens.

Lysogens were isolated of ∅KC900 and ∅901. These phage constructs contain DNA internal to the *actI* gene. Lysogeny occurs via homologus recombination of the internal *actI* DNA present on the phage and the chromosomal *actI*. In lysogens the phage DNA has become inserted into the chromosome disrupting the *actI* gene which abolishes Act production as can be seen in the photograph. The Act⁻ lysogens appear Red.

The phage ∅KC902 and ∅KC903 contain DNA internal to the *redX* gene. Red production is abolished in lysogens of ∅KC902 and ∅KC903 as can be seen on the facing page.

Figure 3.3

Photograph of Act⁻ and Red⁻ US218 lysogens.

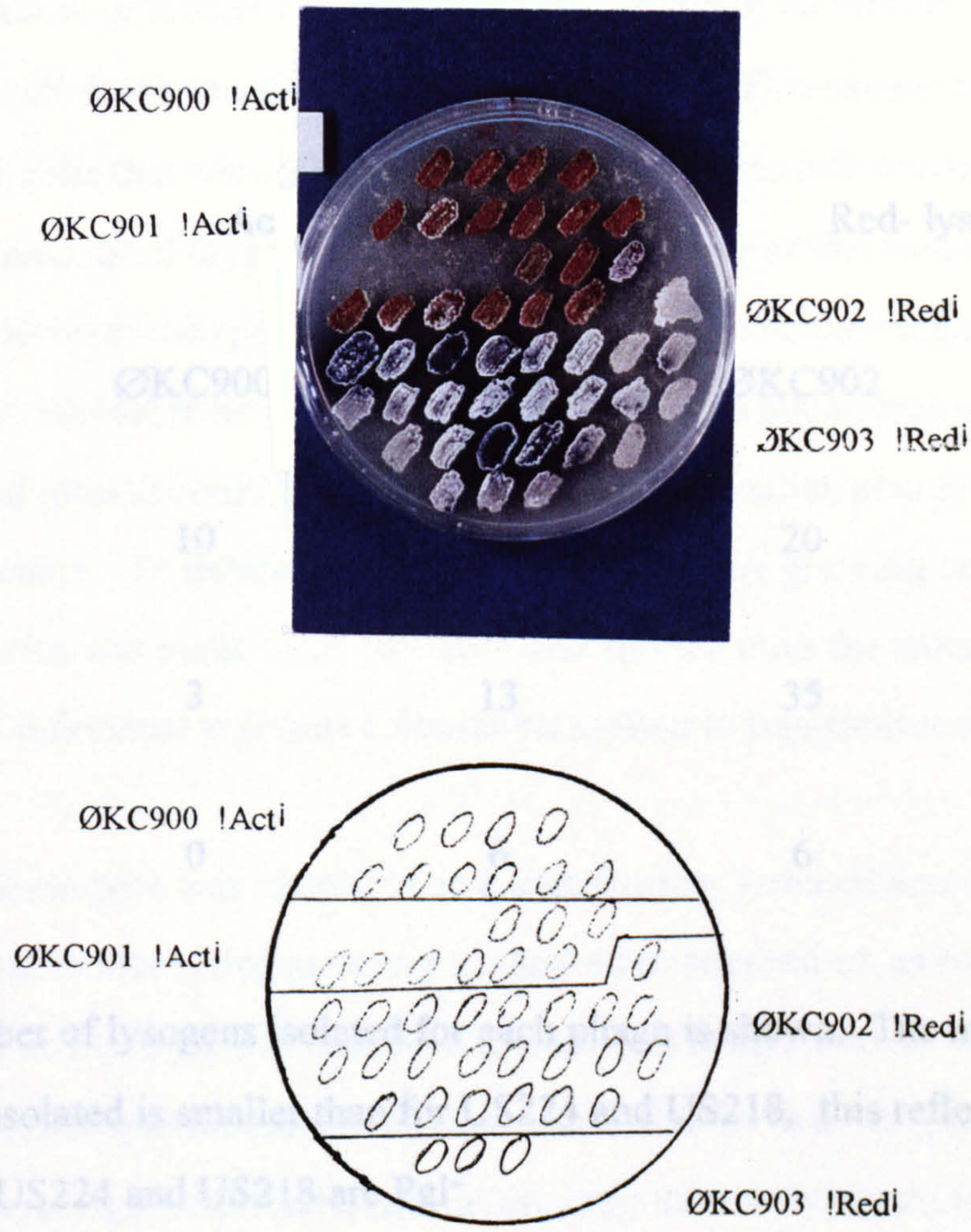


Table 3.2

Isolation of lysogens of US218 US224 and J802 with phages ØKC900, ØKC901, ØKC902 and ØKC903.

	Act⁻ lysogens		Red- lysogens	
	ØKC900	ØKC901	ØKC902	ØKC903
US218	10	9	20	42
US224	3	13	35	23
J802	0	6	6	3

The number of lysogens isolated for each phage is shown. The number of J802 lysogens isolated is smaller than for US224 and US218, this reflects that J802 is Pgl^+ and US224 and US218 are Pgl^-.

The medium used contained silica (Syton) as a gelling agent, in order to prevent growth on agar contaminants from interfering with the results. Even using silica based medium and preparing all media and solutions with purified water, a small amount of growth was observed on plates to which no nitrogen source had been added. This was attributed to streptomycetes being very effective scavengers for nitrogenous compounds in the medium, carried over in the inoculum or released from dead spores. Growth on nitrogen released as a result of cell death was defined as cryptic growth. This causes a problem in identifying cells that were unable to grow on proline as sole nitrogen source, as a small amount of cryptic growth is always seen. For this reason great care was taken to wash the spores in purified water before inoculating the silica medium. A control plate where no nitrogen had been added was used so that background growth could be compared to growth on silica plus proline as sole nitrogen source. To determine if patches of cells were growing cryptically, a dilution series was made from 10^{0}-10^{-8} and spotted onto the silica medium. Growth of individual vigorous colonies was taken to represent true growth.

A Put^{-} phenotype was identified as a significantly reduced amount of growth with proline as sole nitrogen source compared to ammonium as sole nitrogen source but no more growth than with no nitrogen source added. Of the lysogens tested approximately 75% were shown to have a Put^{-} phenotype. There was no significant difference in the ratio of Act^{-} lysogens to Red^{-} lysogens shown to be Put^{-}, as can be seen from the summary of results shown in Table 3.3. An example of Put^{+} and Put^{-} cells as determined by growth on silica with ammonium and silica plus proline as sole nitrogen source is given in Figure 3.4.

Table 3.3

A Summary of proline transport tests performed on lysogens

Lysogens of J802 and of Put⁻ strains US218 and US224 were assayed to determine their Put $^{+/-}$ phenotype.

Phage ØKC900 and ØKC901 contained DNA internal to the *actI* gene. Actinorhodin production is abolished in lysogens of these phage. Phage ØKC902 and ØKC903 contained DNA internal to the *redX* gene and lysogens of these phage are unable to synthesise Red.

Growth on Syton silica based medium with proline as sole nitrogen source is shown. Growth is indicated by (+), absence of growth is indicated with (-). Proline transport mutants are unable to grow in media with proline as sole nitrogen source. Some lysogens failed to grow on this media regardless of nitrogen source, these are indicated on the table as ng (no growth).

Proline transport mutants were originally isolated by their resistance to proline analogues. Resistance (R) or sensitivity (S) to the proline analogue 3,4-dehydroproline of the lysogens is shown in the table.

^{14}C proline transport was assayed. Cells that transported proline as efficiently as the parent strain J802 were scored in the table as (+). Proline transport mutants are identified as those which transport proline at a reduced level to J802, this is scored as(-). There were cases where replicate assays gave conflicting results of proline transport. These are marked -/+ on the table to indicate this uncertainty.

The overall Put$^{+/-}$ phenotype, as determined by the three assays is given in the final column.

Table 3.3

A Summary of proline transport tests performed on lysogens.

Lysogen	Strain	Phage	Act^-/Red^-	Syton	$DHP^{R/S}$	^{14}Cproline	$Put^{+/-}$
A1	US218	ØKC900	Act^-	-	nt	nt	-
A2	US218	ØKC900	Act^-	+	nt	nt	+
A3	US218	ØKC900	Act^-	-	R	-/+	+/-
A14	US218	ØKC901	Act^-	ng	S	nt	+
A15	US218	ØKC901	Act^-	ng	R	nt	-
A17	US218	ØKC901	Act^-	-	R	+	+/-
A24	US218	ØKC902	Red^-	-	R	-	-
A25	US218	ØKC902	Red^-	-	S	nt	+
A26	US218	ØKC902	Red^-	-	nt	nt	-
A45	US218	ØKC903	Red^-	-	R	-	-
A47	US218	ØKC903	Red^-	-	R	-	-
A48	US218	ØKC903	Red^-	nt	R	nt	-
B21	US218	ØKC903	Red^-	-	nt	nt	-
B22	US218	ØKC903	Red^-	ng	R	nt	-
B23	US218	ØKC903	Red^-	-	nd	nt	-
B37	US218	ØKC902	Red^-	ng	R	nt	-
B38	US218	ØKC902	Red^-	-	nt	nt	-
B39	US218	ØKC902	Red^-	-	R	nt	-
Q5	US224	ØKC900	Act^-	nt	R	-	-
I45	US224	ØKC901	Act^-	nt	S	nt	+
C34	US224	ØKC901	Act^-	nt	R	-	-
G43	US224	ØKC902	Red^-	nt	R	-	-
H16	US224	ØKC902	Red^-	nt	R	+/-	+/-
H37	US224	ØKC903	Red^-	nt	R	+/-	+/-
H38	US224	ØKC903	Red^-	nt	R	nt	-

Table 3.3 continued

Lysogen	Strain	Phage	Act⁻/Red⁻	Syton	DHP	^{14}Cproline	
I13	US224	ØKC903	Red⁻	nt	R	nt	-
I14	US224	ØKC903	Red⁻	nt	R	+	+
I15	US224	ØKC903	Red⁻	nt	R	nt	-
L38	J802	ØKC902	Red⁻	nt	S	nt	+
G12	J802	ØKC903	Red⁻	nt	S	nt	+

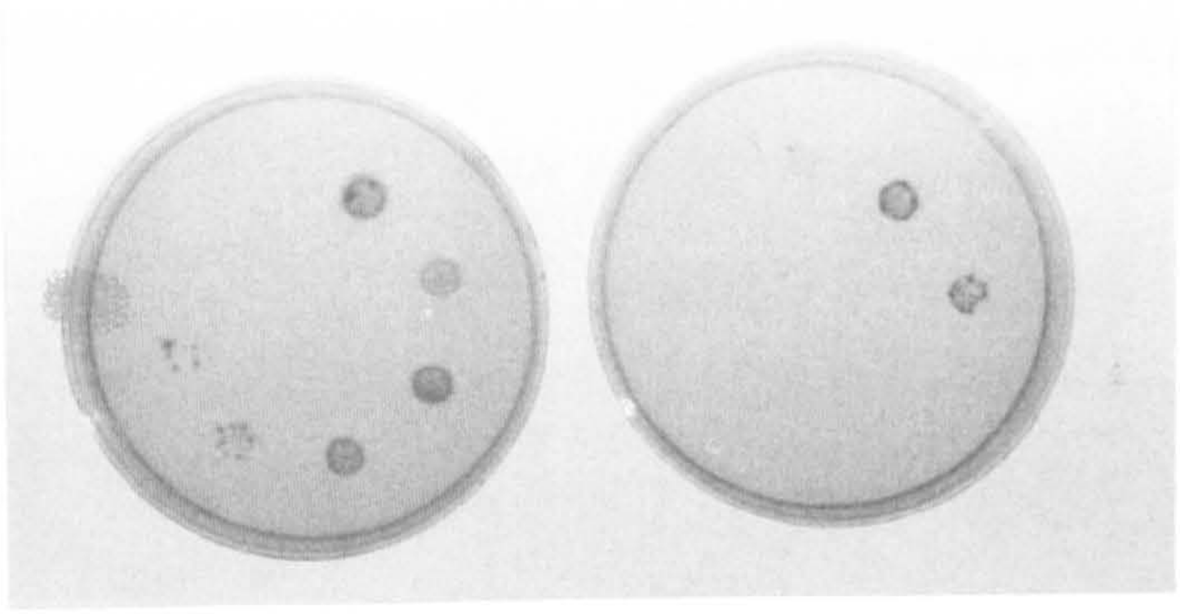
Figure 3.4

Photograph of J802 and US224 growth on Syton silica medium with proline as sole nitrogen source

The photograph shows J802 and US224 grown on Syton silica with proline as sole nitrogen source. The spots of growth represent 5µl innoculation with a dilution series strating at 10^{-2} at the top centre point of the plate and finishing at 10^{-7} for J802 and 10^{-3} for US224. The comparison of US224 on nitrogen as sole nitrogen source and proline as sole nitrogen source is the test to demonstrate a Put^- phenotype.

Figure 3.4

Photograph of J802 and US224 growth on Syton silica medium with proline as sole nitrogen source



3.3.2 Resistance to proline analogues

The second method to identify the Put$^-$ phenotype was to show resistance or sensitivity to the proline analogues AZC and DHP (2.3.3). When isolated, the Put$^-$ strains US218 and US224 showed resistance to AZC and DHP. The parent strain J802 was sensitive to both proline analogues. On testing J802, it was found to be resistant to the concentrations of AZC used, as shown in Figure 3.5, but still retained sensitivity to DHP. Few of the US218 and US224 lysogens tested were not resistant to DHP i.e. most were shown to be resistant to DHP and were therefore Put$^-$. Examples of Put$^+$ and Put$^-$ phenotypes as determined by resistance and sensitivity to DHP respectively are shown in Figures 3.5 and 3.6. A summary of results is given in Table 3.3.

3.3.3 Transport of ^{14}C proline

The third method used to examine the Put$^{-/+}$ phenotype was to assay directly for the transport of ^{14}C proline into the cell (2.3.3). In the process of assaying ^{14}C proline uptake, modifications were made to improve the assay. These included varying the amount of proline used, mixing ^{14}C proline and unlabelled proline to increase the concentration of proline from 2μM - 5μM. This also served to investigate if higher proline concentrations stimulated proline import via another, low affinity, transport system. The age of cells harvested for use in the transport assays was also varied, neither of these modifications had any effect on the results. Fewer lysogens were tested using the ^{14}C proline assay because of the length of time required to perform each assay. Proline transport of US218, US224 and J802 is shown in Figure 3.7. One of the lysogens which showed decreased proline transport compared to J802 was US218::∅KC902 A24 as is shown in Figure 3.8. This experiment

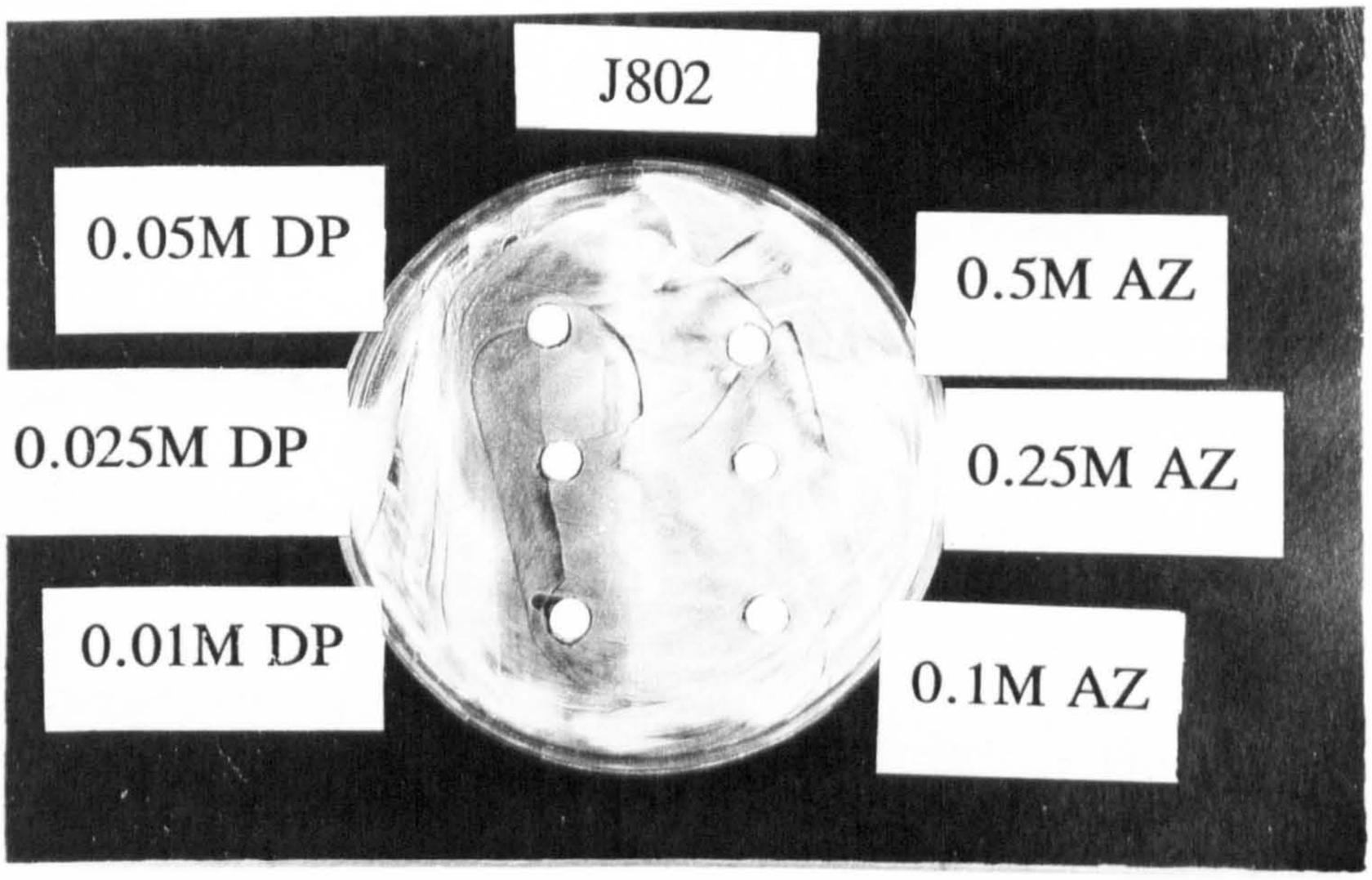
Figure 3.5

Resistance of US224 and J802 to proline analogues

Proline transport mutants Put⁻ were isolated by resistance to toxic proline analogues. Azetidine-2-carboxylate (AZC) and 3,4-dehydroproline (DHP) were the proline analogues used. US224 is a Put⁻ mutant of J802. US224 shows resistance to DHP whereas J802 is sensitive. J802 does not show sensitivity to AZC at the concentrations used here.

Figure 3.5

Resistance of US224 and J802 to proline analogues



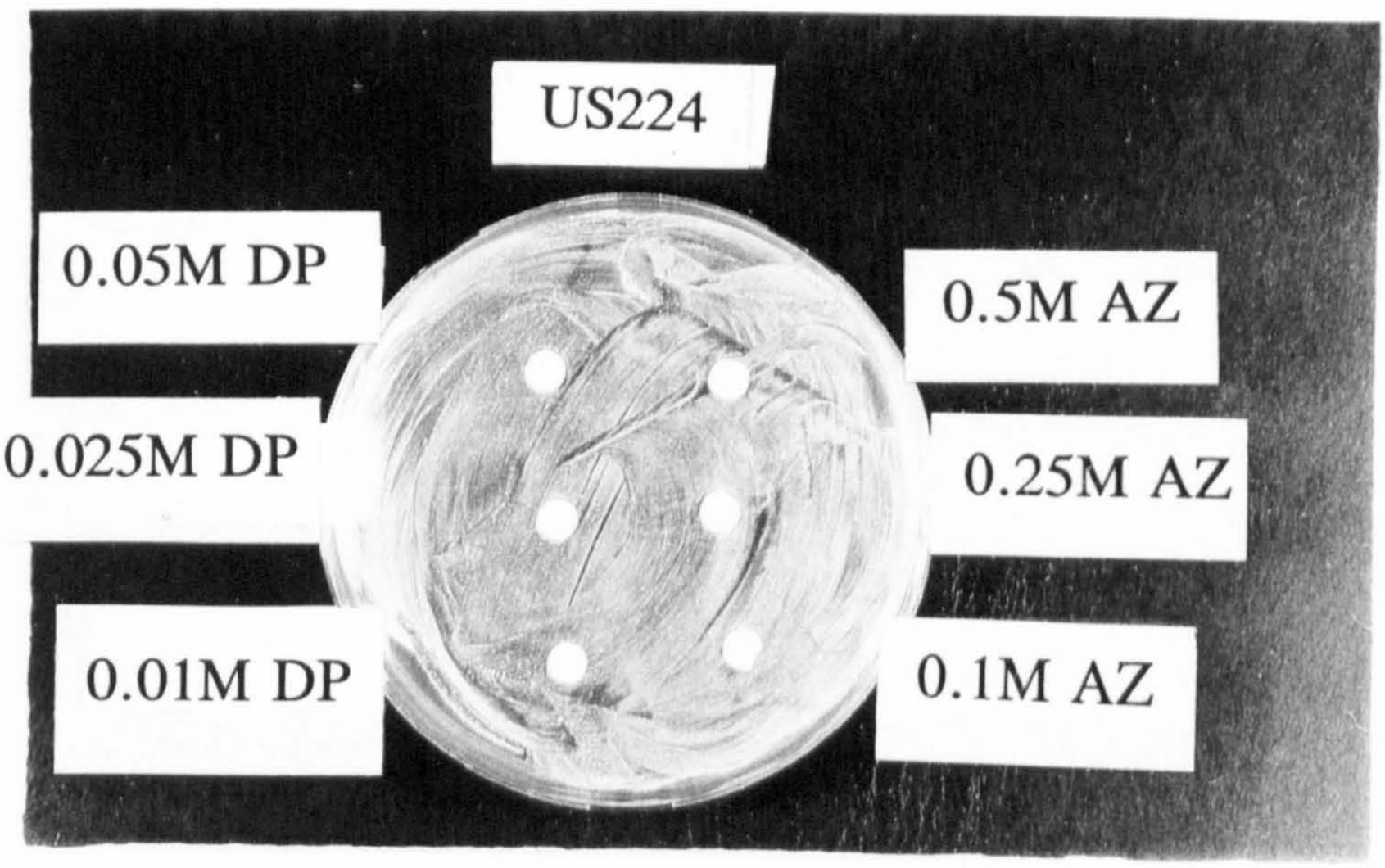


Figure 3.6

Resistance of US218 lysogens to the proline analogue DHP

Resistance to proline analogues is a Put⁻ phenotype. Resistance of US218 lysogens to proline analogues is shown on the facing page.

B22 is a lysogen of US218 with phage ∅KC903.

Figure 3.6

Resistance of US218 lysogens to the proline analogue DHP

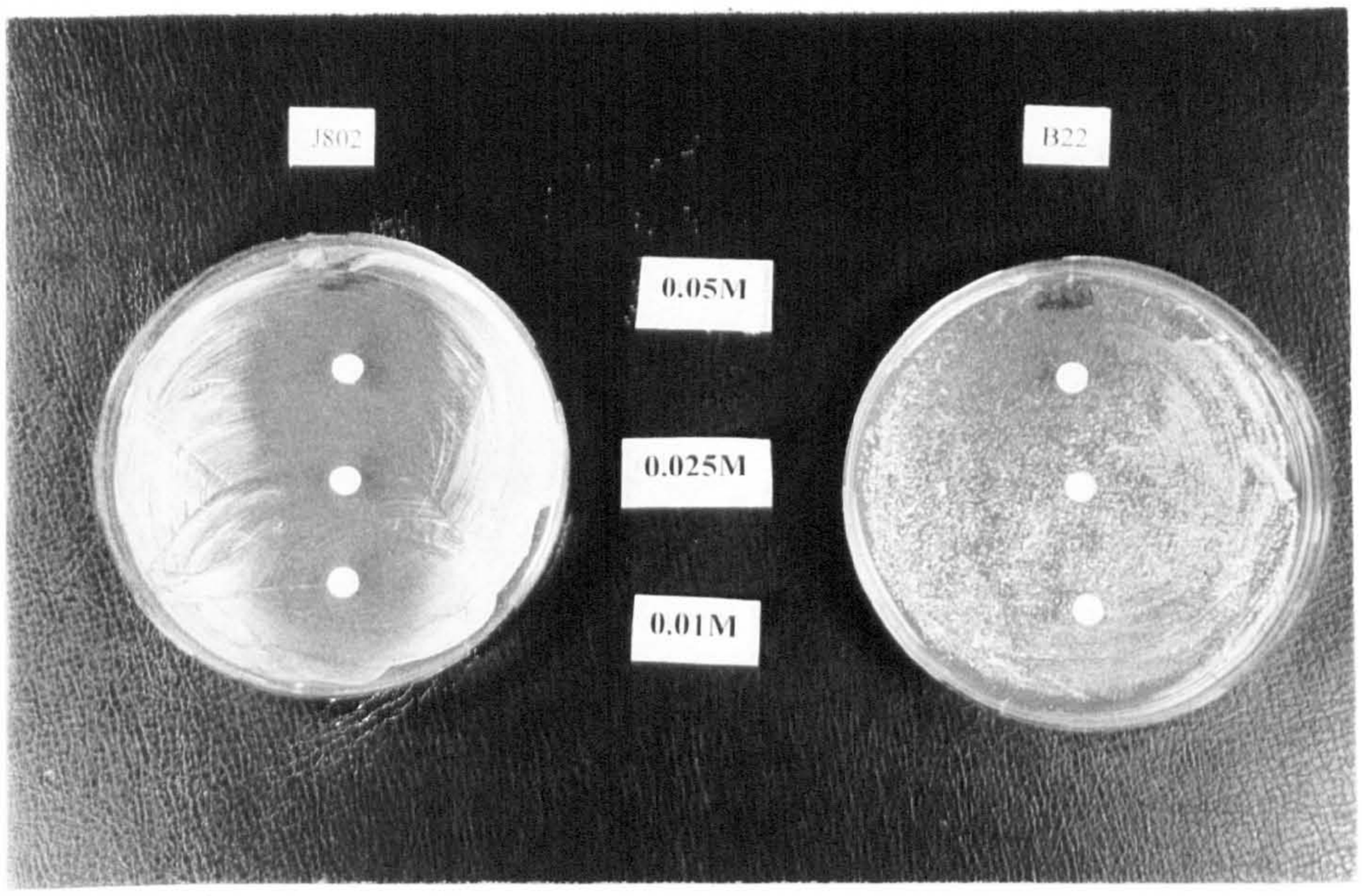


Figure 3.7

^{14}C Proline transport of US218, US224 and J802

Proline transport was assayed in cells directly by monitoring the amount of ^{14}C proline taken up by cells at timed intervals. ^{14}C proline concentration was monitored by a liquid scintillation counter. The amount of proline in ρ moles was calculated using the equation presented in the Appendix.

Proline transport by US218 and US224, both Put$^-$ mutants, is much reduced by comparison with the progenitor strain J802.

Figure 3.7

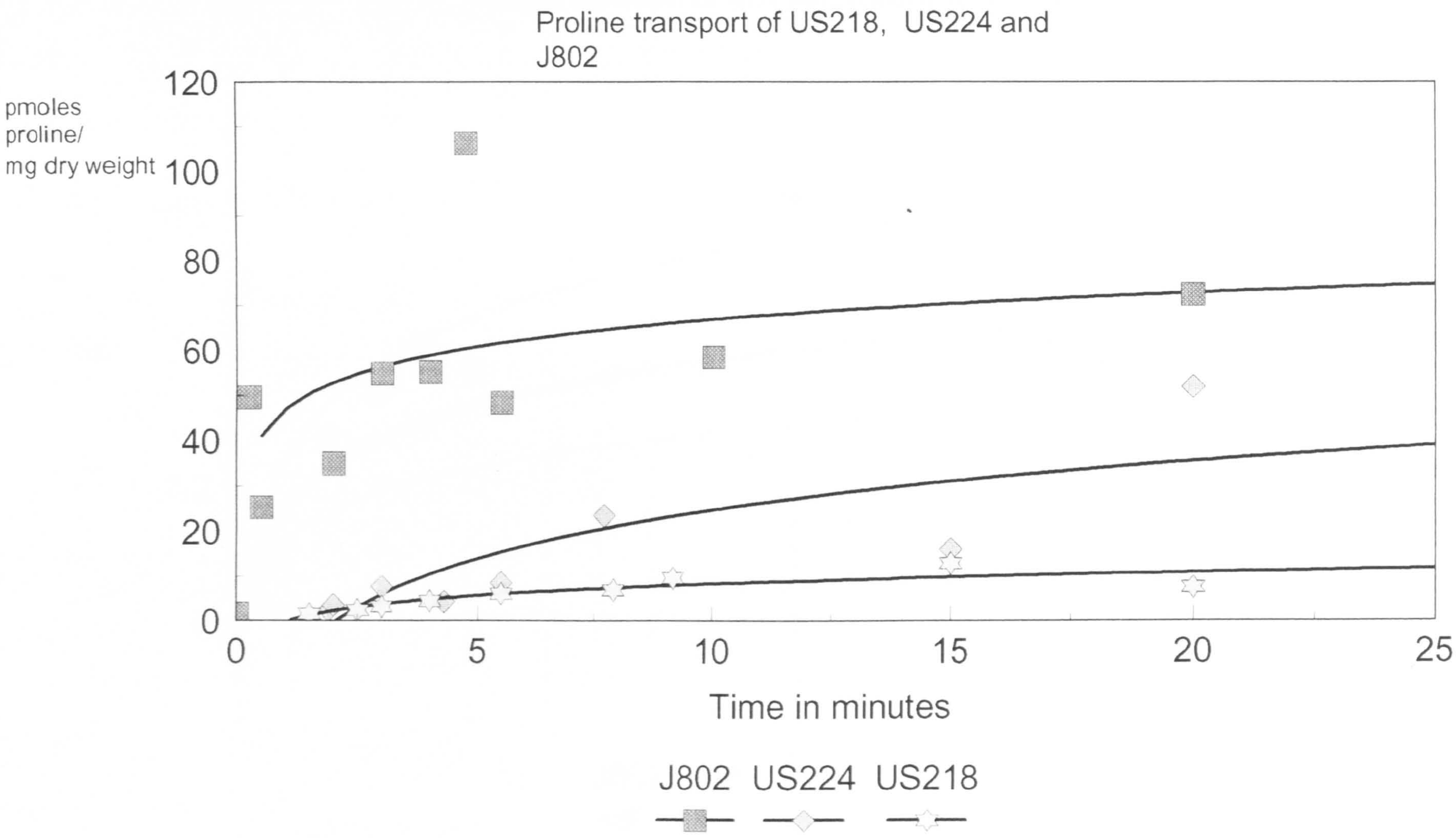

Proline transport of US218, US224 and J802
pmoles proline/ mg dry weight
120
100
80
60
40
20
0
0
5
10
15
20
25
Time in minutes
J802
US224
US218

Figure 3.8

^{14}C Proline transport of US218 lysogens and J802.

Lysogens of the Put$^-$ strain US218 show reduced proline transport by comparison to the parent strain J802. In particular A24 and A47, US218 lysogens of ∅K902 and ∅K903 respectively, both Red$^-$, show the Put$^-$ phenotype of reduced proline transport. This demonstrates that a Red$^-$ Put$^-$ lysogen is viable. The lysogens tested here are of :-

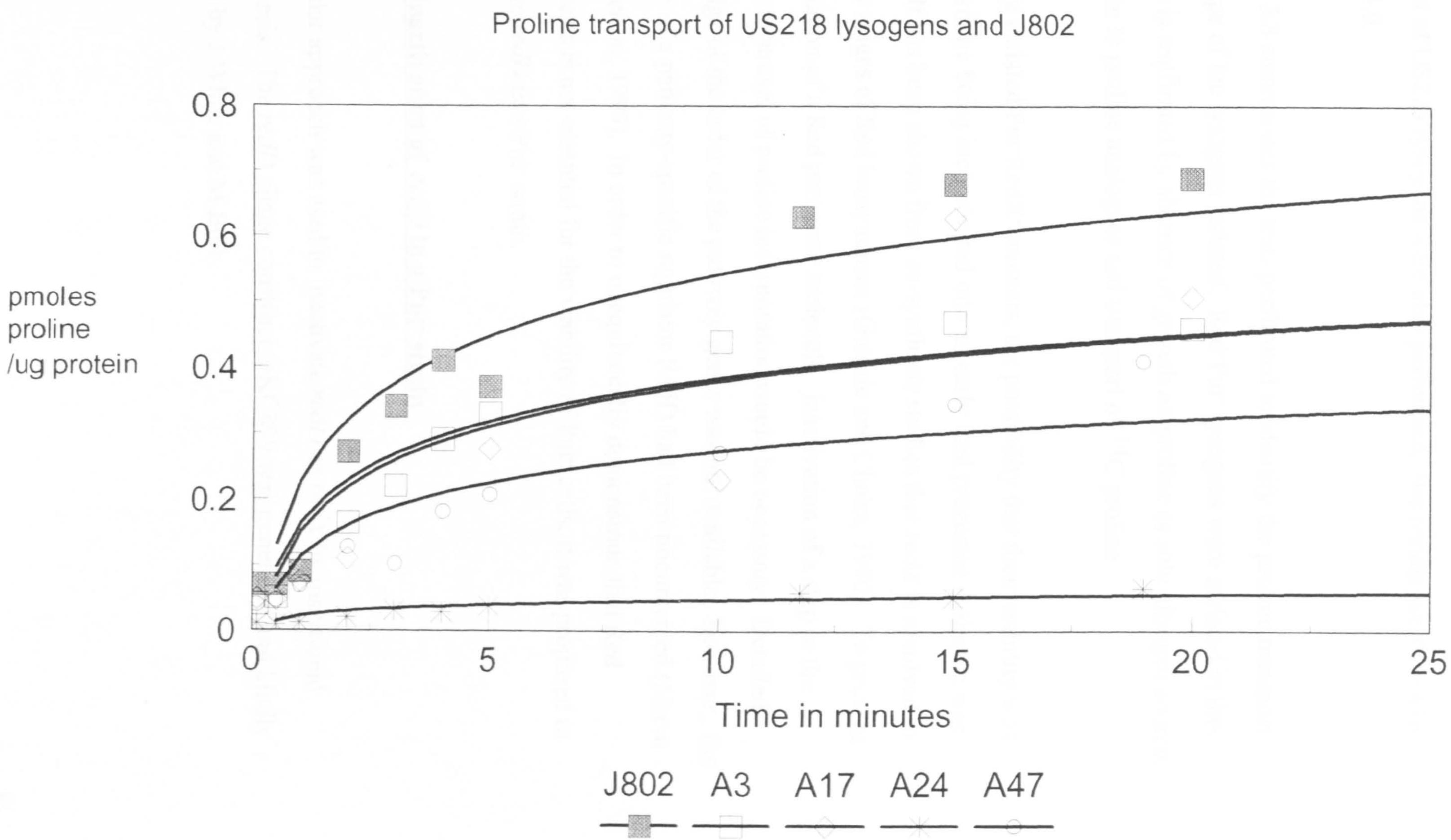
A3 US218::∅KC900

A17 US218::∅KC901

A24 US218::∅KC902

A47 US218::∅KC903

Figure 3.8


Proline transport of US218 lysogens and J802
pmoles proline /ug protein
0.8
0.6
0.4
0.2
0
0
5
10
15
20
25
Time in minutes
J802
A3
A17
A24
A47

was repeated three times and a similar result obtained each time. Proline transport of US224 lysogens were also performed, the results are shown in Figure 3.9.

Table 3.3 summarises the tests performed to identify the proline transport phenotype of the lysogens isolated. Red$^-$Put$^-$ lysogens were isolated in this study as is confirmed by absence of growth on proline as sole nitrogen source, resistance to proline analogues and transport of ^{14}C proline.

Having isolated Put$^-$RedX$^-$ mutants, the possibility that their viability was due to proline being incorporated into an early Red precursor molecule was raised. It has been shown from co-synthesis studies that *redX* is involved in the early stages of Red biosynthesis (Guthrie and Chater, 1990). To prevent the formation of a Red precursor molecule, inactivation of a step in the pathway upstream of proline incorporation would be necessary. Detailed knowledge of the order of the pathway genes was not available, however, the presence of a pathway-specific regulator RedD had been documented (Narva and Feitelson, 1990). In order to unequivocally demonstrate that Red biosynthesis is not essential for the viability of Put$^-$ cells, it was proposed to inactivate *redD* in a Put$^-$ strain.

3.3.4 Inactivation of *redD* in a Put$^-$ strain

A similar approach was used to inactivate *redD* as *redX* by insertional mutagenesis. The *redD* phage construct ØKC899 was generated and kindly supplied by J.White and M.Bibb.

Figure 3.9

^{14}C Proline transport of US224 and US218 lysogens and J802

Lysogens of the Put⁻ strain US224 show reduced transport by comparison to the progenitor strain J802. The lysogens shown facing are of:-

Q5 US224::∅KC900 (Act⁻)

C34 US224::∅KC901 (Act⁻)

G43 US224::∅KC902 (Red⁻)

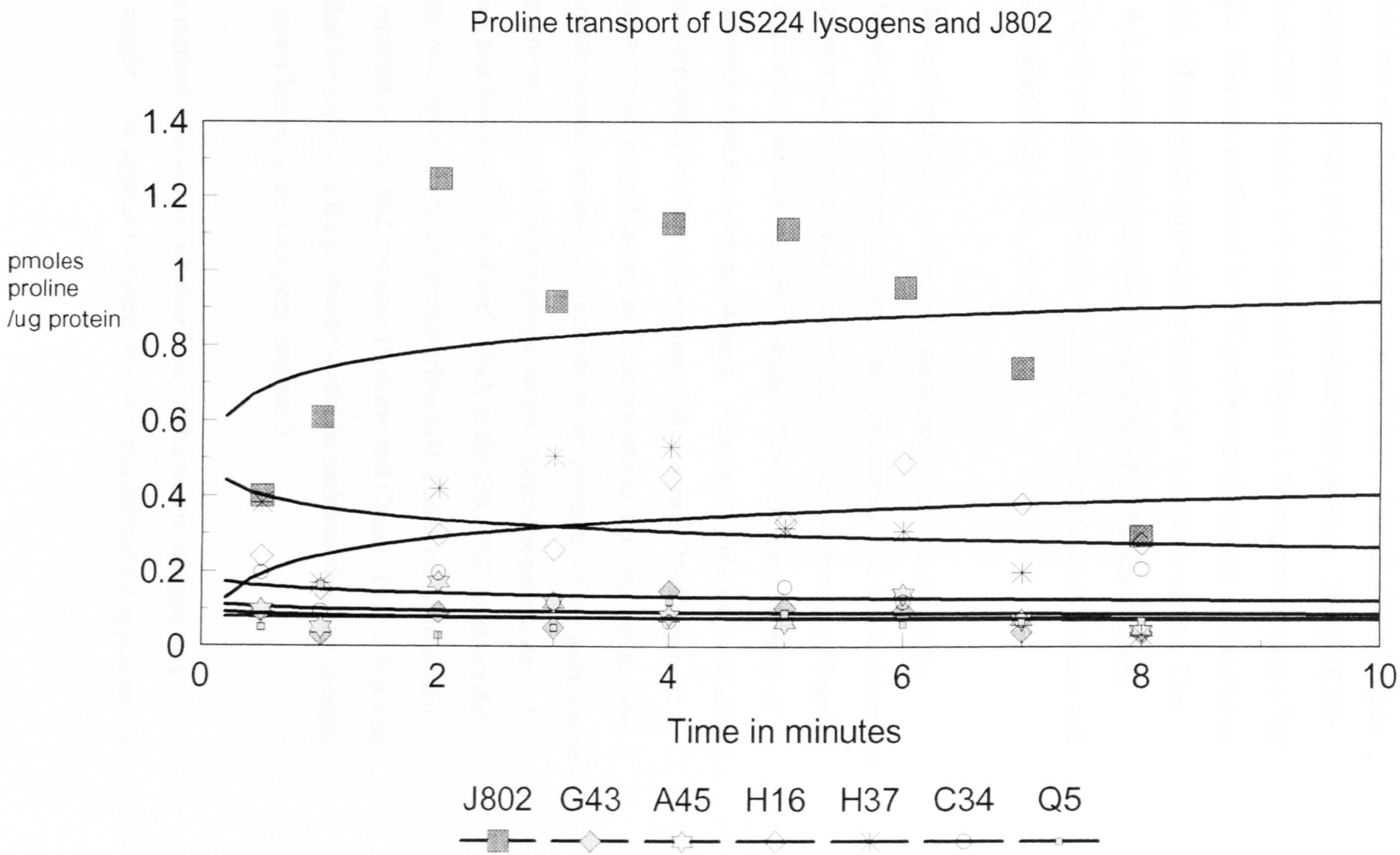
H16 US224::∅KC902 (Red⁻)

H37 US224::∅KC903 (Red⁻)

A45 US218::∅KC903 (Red⁻)

Of special interest is the reduced proline transport as demonstrated by the Red⁻ lysogens G43, H16, H37 and A45. This shows that Red⁻Put⁻ lysogens are viable.

Figure 3.9


Proline transport of US224 lysogens and J802
pmoles proline /ug protein
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
2
4
6
8
10
Time in minutes
J802
G43
A45
H16
H37
C34
Q5

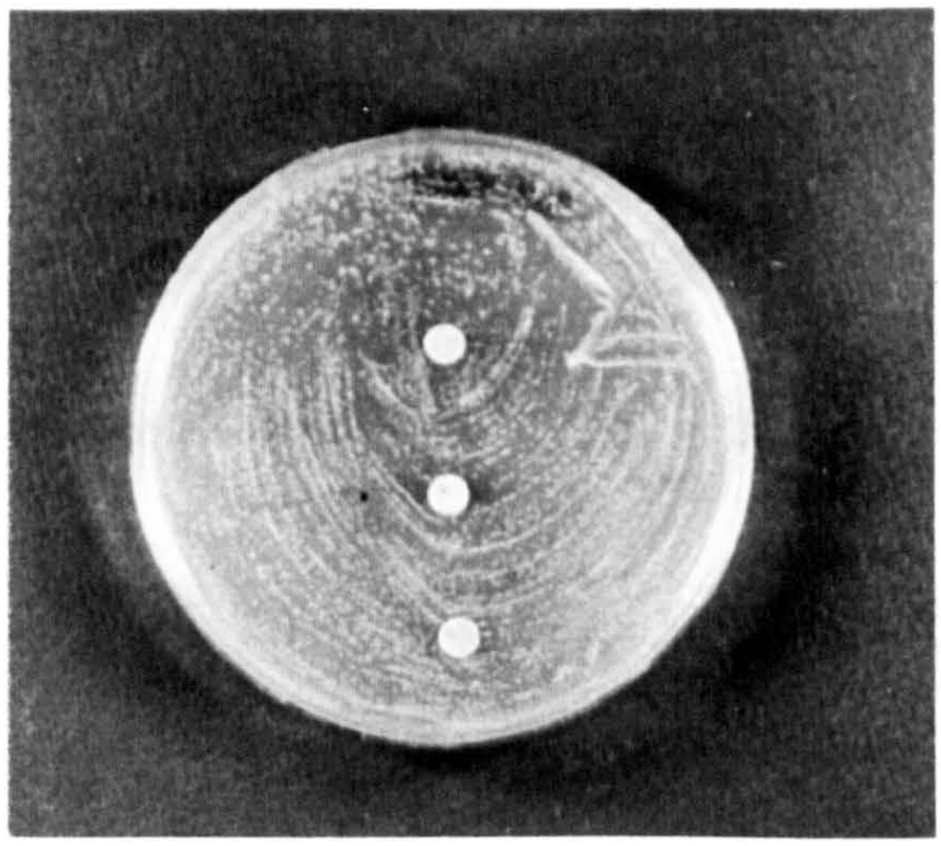
The strain US224 was lysogenised using ØKC899 (2.4.5). One lysogen was isolated and was tested to show that it was still Put$^-$. The ØKC899 lysogen was thiostrepton resistant and did not produce Red pigment. The ØKC899 lysogen was DHP resistant as is shown in Figure 3.10 suggesting it has a Put$^-$ phenotype. This was confirmed by ^{14}C proline uptake assays as is shown in Figure 3.11. These results show that a RedD$^-$Put$^-$ lysogen is viable. This suggests that Red biosynthesis is not essential in a Put$^-$ background and disproves the hypothesis that Red is essential in a Put$^-$ mutant to act as a sink for excess proline (Hood *et al.*, 1992).

Using the *redX* and *redD xylE* fusions made previously, it should be possible to assay the promoter activity of the *red* genes. This would determine if the genes were activated to a greater extent or if they were activated earlier in a Put$^-$ mutant as compared to a Put$^+$ strain. However, the reliability of *xylE* as a reporter gene has been questioned. Activity of *xylE* has been shown in constructs irrespective of the orientation of the internal DNA fragment inserted upstream of the *xylE* gene. In these constructs only one orientation should transcriptionally couple *xylE* activity to the promoter of the gene whose internal fragment is cloned into the phage vector. Recent sequence data of *redX* shows that the orientation of *redX* DNA in the ØKC902 construct did not fuse the *redX* promoter to the promoterless *xylE* gene, yet XylE activity has been reported in ØKC902 lysogens (Guthrie and Chater, 1990). It is also possible that the activity of the promoter would be undetected, if the strength of the promoter linked to the *xylE* gene was weak.

Another method of assessing whether Red production is initiated was therefore sought. The approach chosen involved monitoring the appearance of

Figure 3.10

Proline analogue resistance of *redD* US224 lysogen.



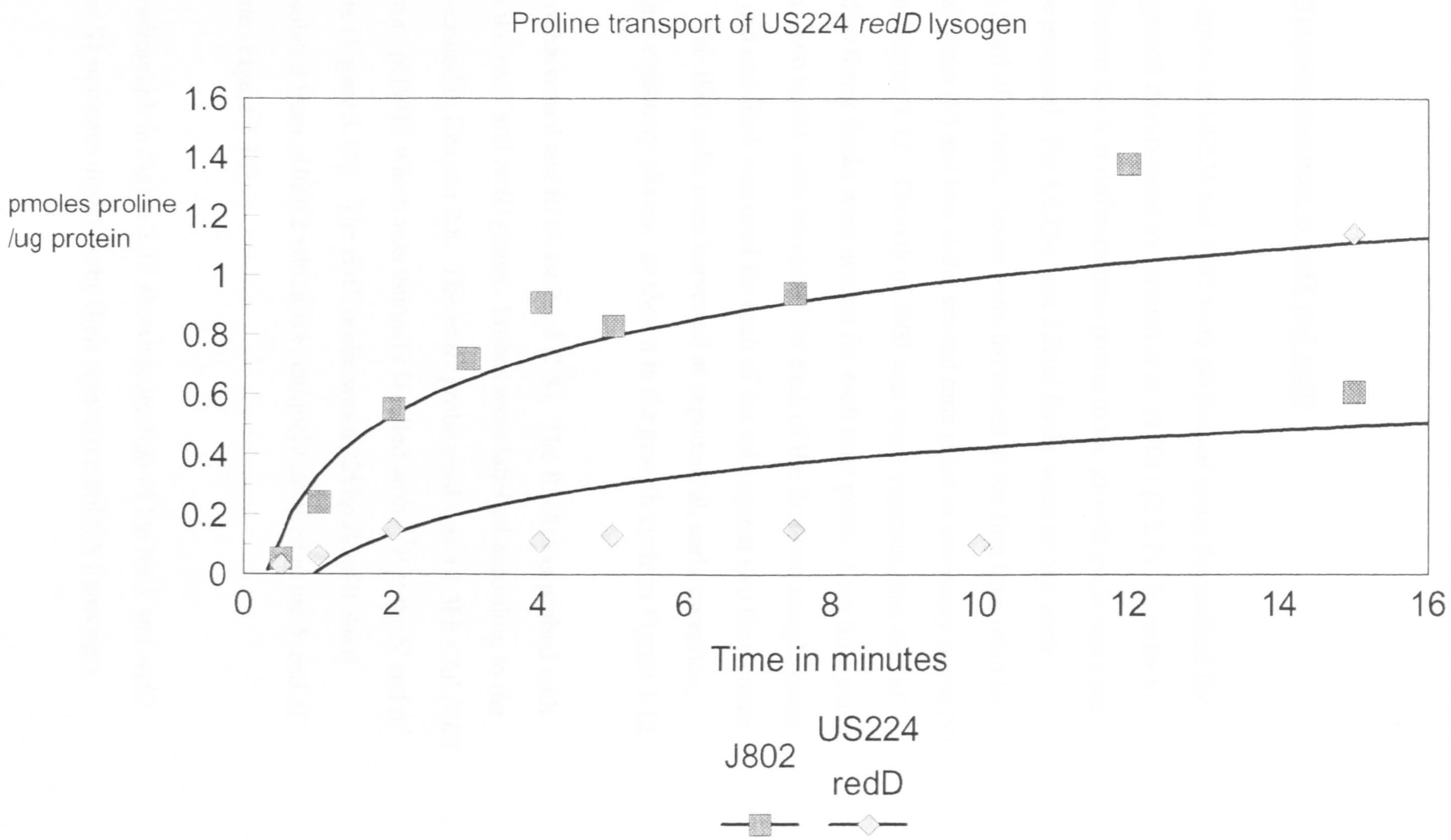
A ∅KC899 lysogen of US224 was isolated. ∅KC899 is a ∅C31 altered phage containing DNA internal to *redD*. The resistance of the *redD* lysogen to the proline analogue 3,4-dehydroproline (DHP) is shown on the facing page. Resistance to the proline analogue DHP is a Put$^-$ phenotype.

Figure 3.11

^{14}C proline transport of a US224 *redD* lysogen.

A US224 lysogen of ∅KC899, a phage containing DNA internal to *redD*, was isolated. Proline transport of the US224::∅KC899 lysogen was assayed and the results are shown on the facing page. The US224::∅KC899 lysogen showed reduced proline transport compared to the progenitor J802. This is the phenotype expected of a Put^- mutant. It demonstrates that a $RedD^-$ Put^- mutant is viable.

Figure 3.11


Proline transport of US224 redD lysogen
pmoles proline /ug protein
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
2
4
6
8
10
12
14
16
Time in minutes
J802
US224 redD

red gene mRNA transcripts using radiolabelled DNA probes after S1 nuclease treatment.

3.3.5 S1 Nuclease mapping of *redX* and *redD*

Growth curves of US224 and J802 were performed using the method for dispersed growth documented by Strauch *et al.* (1991) (2.2.2). In order to isolate sufficient RNA at different time points in the growth cycle replicate flasks were prepared. For US224 nine 250ml flasks were set up, each containing 50ml of culture. Seven were harvested at the first time point in exponential phase (E) and two at the second time point in stationary phase (F), as shown in Figure 3.12. Growth of J802 was more vigorous than that of US224 and so fewer flasks were needed for each time point. Four time points were taken; two flasks were harvested for each of the first two sample times (A and B) and one flask was used for each of the subsequent two time points (C and D). The J802 cells were harvested at exponential, early transition, transition and stationary phases, as shown in the growth cycle in Figure 3.12.

Cells were harvested and RNA isolated (2.8). The RNA was probed with DNA from the *redX* and *redD* genes. Probes were labelled according to the method described in Chapter 2.8. The *redD* probe used was a 1.3kb *Cla*I *Nde*I fragment from pIJ6013 which was uniquely labelled with ^{32}P at the 5' end of the *Cla*I site (Figure 3.13). The *redX* probe was a 229bp *Bam*HI *Sma*I fragment isolated from pIJ6012 which was uniquely labelled at the 5' end of the *Sma*I site (Figure 3.14).

The autoradiograph in Figure 3.15 shows protection of the *redX* and *redD* probes from S1 nuclease digestion by their respective mRNA transcripts.

Figure 3.12

Growth curve of J802 and US224 from which RNA was isolated.

Growth of US224 and J802 in SMM medium using the method as Strauch *et al.* (1991) is shown facing. The time points labelled A-F are those at which cells were harvested. Two flasks were harvested at each time points A and B for J802, subsequently one flask was harvested at each time point (C and D). For US224 nine flasks of cultures were grown, seven were harvested at time point E the remaining two were harvested at time point F.

Figure 3.12

Growth curve of J802 and US224 from which RNA was isolated.

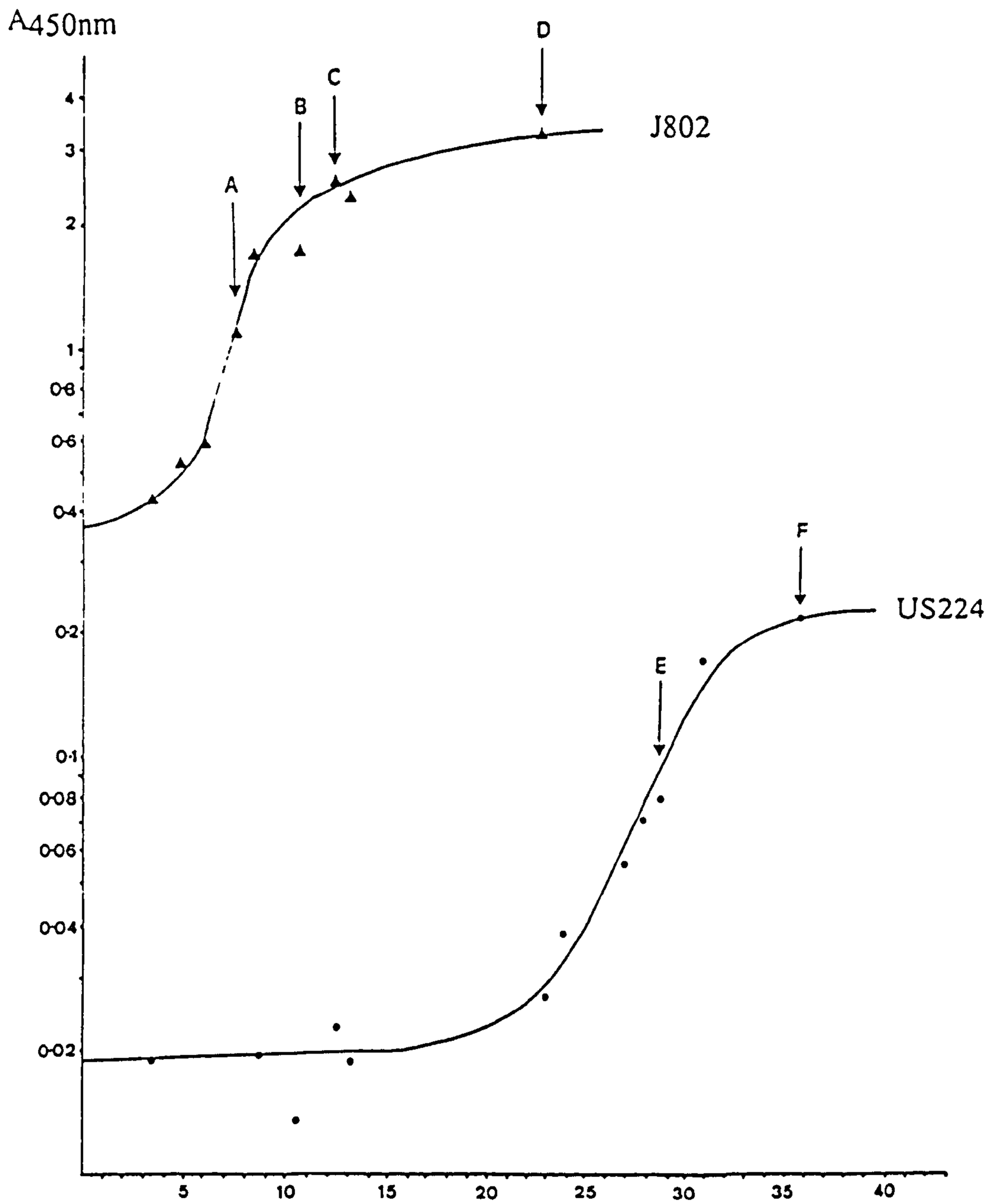


Figure 3.13

Representation of the *redD* DNA used as a probe in S1 nuclease analysis.

The *redD* DNA used as a probe was isolated from pIJ6013. The plasmid pIJ6013 was created by cloning a 2.2kb *Bam*HI *Pst*I fragment from pIJ2341 into pIJ2926 (J.White personal communication). The *redD* probe was made by uniquely labelling the 1.3kb *Cla*I *Nde*I fragment from pIJ6013, at the 5' end of the *Cla*I site with ^{32}P. The S1 nuclease protected fragment expected is 229bp.

Figure 3.13

Representation of the *redD* DNA used as a probe in S1 nuclease analysis

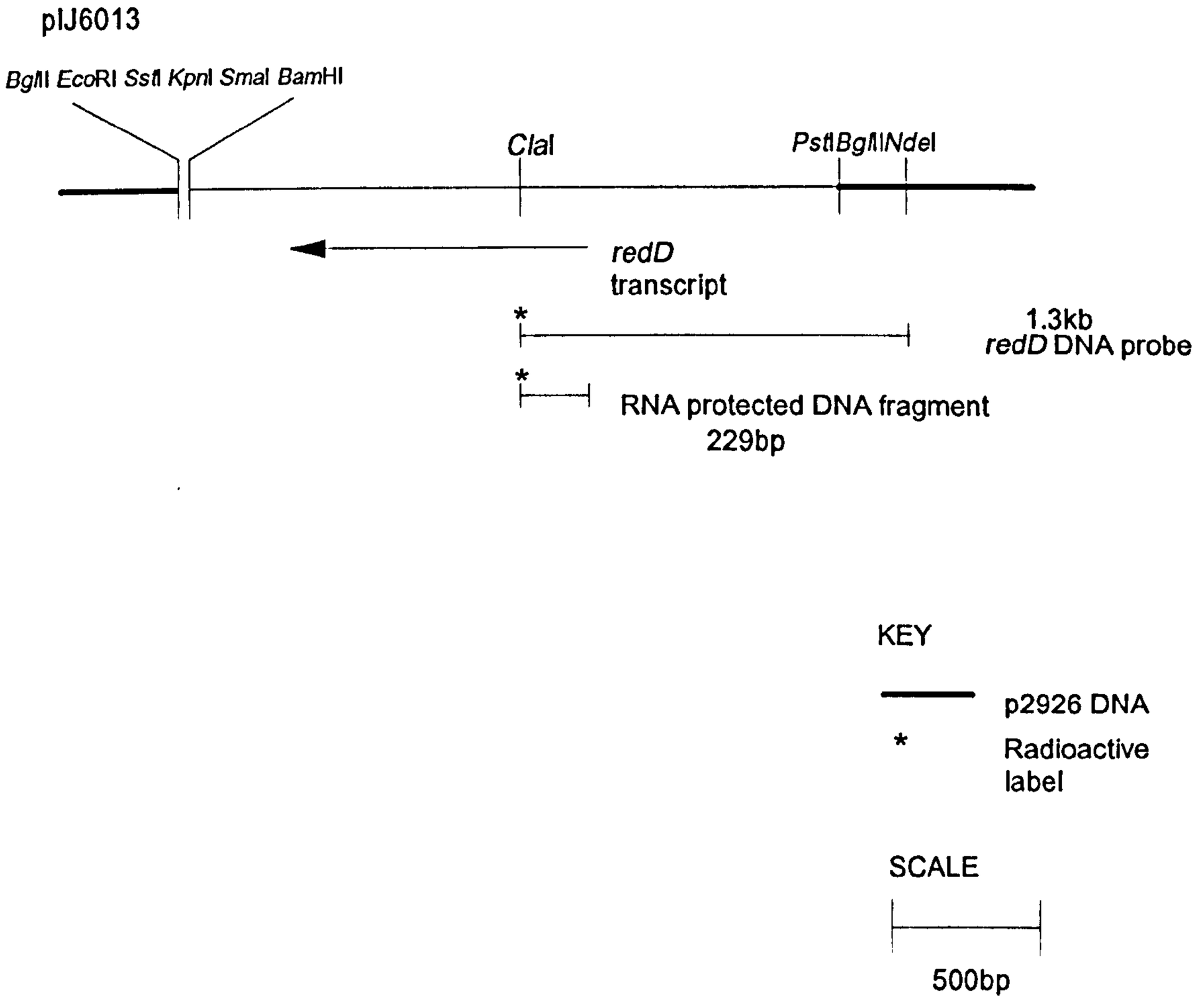


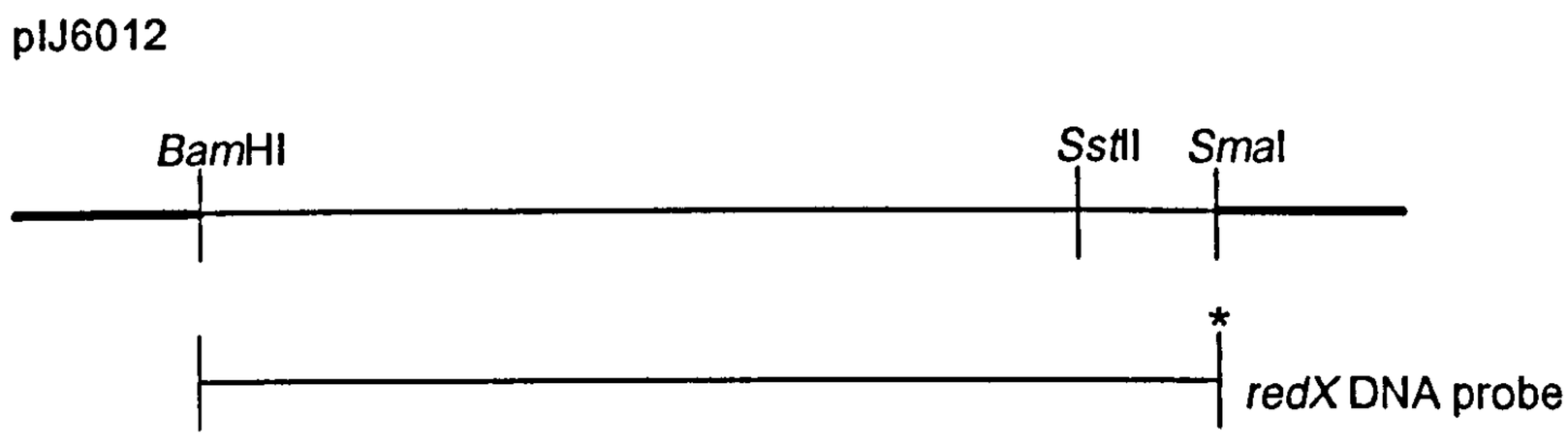
Figure 3.14

Representation of the *redX* DNA used as a probe in S1 nuclease analysis.

The *redX* DNA used as a probe was prepared from pIJ6012. The plasmid pIJ6012 was constructed by cloning a 220bp *Sma*I *Bam*HI fragment from pIJ6000 into pIJ2926 (J.White personal communication). The *redX* probe was made by uniquely labelling the 220bp *Bam*HI *Sma*I fragment from pIJ6012 with ^{32}P at the 5' end of the *Sma*I site. The S1 nuclease protected fragment expected is 220bp.

Figure 3.14

Representation of *redX* DNA used as a probe in S1 nuclease analysis



Direction of transcription of *redX*
(believed direction of *redX* transcription prior to sequence data being available)

KEY

Vector DNA

* Radioactive label

SCALE

100bp

Figure 3.15

Autoradiograph of S1 nuclease protected probes

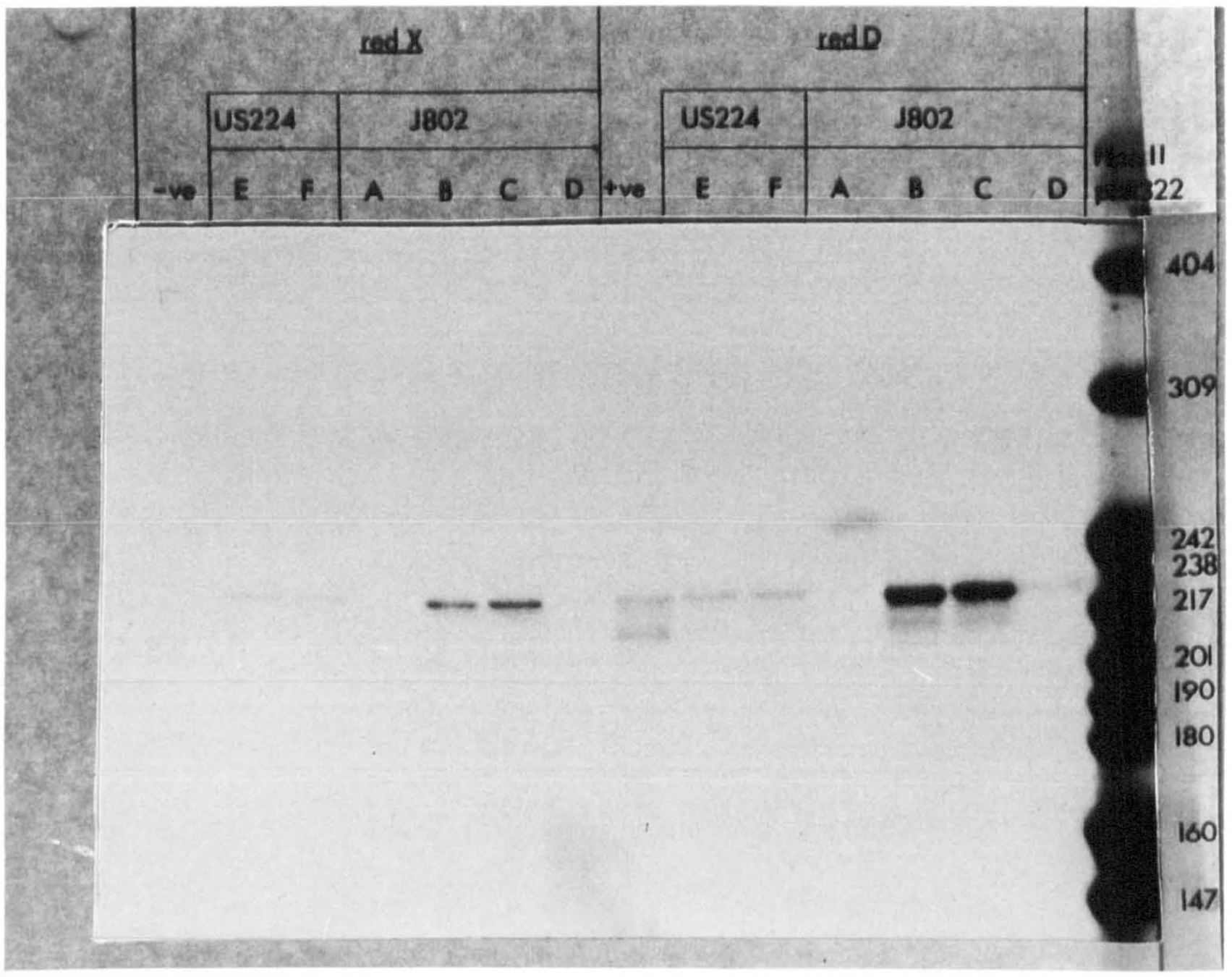
Radiolabelled DNA probes of *redD* and *redX* were hybridised with the RNA isolated from cells sampled at time points A-F (Figure 3.11). RNA transcripts of *redD* and *redX* hybridised to the DNA probes, forming a double stranded structure that is protected from S1 nuclease digestion. The protected DNA probes were visualised by autoradiography following acrylamide gel electrophoresis of S1 nuclease treated samples.

As can be seen on the facing page a fragment of approximately 223bp of the *redD* probe was protected from S1 nuclease digestion. The presence of protected *redD* probe DNA indicates the presence of *redD* transcripts. RedD is a positive regulator of Red biosynthesis. From the work of Takano *et al.* (1992) it has been established that the appearance of *redD* transcripts precedes the appearance of *redX* transcripts by approximately an hour.

A fragment of 220bp of the *redX* DNA probe was protected from S1 nuclease digestion. In J802 the appearance of *redX* and *redD* transcripts at the onset of stationary phase is in agreement with the findings of Takano *et al.* (1992)

Figure 3.15

Autoradiograph of S1 nuclease protected probes



For *redX* a protected fragment of 220bp was seen. A *redD* DNA fragment of 223bp was protected. The results obtained were in agreement with those obtained by Takano *et al.* (1992), who showed that *redD* transcripts appeared during transition into stationary phase, followed by the appearance of *redX* transcripts on entry into stationary phase. In these experiments J802 was used as compared to M145 in the experiments of Takano *et al.* (1992). Although the actual time taken to enter transition and stationary phase was longer for J802 than M145, the pattern of appearance of the *redD* and *redX* transcripts remained similar. In US224 *redD* and *redX* transcripts were detected in exponential phase. This would suggest from the results of Takano *et al.* (1992), that the *redD* transcripts would have been present even earlier in the growth phase because the appearance of *redX* transcripts follows by a lag of at least an hour the appearance of *redD* transcripts.

A sample of the cells were harvested at time points E and F for US224 and time point C for J802 as shown on Figure 3.12, in order to perform ^{14}C proline uptake assays, to show that US224 cells used in this experiment were Put^-. The proline uptake assay showed that US224 did not transport proline, as can be seen in Figure 3.16.

These results suggest that transcription of *red* genes occurs earlier in the Put^- strain US224 than in the progenitor Put^+ strain J802.

Figure 3.16

^{14}C proline transport of cells from which RNA was isolated for S1 analysis

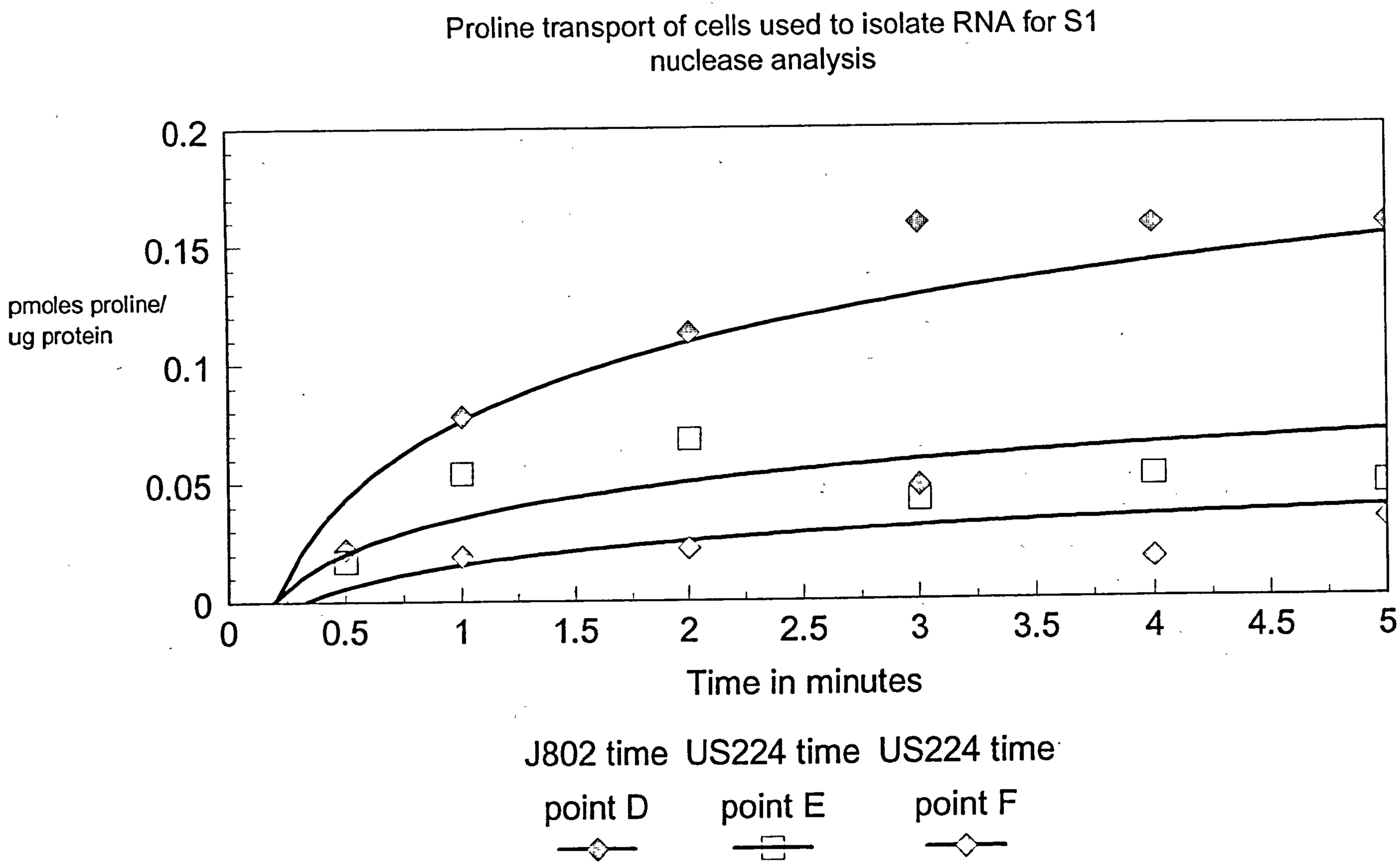
^{14}C proline uptake was assayed in cells grown from a sample taken at time points D, E and F of J802 and US224 (Figure 3.11). Proline transport was reduced in US224 as compared to J802. This shows the RNA isolated was from a Put$^-$ mutant. Points D, E and F correspond to samples taken at time points in the growth curve of US224 and J802 (Figure 3.11).

Point D J802 stationary phase

Point E US224 mid-log

Point F US224 stationary phase

Figure 3.16


Proline transport of cells used to isolate RNA for S1 nuclease analysis
pmoles proline/ ug protein
0.2
0.15
0.1
0.05
0
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
Time in minutes
J802 time point D
US224 time point E
US224 time point F

3.4 Discussion

This study has shown that inactivation of Red biosynthesis in a *put44* mutant is not a lethal event. Previous data had implied that inactivation of Red was a lethal event (Hood *et al.*, 1992). The Put mutant allele studied, *put-44*, was the weakest of the Put⁻ alleles isolated. The *put-44* mutant PTM44f only marginally over-produced Red with respect to J802 the progenitor strain. The activity of the proline catabolic enzymes proline oxidase and P5C dehydrogenase are only reduced in PTM44f. In other Put⁻ mutants, such as PUM7r, the activities of proline oxidase and P5C dehydrogenase were undetectable. There is a correlation between the Put mutants which had the least activities of the proline catabolic enzymes and those which showed the greatest production of Red. The strain PUM7r produced the most Red of all the Put⁻ mutants and could be described as the most extreme Put mutant. Attempts to cross PUM7r with J1506 to isolate a *put-7* allele in a more amenable strain were unsuccessful. Lysogeny of PUM7r directly with ØKC900, ØKC901, ØKC902 and ØKC903 was attempted but no thiostrepton resistant lysogens were isolated. Further inactivation of Red biosynthesis in other Put⁻ mutants carrying other mutant alleles is needed before it can be concluded that loss of Red production and proline catabolism is not possible in the same cell.

S1 nuclease mapping experiments of *redD* and *redX* genes has shown that Red is produced earlier in the *put-44* mutant US224 than the progenitor strain J802. It could be argued that due to the slow growth rate of US224 antibiotic production is stimulated. Work of Takano *et al.* (1992) showed that increased ppGpp due to nutritional shift down caused a reduction in growth rate and stimulated precocious antibiotic production. It has been proposed that

reduction in growth rate, as seen on entry into stationary phase in liquid cultures is a factor in inducing antibiotic production. Unlike Red biosynthesis, Act production was not stimulated in the Put⁻ mutants. In fact, Act production was reduced in some of the Put⁻ mutants. It could be argued that biosynthesis of the antibiotic Red to reduce proline levels would cause a metabolic load on the cells thus causing a slower growth rate.

Since S1 nuclease mapping experiments were performed the *redX* gene has been sequenced. From the sequence data the direction of transcription appears to be in the opposite direction to that previously believed. The implications on this work are that the *redX* DNA probe was designed to protect a transcript running in the opposite direction to that predicted from sequence data. In view of this, the protection of the *redX* DNA probe seen in Figure 3.15 would not be expected. However, protection of the labelled DNA did occur.

Transcripts of *redX* were also seen in the work of Takano *et al.* (1992) based on the same *redX* transcript orientation as this work. The protection of the *redX* DNA probe must be occurring by some reproducible phenomenon. One possible explanation could be that protection of the labelled DNA occurred by hybridisation with the true *redX* transcript and homologus DNA with the DNA probe. The reason for believing that the DNA probe protection seen was a reflection of true *redX* transcription is that the appearance of *redX* transcripts correlated with the appearance of *redD* transcripts. The *redX* transcripts were seen an hour after the appearance of *redD* transcripts. This is the pattern of transcription that would correlate with the proposal that *redD* is a positive regulatory element which activates transcription of the Red structural genes including *redX*.

CHAPTER 4

DNA SEQUENCE AND DNA SEQUENCE ANALYSIS OF THE Pwb REGION

4.1 Introduction

Red biosynthesis is observed in *bldA* mutants grown on media with a low concentration of less than 0.5mM phosphate (Guthrie and Chater, 1990). Restoration of production of the other antibiotics; methlylenomycin, actinorhodin and the calcium dependent antibiotic, is not seen under these conditions. Restoration of Act biosynthesis in a *bldA* mutant has been achieved by altering a TTA codon to a TTG codon within the *actII*ORF4 gene.

The *bldA*-dependence of Red biosynthesis is less well understood. The pathway-specific, positive regulator encoded by *redD*, although homologous to actII-ORF4, does not contain a TTA codon. The ability to induce Red biosynthesis in low phosphate conditions in *bldA* mutants, suggests that the Red structural genes also do not contain a TTA codon (Guthrie and Chater, 1990). This leads to the hypothesis that there is another regulatory gene that contains a TTA codon which is responsible for Red activation in high phosphate conditions. To investigate this link between *bldA* and Red production, mutants of *bldA* strains that had restored Red biosynthesis in high phosphate conditions were isolated.

4.1.1 Isolation of Pwb mutants

Mutants of J1700 (*bldA39*) which had acquired the ability to produce Red were isolated by E.Guthrie (1990). During screening for pigmented antibiotic production Red-producing, but no actinorhodin producing mutants were isolated. The mutant phenotype was called Pwb, pigmented whilst bald. Three mutants were studied in greater detail and these are listed in Table 4.1. Mapping of these mutations was performed, but was inconclusive. A library

Table 4.1

Isolation and original cloning of Pwb mutant DNA

Mutant	Mutagen	Strain	Allele	Plasmid	M13 clones
EG6	UV	EG311	*pwb-6*	pIJ2530	ØCF1 ØCF2
EG9	UV	EG111	*pwb-9*	pIJ2520	ØCF3 ØCF4
EG16	NTG	EG441	*pwb-16*	pIJ2540	ØCF5 ØCF6
Wild-type	-	-	*pwb*$^+$	pCNB1003	ØCF7 ØCF8

Pwb mutants were isolated by E.Guthrie (Guthrie and Chater, 1990). Three mutants Pwb-6, Pwb-9 and Pwb-16 were the subject of further study. Plasmids pIJ2520, pIJ2530 and pIJ2540 which conferred the Pwb phenotype were constructed by E.Guthrie. The wild-type *red* cluster was made available by F.Malpartida (Malpartida *et al.*, 1990). M13 clones containing the 800bp *Eco*RI fragment in both orientations were contructed as described in this study.

of total DNA of the three Pwb mutants generated using a low copy-number, transmissible streptomycete plasmid vector (pIJ698) was constructed. The DNA was introduced into J1700 by plate mating and thiostrepton resistant, pigmented colonies isolated. A number of Red transconjugants were isolated from each mutant. The production of Red by the transconjugant implies that the Pwb phenotype is dominant to the chromosomal copy of the gene. Plasmid DNA was isolated from one transconjugant of each Pwb mutant. The three plasmids were transformed back into J1700 to show that the Pwb phenotype was associated with the plasmids. The plasmids were named pIJ2520, pIJ2530 and pIJ2540 (Table 4.1). From DNA restriction analysis it was shown that the three plasmids contained several fragments in common. Notably three *Bam*HI fragments were present in all three plasmid inserts. The *Bam*HI fragments were of a similar size to those located to the right hand end of the *red* cluster (Figure 4.1). Southern blotting analysis gave evidence that the plasmid inserts did indeed contain *red* DNA. Partial *Sau*3AI digests of the plasmid inserts were sub-cloned into the ØC31 derivative ØKC861. Red production by *bldA* strains lysogenised with insert containing phages was recorded. An 800bp *Eco*RI fragment was common to all the Pwb inducing phages suggesting that the Pwb mutations lay within or close to this DNA fragment (E.Guthrie unpublished results).

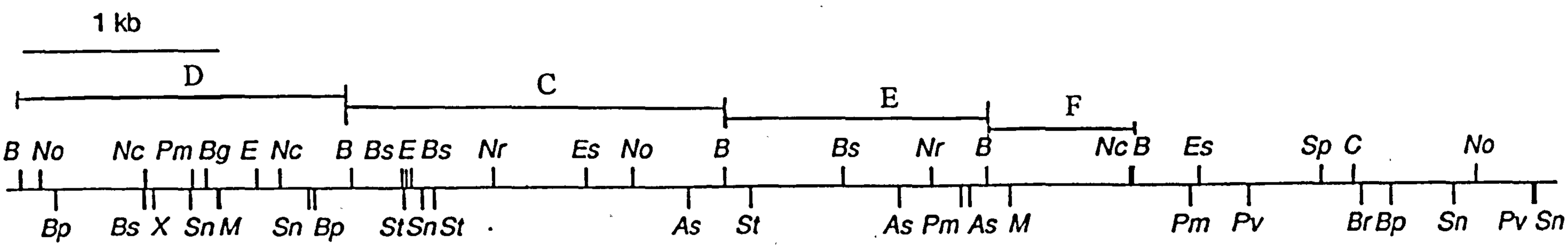
Figure 4.1

Location of DNA inserts of plasmids pIJ2520, pIJ2530 and pIJ2540

The DNA inserts of plasmids pIJ2520, pIJ2530 and pIJ2540 were derived from Pwb-9, Pwb-6 and Pwb-16 respectively and conferred the Pwb phenotype when transferred to a *bldA* strain. The inserts contained fragments C, E and F in common.

Figure 4.1

Location of the DNA inserts of plasmids pIJ2520, pIJ2530 and pIJ2540



As, AscI; B, BamHI; Bg, BglII; Bp, BspEI; Br, BsrGI; Bs, BstEII; C, ClaI; E, EcoRI; Es, EspI; M, MluI; Nc, NcoI; No, NotI; Nr, NruI; Pm, PmlI; Pv, PvuII; Sn, SnoI; Sp, SphI; St, StuI; X, XhoI

Reproduced with permission M.Bibb

4.2 Cloning of the 800bp *Eco*RI regions containing the Pwb mutations

The aim of this study was to further characterise the Pwb mutations. To achieve this the initial aim was to clone and sequence the 800bp *Eco*RI fragment of pIJ2530, pIJ2520, pIJ2540 and the wild-type copy of the region, in order to determine the nature and location of the mutations that caused the Pwb phenotype. The entire *red* cluster was received from F.Malpartida on plasmid pCNB1003 (Malpartida *et al.*, 1990). The Pwb plasmids pIJ2530, pIJ2520 and pIJ2540 were received from K.F.Chater.

4.2.1 Cloning of the 800bp *Eco*RI fragment of pIJ2530

The strategy used to isolate the 800bp *Eco*RI fragment of pIJ2530, pIJ2520, pIJ2540 and pCNB1003 was by restriction, agarose gel electrophoresis and DNA fragment purification using the Geneclean system (2.5.2). The next stage was to clone the DNA fragment into a high copy number *E.coli* plasmid vector. The vector pIC20H was chosen as clones with inserts can be visualised using blue/white colony selection on Xgal medium.

This strategy was successful for the cloning of the pIJ2530 *Eco*RI fragment to create pPWB6. To facilitate sequencing, the *Eco*RI fragment was subcloned into M13 mp18 creating ØCF1 and ØCF2 (Table 4.1 and Table 4.2). The strategy was unsuccessful for the cloning of the same fragment of pIJ2520, pIJ2540 and pCNB1003 due to problems in obtaining sufficient quantities of the plasmid DNA to be able to visualise the 800bp *Eco*RI fragment by agarose gel electrophoresis.

Table 4.2

A summary of clones and deletion clones used in DNA sequencing.

Clone name	Description	Origin of insert	Allele	Trivial name
ØCF1	800bp *Eco*RI fragment in mp18	pIJ2530	*pwb-6*	13
ØCF2	800bp *Eco*RI fragment in mp18	pIJ2530	*pwb-6*	8
ØCF1*Bam*D	*Bam*HI deletion of CF1	CF1	*pwb-6*	B13
ØCF2*Sma*D	*SmaI* deletion of CF2	CF2	*pwb-6*	S8
ØCF2*Bam*D	*Bam*HI deletion of CF2	CF2	*pwb-6*	B8
ØCF2*NcoAsp*D	*Nco*I *Asp*718I deletion of CF2	CF2	*pwb-6*	NA8
ØCF3	800bp *Eco*RI fragment in mp18	pIJ2520	*pwb-9*	51
ØCF4	800bp *Eco*RI fragment in mp18	pIJ2520	*pwb-9*	52
ØCF5	800bp *Eco*RI fragment in mp18	pIJ2540	*pwb-16*	94
ØCF6	800bp *Eco*RI fragment in mp18	pIJ2540	*pwb-16*	99
ØCF7	800bp *Eco*RI fragment in mp18	pIJ4136	*pwb+*	E11
ØCF8	800bp *Eco*RI fragment in mp18	pIJ4136	*pwb+*	E3 or E6
ØCF9	1.7kb *Bam*HI fragment in mp19	pIJ2530	*pwb-6*	3D3
ØCF10	1.7kb *Bam*HI fragment in mp19	pCNB1003	*pwb+*	1003
ØCF9*Hinc*D	*Hinc*II deletion of CF9	CF9	*pwb-6*	Hinc 3D3
ØCF10*Hinc*D	*Hinc*II deletion of CF10	CF10	*pwb+*	Hinc 1003
ØCF11	*Bgl*II *Eco*RI fragment in mp18	pIJ4132	*pwb+*	BglEco18
ØCF12	*Bgl*II *Eco*RI fragment in mp19	pIJ4132	*pwb+*	BglEco19

4.2.2 Cloning of the 800bp *Eco*RI fragment of pIJ2540 and pIJ2520

A second strategy was employed to clone the *Eco*RI fragment of pIJ2520 and pIJ2540, which exploits the observation that smaller DNA fragments are subcloned into linear vectors at a higher frequency than are larger ones. Digestion of pIJ2530 releases the 800bp Pwb containing fragment and two much larger fragments of approximately 9kb and 25kb. Ligation of *Eco*RI restriction digests of pIJ2520 and pIJ2540 into M13 mp18 directly was attempted. The 800bp fragment was ligated into the vector at a much higher frequency than the two much larger fragments. The 800bp *Eco*RI fragments of pIJ2520 and pIJ2540 were successfully cloned into the *Eco*RI site of M13 mp18 using this method to create ØCF3, ØCF4, ØCF5 and ØCF6 (Table 4.1 and Table 4.2).

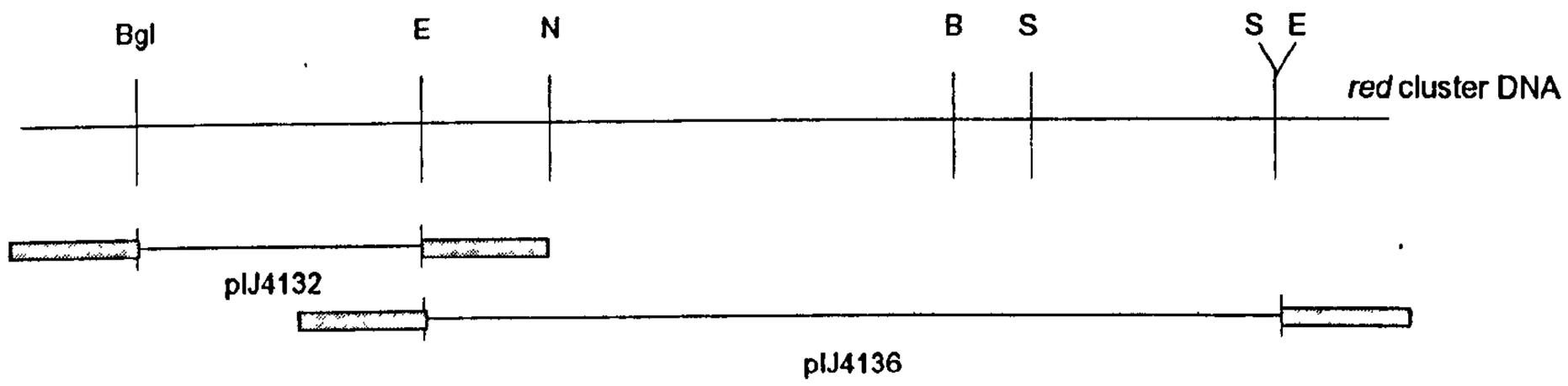
4.2.3 Cloning of the wild-type 800bp *Eco*RI fragment region of the *red* cluster

Cloning the 800bp *Eco*RI fragment from the wild-type *red* cluster was unsuccessful. A smaller clone pIJ4136, containing only the *Eco*RI fragment was made available to us by J.White (Figure 4.2). The *Eco*RI fragment was cloned into M13 mp18 in both orientations to create ØCF7 and ØCF8, in order to facilitate DNA sequencing.

In summary, the 800bp *Eco*RI fragment from the three Pwb mutants and the wild-type were cloned in both orientations into M13 mp18 (Table 4.1 and Table 4.2).

Figure 4.2

Restriction map to show the location of DNA in wild-type clones pIJ4136 and pIJ4132



The plasmids pIJ4132 and pIJ4136 containing Pwb $^{+}$DNA were kindly received from J.White. The *Eco*RI fragment of pIJ4136 was cloned into M13mp18 in both orientations to create ϕCF7 and ϕCF8 (4.2.3). The *Bgl*II *Eco*RI fragment of pIJ4132 was cloned into *Bam*HI *Eco*RI restricted M13mp18 and M13mp19 to create ϕCF11 and ϕCF12 respectively (4.3.3).

4.3 Sequencing Strategy

Restriction analysis of pPWB6 (pIC20H::800bp*Eco*RI, Pwb-6) using enzymes that had a single site in the polylinker of the vector pIC20H revealed the 800bp *Eco*RI fragment contained apparently unique *Bam*HI and *Sma*I sites (Figure 4.3). However, sequence analysis later revealed a second *Sma*I site separated by one base from the *Eco*RI site at the right hand end of the DNA fragment (Figure 4.12). The clones ØCF1 and ØCF2 containing the 800bp *Eco*RI fragment of pIJ2530 were the first clones available for sequencing. DNA sequencing of these clones proceeded while attempts were made to clone the *Eco*RI region of pIJ2520 (Pwb-9), pIJ2540 (Pwb-16) and pIJ4136 (wild-type). A directed approach was taken to sequence ØCF1 and ØCF2 using the unique *Bam*HI site and the internal *Sma*I site to construct deletion clones (Figure 4.4).

4.3.1 Construction of deletion clones of ØCF1 and ØCF2

The *Bam*HI deletion clones were constructed by restriction of the replicative form (RF) of ØCF1 and ØCF2 with this endonuclease. There are two *Bam*HI sites in these clones, one lies in the polylinker of M13mp18 and the other within the DNA insert. Restriction of ØCF1 releases a 500bp fragment and ØCF2 a 300bp fragment. The linear *Bam*HI digested RF DNA was removed from an agarose gel after electrophoresis to separate the two fragments. The DNA was purified from the agarose and re-ligated (Chapter 2). All other deletion clones were constructed by a similar method. As can be seen from Figure 4.4 *Bam*HI deletion of both ØCF1 and ØCF2 were obtained to create ØCF1*Bam*D and ØCF2*Bam*D. A *Sma*I deletion ØCF2 was obtained,

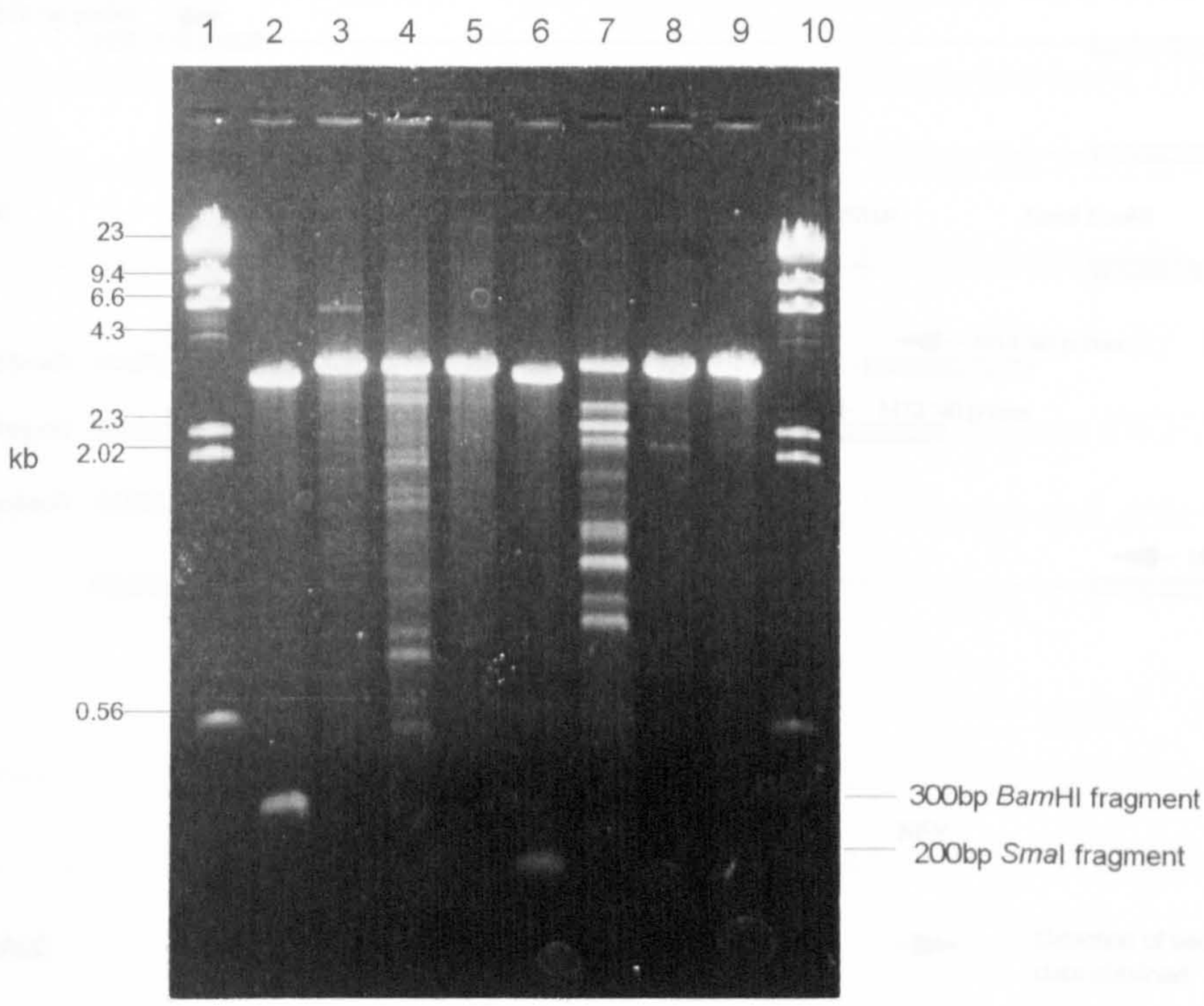
Figure 4.3

Restriction analysis of pPWB6

The pPWB6 plasmid was digested with restriction enzymes that have a single site in the polylinker of the vector. The number of bands released by digestion with a restriction enzyme is an indication of the number of sites within the 800bp *Eco*RI DNA insert. As can be seen restriction with *Bam*HI and *Sma*I released a single band indicating a single site in the DNA insert

Figure 4.3

Restriction analysis of pPWB6

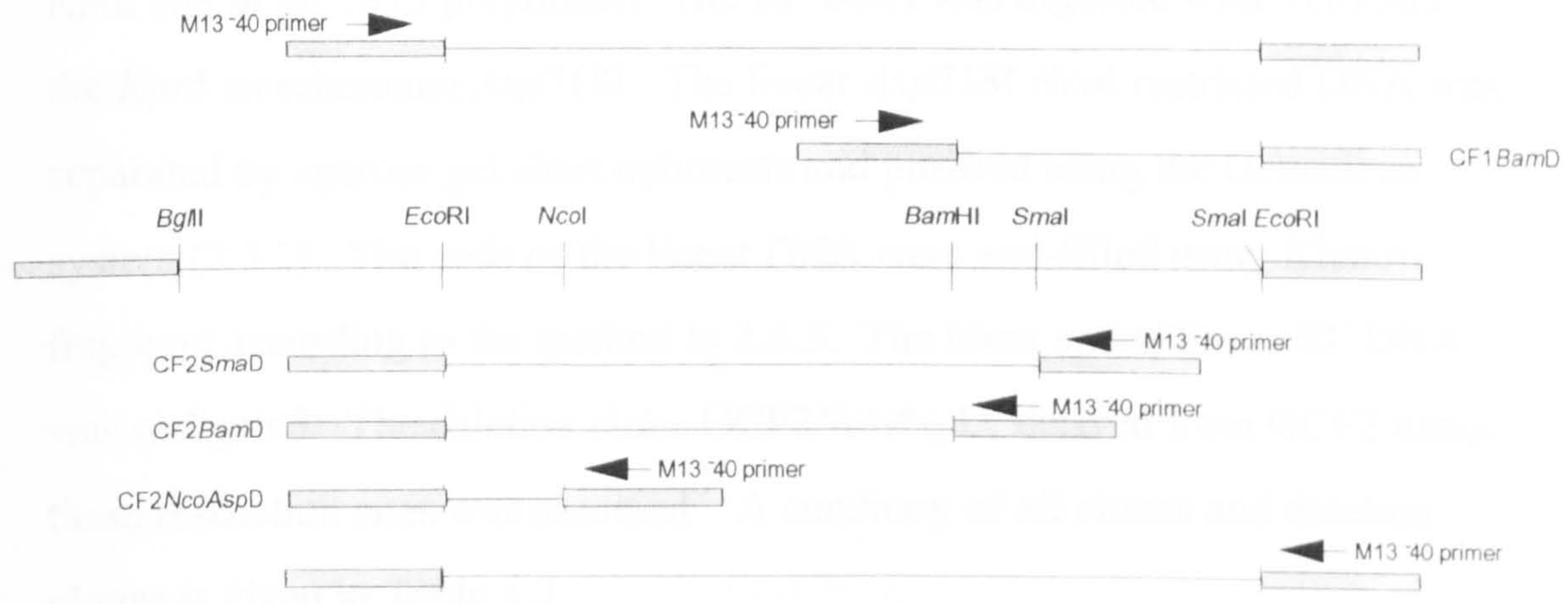


Lane order
1 *Hind*III λ
10 *Hind*III λ

pPWB6 digested with:-
2 *Bam*HI
3 *Pst*I
4 *Xho*I
5 *Kpn*I
6 *Sma*I
7 *Eco*RV
8 *Sal*I
9 *Bgl*II

Figure 4.4

Restriction map to show clones and deletion clones used in DNA sequencing



A directed approach was taken to sequence the 800bp *Eco*RI Pwb-6 DNA fragment. Deletion clones of ϕCF1 and ϕCF2 were constructed and sequence data obtained using the M13 -40 primer

ØCF2*Sma*D (Figure 4.4). The M13 -40 universal primer was used in sequencing reactions of these deletion clones.

Sequence data obtained from ØCF1 using the M13 -40 primer revealed an *Nco*I restriction site. Restriction of the DNA clone suggested the *Nco*I was unique. A deletion clone was constructed using this unique *Nco*I site and the *Kpn*I site in the M13 polylinker. The RF DNA was digested with *Nco*I and the *Kpn*I isoschizomer *Asp*718I. The linear *Asp*718I *Nco*I restricted DNA was separated by agarose gel electrophoresis and purified using the Geneclean system (2.5.2). The ends of the linear DNA were end-filled using Klenow fragment according to the method in 2.5.5. The blunt ended linear RF DNA was re-ligated. The deletion clone ØCF2*NcoAsp*D, derived from ØCF2 using these restriction sites was obtained . A summary of all clones and deletion clones is given in Table 4.2.

Oligonucleotides were designed so that complete double stranded sequence could be obtained (Table 4.3). The location of oligonucleotides is shown in Figure 4.5.

4.3.2 Analysis of the sequence data obtained of the 800bp *Eco*RI fragment of Pwb-6

Computer analysis of the DNA sequence data of the 800bp *Eco*RI fragment revealed two incomplete reading frames. The first was 201 amino acids long stretching over 605 bp from the left hand *Eco*RI site (Figure 4.6). The second open reading frame was 40 amino acids long and transcribed in the opposite direction to ORF1. Similarity of ORF1 to the UhpA-LuxR family of transcriptional activators suggested that this was the ORF of interest. The

Table 4.3

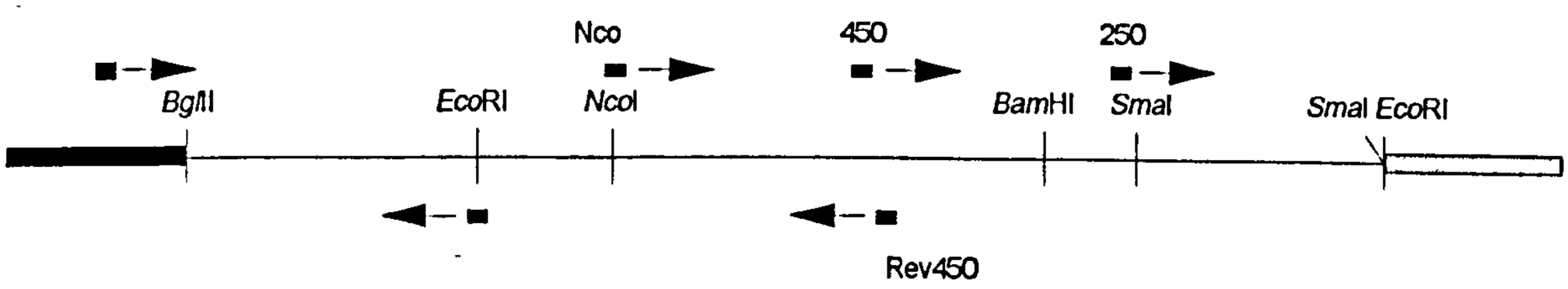
Oligonucleotides used in DNA sequencing

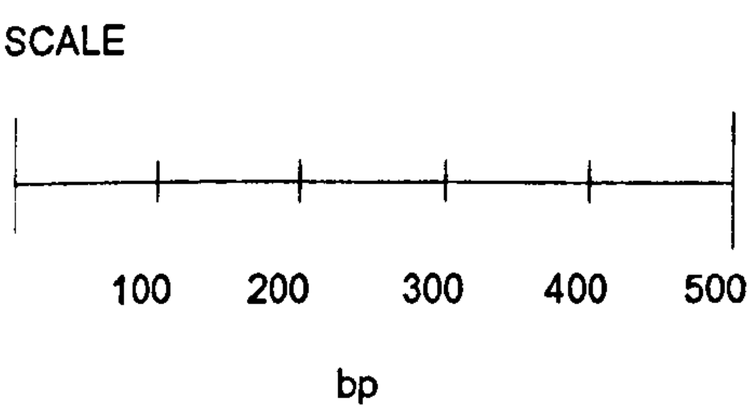
Name	Sequence	Length in bp	Position on sequence
250	GGG CCT CCT GGG CAT CT	17	818 - 834
450	AGC AAG GCC CCG AAG AA	17	615 - 631
Nco	TTC ACC ATG GAC AGC AT	17	378 - 394
Rev 450	TTC CCG AGG GGT CAG TT	17	660 - 643
Norwich 1	GTG CAC GTG TCC TTC TGA G	19	75bp upstream of *Bgl*II site
Norwich 2	AAT TCC CTG GCC CAG GAT GAC	21	263 - 243

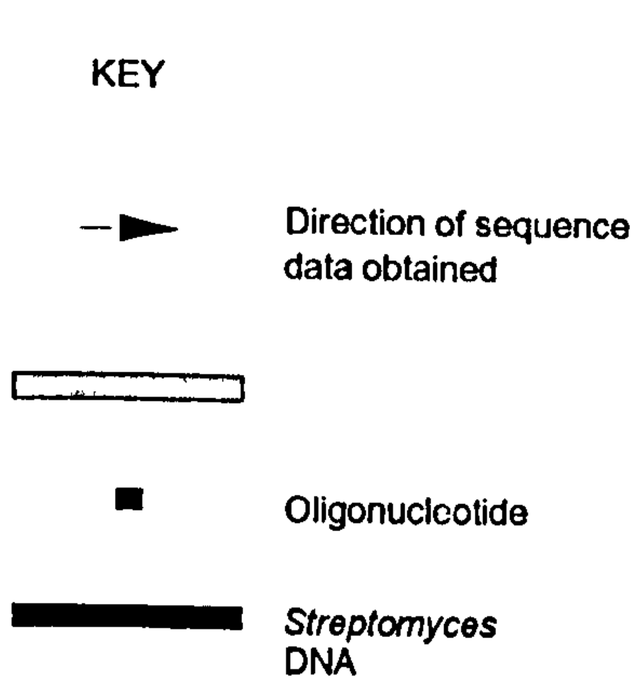
Table 4.3 shows the nucleotide sequence, length and position on the complete DNA sequence (Figure 4.10) of oligonucleotides used in DNA sequencing.

Figure 4.5

Restriction map to show location of oligonucleotides used in DNA sequencing







Oligonucleotides were designed to obtain double stranded sequence of the Pwb-6 800bp *Eco*RI region. The location of oligonucleotides is shown.

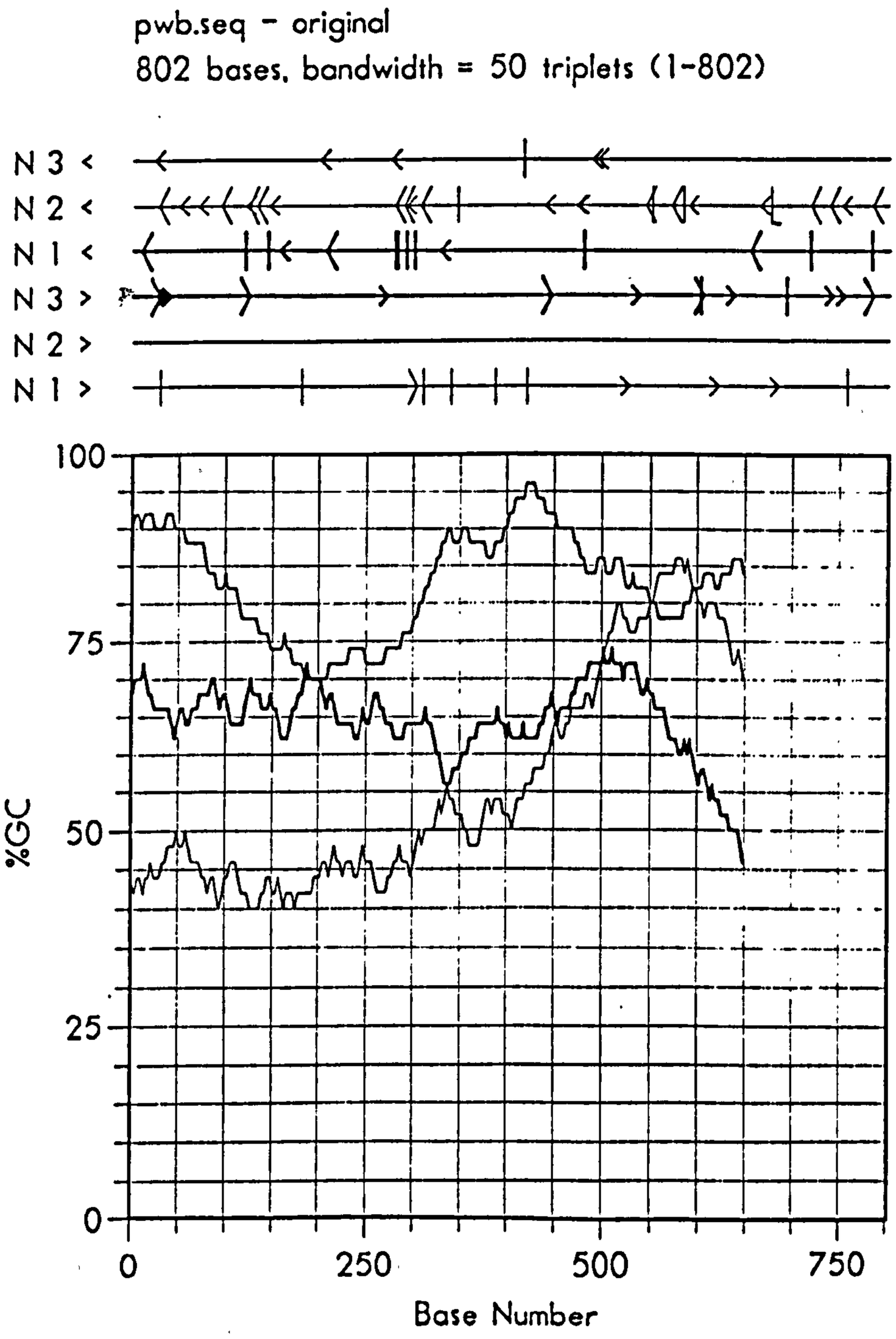
Figure 4.6

FRAME analysis of the 800bp *Eco*RI fragment

FRAME analysis identifies open reading frames. Streptomycete DNA has 74% mol G + C content. If bases are randomly arranged the frequency of distribution of guanidine and cytosine nucleotides at any base position is in the region of 74%. In a coding region this distribution changes, the likelihood of finding a G or C base at position 1 in a codon is 70%, at position 2 in the codon the frequency expected is 50% and at position 3 the expected frequency is 90%. The DNA sequence is divided into triplets starting at base 1 to perform the FRAME analysis. From the FRAME plot on the facing page at position 0, the frequency of distribution of finding a G or C nucleotide is 90% for the position N2 of the assigned triplet is 90%, 70% for position N3 and 50% for position N1. This distribution is similar to that expected of a coding region. the reading frame of the coding region can be predicted. The first base in each codon of the predicted open reading frame corresponds to the bases designated N3.

Figure 4.6

FRAME Analysis of the 800bp *Eco*RI fragment



ORF1 was incomplete so it was necessary to clone the DNA upstream of the left hand *Eco*RI fragment in order to obtain the sequence of the whole gene.

4.3.3 Cloning of the DNA upstream of ORF1

The wild-type clone containing DNA of the region upstream of the left hand *Eco*RI site, pIJ4132 was supplied by J.White (Figure 4.2). The *Bgl*II *Eco*RI 260bp fragment was cloned into *Bam*HI *Eco*RI digested M13 vectors mp18 and mp19 to create ØCF11 and ØCF12 respectively. Cloning of the *Bgl*II *Eco*RI region of the Pwb mutants was less straightforward as it required cloning from the SCP2-derived plasmids pIJ2530, pIJ2520, pIJ2540 and pCNB1003. Obtaining DNA of these plasmids in sufficient quantity for cloning had previously caused difficulties. The strategy employed was one of random cloning from *Bam*HI digested pIJ2530, pIJ2520 and pIJ2540 plasmids into *Bam*HI digested M13 mp18. The *Bam*HI fragment containing the Pwb region is 1.7kb. The desired insert would contain unique *Bgl*II and *Eco*RI sites. Restriction of the fragment with *Bam*HI and *Bgl*II would release fragments of 900bp and 800bp. Restriction with *Bam*HI *Bgl*II and *Eco*RI would release fragments of 300bp, 500bp and 900 bp (Figure 4.7).

The M13 plaques showed inactivation of the *lacZ* gene if a DNA insert had been successfully ligated into the vector. White plaques were screened by preparing RF DNA and examining the restriction pattern with *Bam*HI to determine the presence and size of the DNA insert. Of 180 white M13 plaques screened, two (had a 1.7kb *Bam*HI insert) with the correct restriction pattern (Figure 4.8). The clones ØCF9 and ØCF10 contained inserts of pIJ2530 (Pwb-6) and pCNB1003 (wild-type) DNA respectively. Cloning of the *Bam*HI fragments of pIJ2520 and pIJ2540 was unsuccessful.

Figure 4.7

Cloning strategy used to obtain the 1.7kb *Bam*HI DNA fragment of the *red* cluster

DNA sequencing of the 800bp *Eco*RI fragment revealed an incomplete open reading frame. The strategy proposed to obtain the upstream region of the 800bp *Eco*RI fragment was to clone the 1.7kb *Bam*HI fragment of the red cluster into *Bam*HI restricted M13mp18. The expected restriction pattern of the 1.7kb *Bam*HI fragment is shown on the facing page. Clones with the desired DNA insert are expected to show this restriction pattern.

Figure 4.7

Cloning strategy used to obtain the 1.7kb *Bam*HI DNA fragment of the *red* cluster

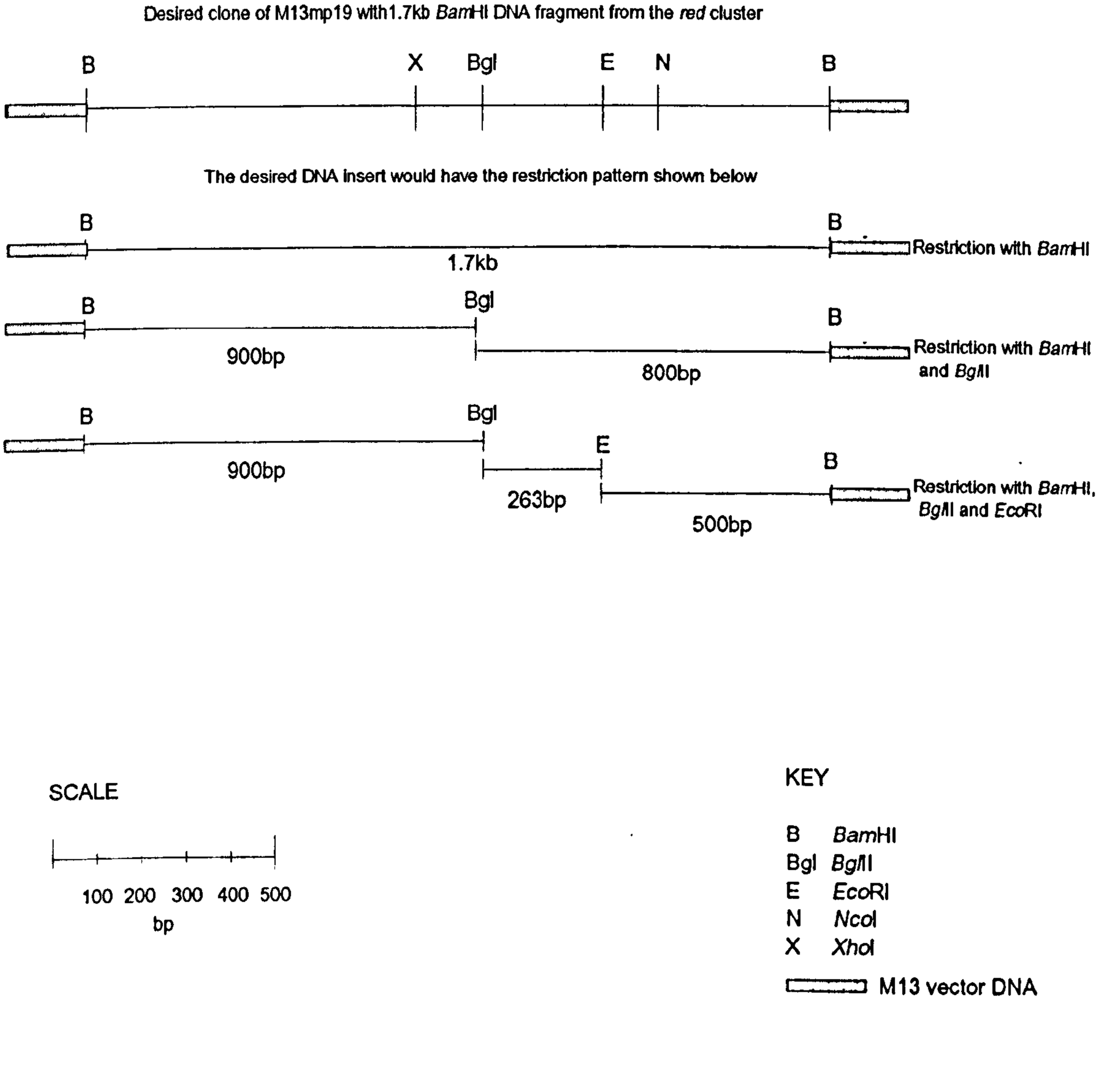


Figure 4.8

Photograph of restriction digests of ∅CF9 and ∅CF10

The clones ∅CF9 and ∅CF10 showed the predicted restriction pattern of the desired 1.7kb BamHI DNA fragment of the red cluster (Figure 4.7). On restriction with BamHI a 1.7kb fragment should be released in addition to the a band corresponding to the linear vector of approximately 8kb. Restriction of ∅CF9 and ∅CF10 with *Bam*HI and *Bgl*II should release two fragments from the 1.7kb DNA insert 900 and 800bp in length. Restriction of ∅CF9 and ∅CF10 with *Bam*HI, *Bgl*II and *Eco*RI should release three fragments from the DNA insert of 900bp, 500bp and approximately 260bp. Restriction of ∅CF9 and ∅CF10 with *Eco*RI allows the orientation of the DNA inserts to be identified (Figure 4.9).

Table to show size of restriction fragments obtained from ∅CF9 and ∅CF10

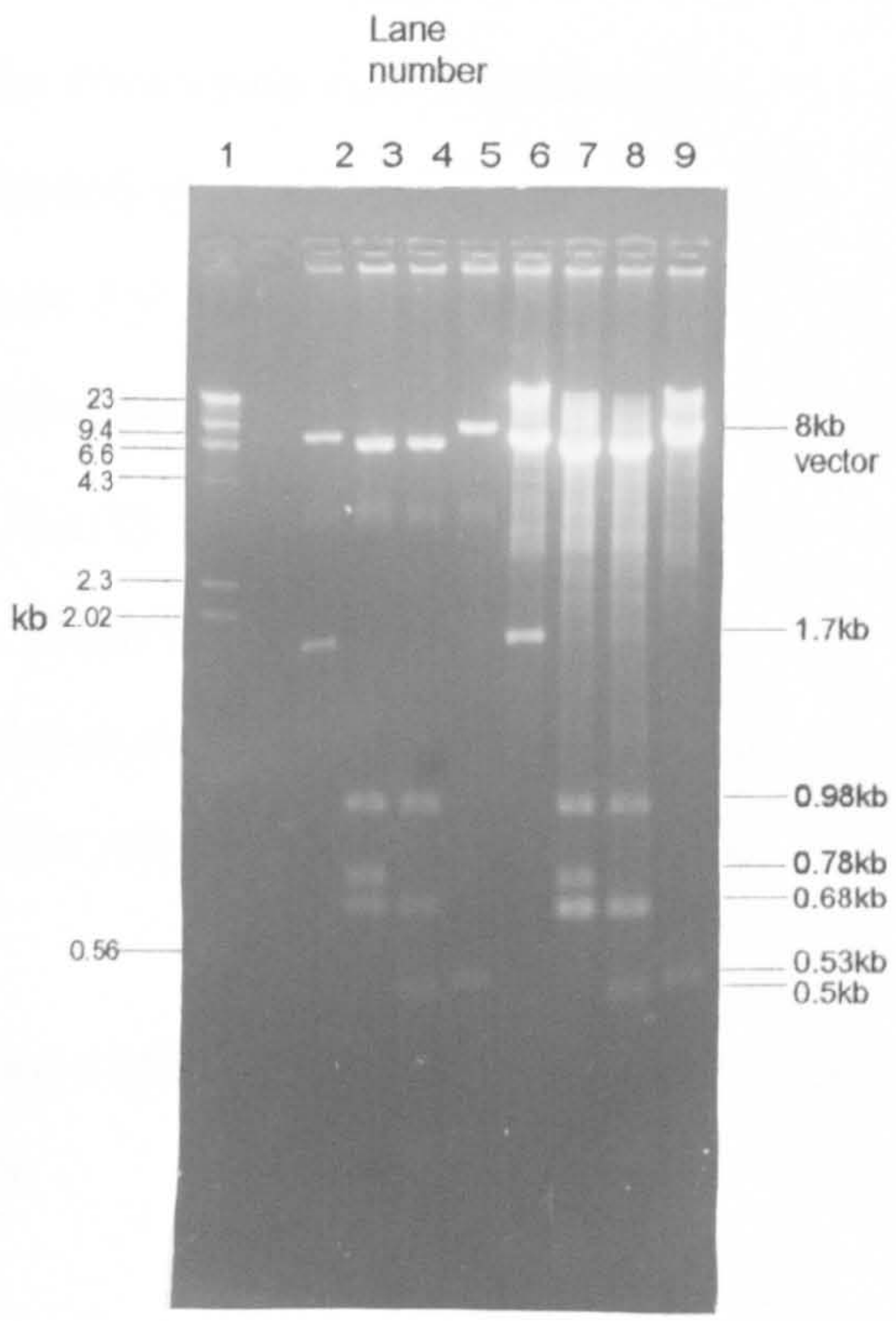
*Bam*HI	*Bam*HI/*Bgl*II	*Bam*HI/*Bgl*II/*Eco*RI	*Eco*RI
6.8 (M13)	6.8 (M13)	6.8 (M13)	8 (M13)
1.75 **(1.7)**			
	0.98 **(0.9)**	0.98 **(0.9)**	
	0.78 **(0.8)**		
	0.68 (M13)	0.68 (M13)	
			0.53 **(0.5)**
	0.5 **(0.5)**		
	(0.26 not seen)		

Figures given are in kilobases

Figures in bold in parenthesis shown the expected size of restriction fragments

Figure 4.8

Photograph of restriction digests of
ϕ CF9 and ϕ CF10



Lane Order

1 *Hind*III λ

2 *Bam*HI
3 *Bam*HI/*Bgl*II
4 *Bam*HI/*Bgl*II/*Eco*RI
5 *Eco*RI
} ϕ CF9

6 *Bam*HI
7 *Bam*HI/*Bgl*II
8 *Bam*HI/*Bgl*II/*Eco*RI
9 *Eco*RI
} ϕ CF10

Sequence of the DNA upstream of the left hand *Eco*RI fragment was obtained by constructing deletion clones of ØCF9 and ØCF10. The orientation of the *Bam*HI fragment with respect to the M13 polylinker was determined by restriction endonuclease digestion with *Eco*RI . The size of the fragment released was 500bp (Figure 4.8). The orientation of the fragment is given in Figure 4.9. The 1.7kb *Bam*HI DNA inserts of ØCF9 and ØCF10 were both in the same orientation. In order to obtain DNA sequence that spanned the *Eco*RI site *Hinc*II deletion constructs were obtained. It was necessary to identify whether there were two *Eco*RI sites located closely together. Further sequence data of the upstream region was obtained using oligonucleotide primers Norwich 1 and Norwich 2 (Table 4.3).

4.3.4 Strategy used to sequence ØCF3, ØCF4, ØCF5, ØCF6, ØCF7 and ØCF8

Sequence data of the 800bp *Eco*RI fragment was obtained for; the Pwb-9 clones ØCF3 and ØCF4, the Pwb-16 clones ØCF5 and ØCF6 and the Pwb^+ clones ØCF7 and ØCF8. The M13 universal primers and primers designed to sequence Pwb-6 clones were used to obtain sequence data (Figure 4.5). Deletion clones were not constructed in Pwb-9, Pwb-16 or for Pwb^+ clones. The single stranded DNA sequence data obtained for Pwb-9, Pwb-16 and Pwb^+ is represented on Figure 4.10.

4.4 Sequencing Data

DNA sequence was obtained using the dideoxy chain termination method. In summary, DNA sequence was obtained of the 800bp *Eco*RI region of the

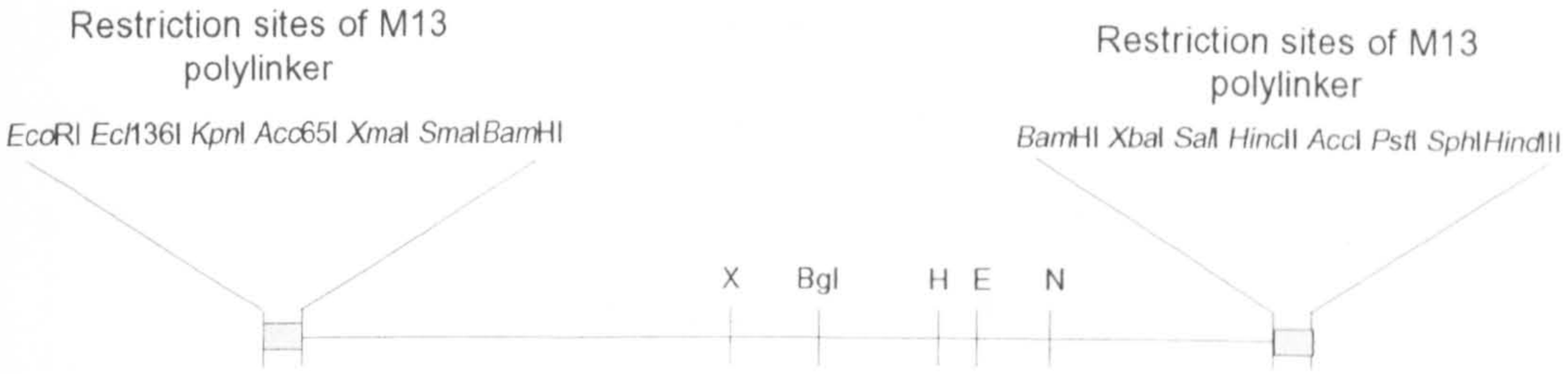
Figure 4.9

Orientation of the DNA inserts of ∅CF9 and ∅CF10

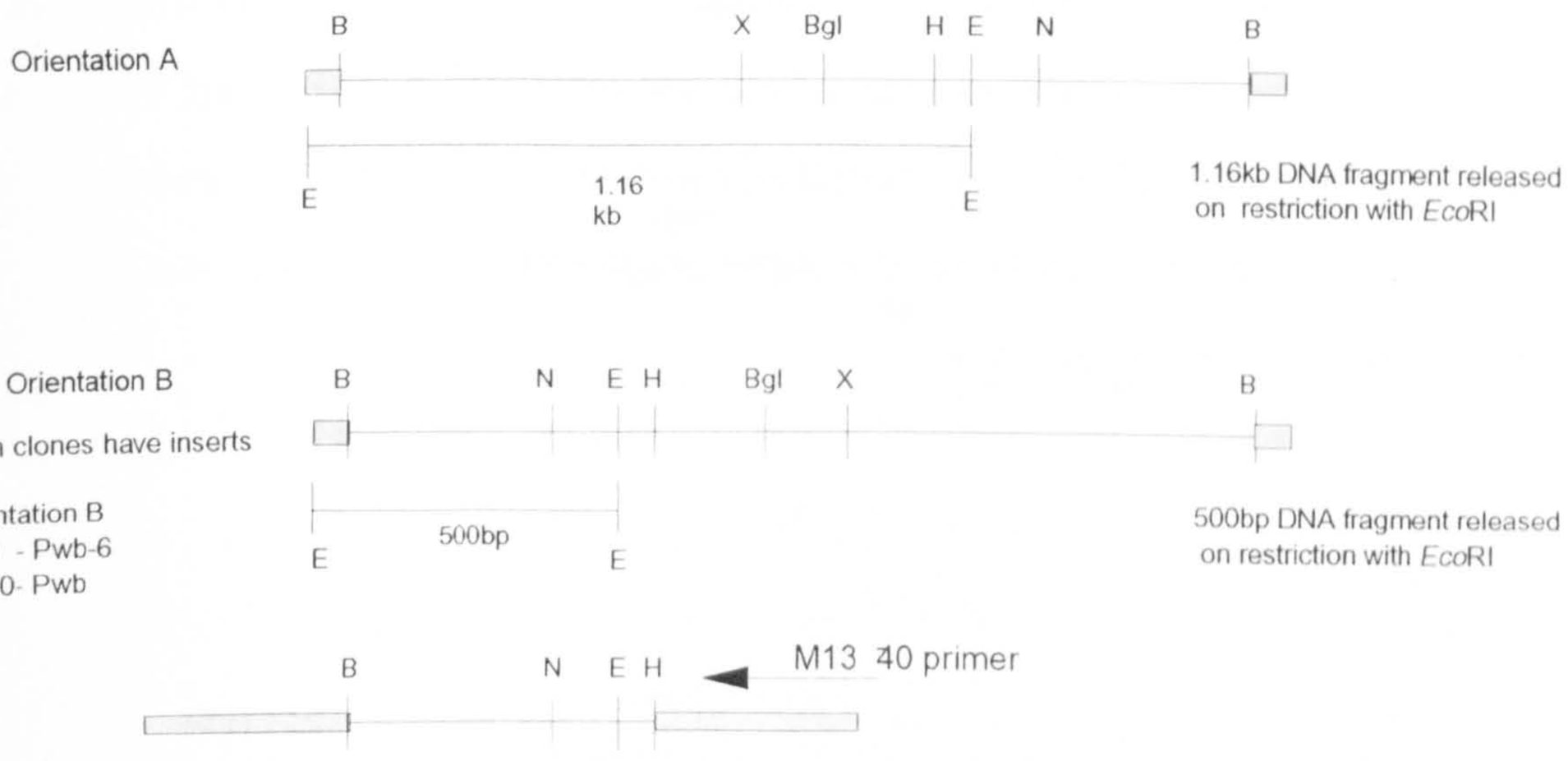
The 1.7kb *red* DNA fragment was cloned into ∅CF9 and ∅CF10. It was necessary to determine the orientation of the DNA insert in order to construct deletion clones so that DNA sequence could be obtained. The clones ∅CF9 and ∅CF10 DNA inserts were of the same orientation, corresponding to orientation B shown on the facing page.

Figure 4.9

Orientation of the DNA inserts of ϕCF9 and ϕCF10



There are two possible orientations of the 1.7kb *Bam*HI insert in the M13mp19 vector. Restriction with *Eco*RI identifies the orientation of clones ¢CF9 and ¢CF10.



*Hinc*II deletion clones of ϕCF9 and ϕCF10 provide DNA sequence data spanning the *Eco*RI site to be obtained using the M13 -40 primer.

Figure 4.10

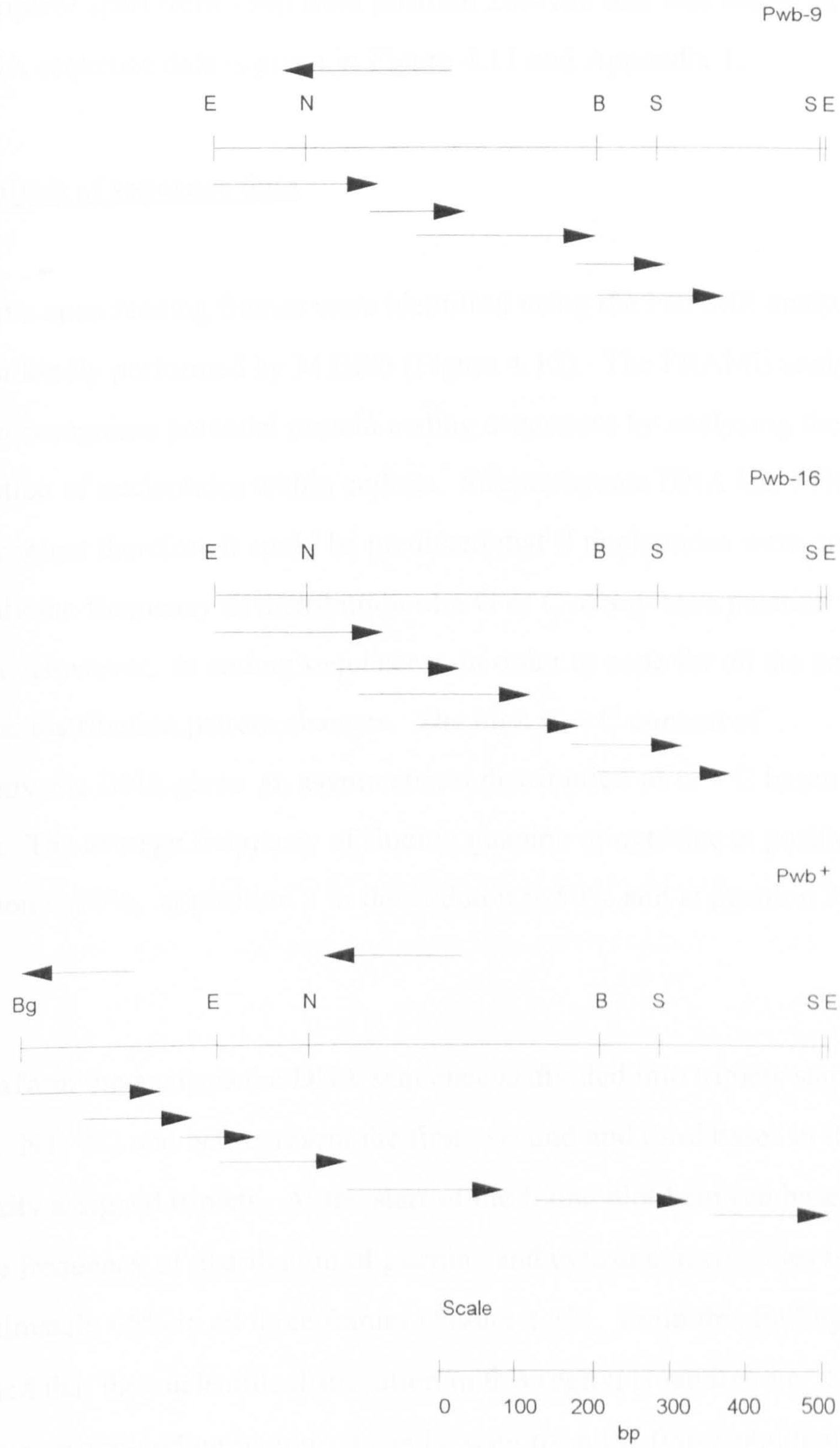
DNA Sequence obtained for Pwb-9, Pwb-16 and Pwb$^+$

The region for which sequence data has been obtained is represented by the arrows on the facing page. The arrows do not represent individual sequences obtained but merely represent the regions for which sequence has been obtained. Except in the case of Pwb-9 where the arrows do represent individual sequences.

Incomplete single stranded sequence has been obtained for Pwb$^+$ clones. Sequence of Pwb-9 and Pwb-16 covers from the *Eco*RI site to past the *Bam*HI site. The open reading frame ORF1 finishes at position 864, this corresponds to between the *Bam*HI and left-hand *Sma*I site on this map. Although single stranded sequence has not been obtained for the entire 800bp fragment for Pwb-9 and Pwb-16 it can be seen that the single stranded sequence that has been obtained covers ORF1.

Figure 4.10

DNA Sequence obtained for Pwb-9, Pwb-16 and Pwb$^+$



wild-type and all three of the Pwb mutants. In the case of Pwb-6 and the wild-type (Pwb^+) contiguous sequence data was obtained of the BglII *Eco*RI 1.06kb fragment. Double stranded sequence was obtained for Pwb-6, which was complete apart from 75bp from position 263-188 that was single stranded. The DNA sequence data is given in Figure 4.11 and Appendix 1.

4.5 Analysis of sequence data

Putative open reading frames were identified using the FRAME analysis program kindly performed by M.Bibb (Figure 4.12). The FRAME analysis program recognises potential protein coding sequences by analysing the distribution of nucleotides within codons. Streptomycete DNA has 74% G + C content therefore it could be predicted that if nucleotides were arranged randomly the frequency of distribution of a G or C at any base position would be 74%. However, in coding sequences, in order to code for all the amino acids the distribution pattern changes. The high G + C content of streptomycete DNA gives an asymmetrical distribution of G + C bases across codons. The average frequency of finding guanine or cytosine at position 1 in the codon is 70%, at position 2 in the codon it is 50% and at position 3 is 90%.

To perform the analysis the DNA sequence is divided into triplets starting at base 1, N1, N2 and N3 represent the first, second and third bases in these arbitrarily assigned triplets. At the start of the frame plot between base 1 to 100 the frequency of distribution of guanine and cytosine nucleotides is approximately 65% in all three frames (Figure 4.12). From this finding it is predicted that the nucleotide distribution in this region is random i.e. it is therefore, a non-coding region. It can be seen from the frame plot that the

Figure 4.11

Nucleotide and predicted amino acid sequence of *redZ*

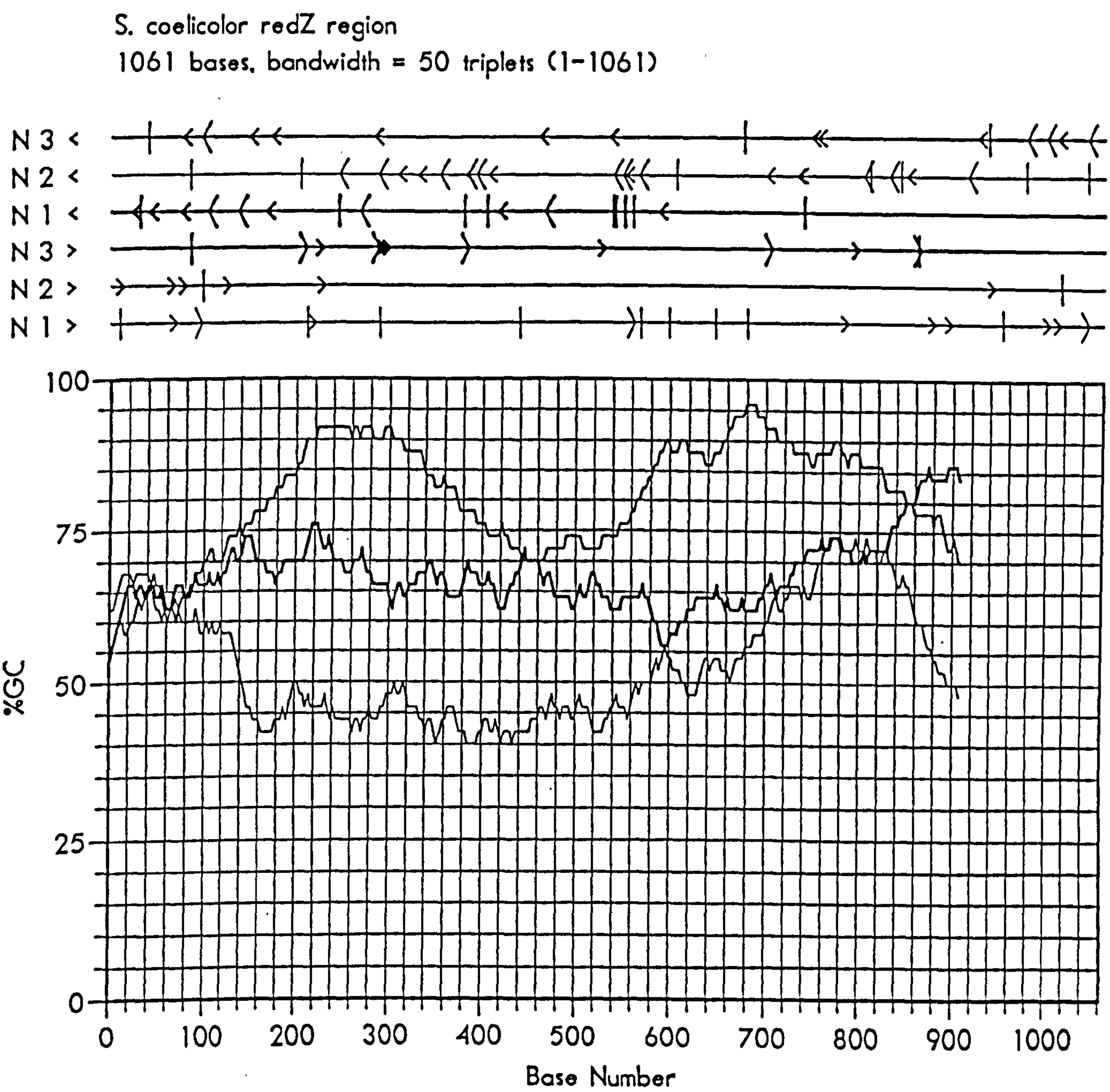
```
     GGATCTTCTTGAGGTGGAAACCACTTCGTATCAGTCTCTCACCAGGGCCTCGCAAGCCCC
1    ---------+---------+---------+---------+---------+---------+ 60
     CCTAGAAGAACTCCACCTTTGGTGAAGCATAGTCAGAGAGTGGTCCCGGAGCGTTCGGGG

      T
     TCTCCAAGTGTGCACACGCGTGCTAAGTTTGGCCGCATGAGGCATATCGGGGGAATCGCA
61   ---------+---------+---------+---------+---------+---------+120
     AGAGGTTCACACGTGTGCGCACGATTCAAACCGGCGTACTCCGTATAGCCCCCTTAGCGT


     AAGAAAGGTGCCCGGCATCGACTTGCCACCGTCGGCGCCGGCCCGCACGCCACGTACGGT
121  ---------+---------+---------+---------+---------+---------+180
     TTCTTTCCACGGGCCGTAGCTGAACGGTGGCAGCCGCGGCCGGGCGTGCGGTGCATGCCA

                                        M  T  T  R  V  L  V  C  C  D
     CCCCATCCTTCCTGGACGAAAGTCAACGTATGACGACCCGTGTCCTGGTGTGCTGCGACC
181  ---------+---------+---------+---------+---------+---------+240
     GGGGTAGGAAGGACCTGCTTTCAGTTGCATACTGCTGGGCACAGGACCACACGACGCTGG

     R  V  I  L  G  E  G  I  R  A  L  L  E  R  H  D  M  K  V  Q
     GCGTCATCCTGGGCGAGGGAATTCGCGCATTGCTGGAGCGGCACGACATGAAAGTGCAGG
241  ---------+---------+---------+---------+---------+---------+300
     CGCAGTAGGACCCGCTCCCTTAAGCGCGTAACGACCTCGCCGTGCTGTACTTTCACGTCC

     V  E  T  T  Q  R  G  S  L  A  T  A  A  E  T  G  P  D  I  L
     TGGAGACCACCCAGCGCGGCTCGCTGGCCACGGCGGCGGAGACCGGGCCCGACATTCTGG
301  ---------+---------+---------+---------+---------+---------+360
     ACCTCTGGTGGGTCGCGCCGAGCGACCGGTGCCGCCGCCTCTGGCCCGGGCTGTAAGTCC

     V  G  V  A  P  L  F  T  M  D  S  I  D  K  L  T  E  L  A  R
     TCGGCGTCGCCCCCCTCTTCACCATGGACAGCATCGACAAGCTCACGGAACTCGCACGAC
361  ---------+---------+---------+---------+---------+---------+420
     AGCCGCAGCGGGGGGAGAAGTGGTACCTGTCGTAGCTGTTCGAGTGCCTTGAGCGTGCTG

     L  G  K  T  L  L  L  T  K  P  E  N  T  H  R  A  F  E  A  L
     TCGGCAAGACGCTTTTGCTGACCAAGCCCGAGAATACCCACCGCGCATTCGAGGCGCTCC
421  ---------+---------+---------+---------+---------+---------+480
     AGCCGTTCTGCGAAAACGACTGGTTCGGGCTCTTATGGGTGGCGCGTAAGCTCCGCGAGG

     R  V  G  V  R  A  V  L  S  A  E  T  S  V  E  E  L  V  H  V
     GCGTCGGAGTACGCGCCGTATTGTCGGCCGAAACGTCGGTCGAAGAACTGGTGCACGTCA
481  ---------+---------+---------+---------+---------+---------+540
     CGCAGCCTCATGCGCGGCATAACAGCCGGCTTTGCAGCCAGCTTCTTGACCACGTGCAGT

     I  R  T  I  T  E  V  N  A  I  I  A  P  E  E  A  Q  E  A  L
     TCAGAACCATCACCGAAGTCAATGCCATAATCGCTCCGGAAGAGGCACAGGAAGCCCTGA
541  ---------+---------+---------+---------+---------+---------+600
     AGTCTTGGTAGTGGCTTCAGTTACGGTATTAGCGAGGCCTTCTCCGTGTCCTTCGGGACT
```

Figure 4.11 continued

```
     T  R  Y  W  R  S  K  A  P  K  N  L  R  P  E  L  T  P  R  E
     CGCGCTATTGGCGCAGCAAGGCCCCGAAGAACCTCCGCCCCGAACTGACCCCTCGGGAAA
601  ---------+---------+---------+---------+---------+---------+660
     GCGCGATAACCGCGTCGTTCCGGGGCTTCTTGGAGGCGGGGCTTGACTGGGGAGCCCTTT

     T  E  V  L  L  L  L  T  Q  G  K  T  N  T  E  M  A  A  T  L
     CGGAAGTCCTCCTCTTACTGACCCAGGGCAAGACCAACACCGAGATGGCCGCGACCCTCT
661  ---------+---------+---------+---------+---------+---------+720
     GCCTTCAGGAGGAGAATGACTGGGTCCCGTTCTGGTTGTGGCTCTACCGGCGCTGGGAGA

     S  V  S  P  T  T  V  R  S  H  V  H  R  I  L  R  K  L  G  A
     CCGTCTCGCCCACCACCGTCAGGTCCCACGTACACCGGATCCTGCGCAAACTGGGCGCGG
721  ---------+---------+---------+---------+---------+---------+780
     GGCAGAGCGGGTGGTGGCAGTCCAGGGTGCATGTGGCCTAGGACGCGTTTGACCCGCGCC

     A  T  R  A  Q  A  V  A  I  A  Y  E  S  G  L  L  G  I  C  P
     CGACCCGTGCGCAGGCCGTGGCCATCGCCTACGAGTCGGGCCTCCTGGGCATCTGCCCGG
781  ---------+---------+---------+---------+---------+---------+840
     GCTGGGCACGCGTCCGGCACCGGTAGCGGATGCTCAGCCCGGAGGACCCGTAGACGGGCC

     G  Y  G  T  P  A  R  *
     GTTACGGCACTCCCGCCCGCTGAATGCCCCGGTCCGGCCGTGCCGCTTCGCGTTCCGTGC
841  ---------+---------+---------+---------+---------+---------+900
     CAATGCCGTGAGGGCGGGCGACTTACGGGGCCAGGCCGGCACGGCGAAGCGCAAGGCACG


     GGCCCGGCCGGGCATCCCGCGCCACGGCCCTTTGTCAGGGCCGTGCCTGGTAGCCGGGTT
901  ---------+---------+---------+---------+---------+---------+960
     CCGGGCCGGCCCGTAGGGCGCGGTGCCGGGAAACAGTCCCGGCACGGACCATCGGCCCAA
                                          *  P  R  A  Q  Y  G  P  E

     CCAGGAGTTCGCCGCTCATCTCCTCGACGACCGTCGCCATCGTGTCGAACACGGTGACCA
961  ---------+---------+---------+---------+---------+---------+1020
     GGTCCTCAAGCGGCGAGTAGAGGAGCTGCTGGCAGCGGTAGCACAGCTTGTGCCACTGGT
       L  L  E  G  S  M  E  E  V  V  T  A  M  T  D  F  V  T  V  L

     GGTCCCGAAAGGCCTCGGTTTTCATGTCCCGGGAGAATTC
1021 ---------+---------+---------+---------+ 1060
     CCAGGGCTTTCCGGAGCCAAAAGTACAGGGCCCTCTTAAG
       D  R  F  A  E  T  K  M  D  R  S  F  E
```

Figure 4.12

FRAME Analysis of *redZ*



distribution of guanine and cytosine nucleotides becomes non-random in the region of nucleotide 150-200. At position 200 the nucleotides at position N3 has a frequency of distribution of guanine and cytosine nucleotides is 70%, at N2 the frequency is 85% and at N1 is 50%. This distribution is indicative of a coding region. It predicts that in this region the nucleotides represented by N3 lie in position 1 in the codon of an open reading frame, N1 lie in position 2 and N2 lie in position 3 of the codon. The frame analysis identifies the reading frame and approximate start of the first open reading frame ORF1. The start codon of ORF1 is an ATG methionine codon at position 210. This is the start site of the predicted protein sequence (Figure 4.11). The predicted protein sequence ends at nucleotide 864 with the opal termination codon (TGA). Due to the location of ORF1 in the *red* cluster, ORF1 has been assigned the gene name *redZ*. Restriction analysis of *redZ* is shown in Figure 4.13.

The second open-reading frame is incomplete and represents the C-terminus of a gene transcribed in the opposite direction to ORF1. Sequence data of the *red* cluster obtained by M.Bibb and co-workers identified the open reading frame represented by ORF2 which in their studies was named *redY* (Figure 4.1). The *redY* gene function is unknown (M.Bibb personal communication).

The prediction made by the FRAME program of the reading frame is supported by initial sequence similarities identified between the potential protein sequence encoded by the 800bp *Eco*RI fragment and other sequences in the databases. A search was made of the OWL database using a combined protein sequence. The DNA sequence was translated in all six reading frames and the six translated sequences were joined together to make one protein sequence. The vast majority of homologus sequences identified in the

Figure 4.13

Restriction map of *redZ*

```
          10          20          30          40          50          60
GGATCTTCTT GAGGTGGAAA CCACTTCGTA TCAGTCTCTC ACCAGGGCCT CGCAAGCCCC
XS M        M                       B     H    B ESHMM       C    M
HA B        N                       S     P    S CAAWN       V    N
OU O        L                       M     H    A OUEOL       I    L
2A 2        1                       1     1    1 96311       1    1

          70          80          90         100         110         120
TCTCCAAGTG TGCACACGCG TGCTAAGTTT GGCCGCATGA GGCATATCGG GGGAATCGCA
          A       MF        D        G HF    N  M                 T
          P       LN        D        D AN    L  N                 F
          A       UU        E        I EU    A  L                 I
          1       12        1        2 31    3  1                 1

          130         140         150         160         170         180
AAGAAAGGTG CCCGGCATCG ACTTGCCACC GTCGGCGCCG GCCCGCACGC CACGTACGGT
       BS    NH  S  M                NH NH S  F          BMSR    N
       AD    CP  F  M                AH AP A  A          SAPS    L
       NU    IA  A  E                RA EA U  U          AELA    A
       11    12  1  1                11 12 6  1          1211    4

          190         200         210         220         230         240
CCCCATCCTT CCTGGACGAA AGTCAACGTA TGACGACCCG TGTCCTGGTG TGCTGCGACC
   F        B   P        HM  M         P         B        M   MT  F
   O        S   S        IM  A         S         S        W   CT  N
   K        T   H        NE  E         H         T        O   RH  U
   1        1   1        21  2         1         1        1   11  2

          250         260         270         280         290         300
GCGTCATCCT GGGCGAGGGA ATTCGCGCAT TGCTGGAGCG GCACGACATG AAAGTGCAGG
H    F  E       M   ET     FH        TMG     F          N       B B
G    O  C       N   CS     NH        TWS     N          L       S S
A    K  O       L   OP     UA        HOU     U          A       G P
1    1  1       1   11     21        211     1          3       1 1

          310         320         330         340         350         360
TGGAGACCAC CCAGCGCGGC TCGCTGGCCA CGGCGGCGGA GACCGGGCCC GACATTCTGG
 BB     T      MFF C       BHBD B   F E  BB    N AS      D
 SS     A      WNN V       AAGS C   N C  SS    C PA      R
 MA     Q      OUU I       LELA E   U I  MA    I AU      D
 11     0      121 1       1311 1   1 1  11    1 16      1

          370         380         390         400         410         420
TCGGCGTCGC CCCCCTCTTC ACCATGGACA GCATCGACAA GCTCACGGAA CTCGCACGAC
  AH          MEM  H    NN         S  M      A                P   M
  HG          NAB  P    CL         F  M      L                L   L
  AA          LRO  H    OA         A  E      U                E   Y
  21          112  1    13         1  1      1                1   1

          430         440         450         460         470         480
TCGGCAAGAC GCTTTTGCTG ACCAAGCCCG AGAATACCCA CCGCGCATTC GAGGCGCTCC
        H         M         C A                  FH M   T  M HH      F
        G         W         V V                  NH W   A  N AH      N
        A         O         I A                  UA O   Q  L EA      U
        1         1         1 1                  21 1   1  1 21      2

          490         500         510         520         530         540
GCGTCGGAGT ACGCGCCGTA TTGTCGGCCG AAACGTCGGT CGAAGAACTG GTGCACGTCA
H        R   FH B          XH        M  TM MT  M     B     A   M
G        S   NH C          MA        A  AC MA  B     S     P   A
A        A   UA E          AE        E  QR EQ  O     R     A   E
1        1   21 1          33        2  21 11  2     1     1   2
```

Figure 4.13 continued

Restriction map of *redZ*

```
        550         560         570         580         590         600
TCAGAACCAT CACCGAAGTC AATGCCATAA TCGCTCCGGA AGAGGCACAG GAAGCCCTGA
   D   H      B          M           EH  M   M  M        C     H
   R   P      C          W           AP  B   N  W        V     G
   D   H      G          O           RA  O   L  O        I     A
   2   1      1          1           12  2   1  1        1     1

        610         620         630         640         650         660
CGCGCTATTG GCGCAGCAAG GCCCCGAAGA ACCTCCGCCC CGAACTGACC CCTCGGGAAA
FM M       H B       EN        M       M E                 BA
NW W       H B       CL        B       N C                 SV
UO O       A V       OA        O       L I                 AA
21 1       1 1       94        2       1 1                 11

        670         680         690         700         710         720
CGGAAGTCCT CCTCTTACTG ACCCAGGGCA AGACCAACAC CGAGATGGCC GCGACCCTCT
       M    M           BB          X           GHF F       M
       N    N           SS          C           DAN N       N
       L    L           AA          M           IEU U       L
       1    1           11          1           231 2       1

        730         740         750         760         770         780
CCGTCTCGCC CACCACCGTC AGGTCCCACG TACACCGGAT CCTGCGCAAA CTGGGCGCGG
 EB                    PN     BM R     H BB     MH M    B     HFF  M
 SS                    PL     SA S     P AI     SH W    S     HNN  W
 PM                    UA     AE A     A MN     TA O    R     AUU  O
 11                    14     12 1     2 11     11 1    1     121  1

        790         800         810         820         830         840
CGACCCGTGC GCAGGCCGTG GCCATCGCCT ACGAGTCGGG CCTCCTGGGC ATCTGCCCGG
       MH     HBD  BH              P     SHB M   B     S       SN
       SH     ACS  AA              L     AAG N   S     F       MC
       TA     EEA  LE              E     UEL L   A     A       AI
       11     311  13              1     631 1   1     1       11

        850         860         870         880         890         900
GTTACGGCAC TCCCGCCCGC TGAATGCCCC GGTCCGGCCG TGCCGCTTCG CGTTCCGTGC
M  B         F    FN      B     BNHR A  HXHB      F       FM         F
A  C         A    AS      S     SCPS V  PMAC      N       NW         N
E  E         U    UP      M     AIAR A  AAEE      U       UO         U
3  1         1    12      1     1122 2  2331      1       21         1

        910         920         930         940         950         960
GGCCCGGCCG GGCATCCCGC GCCACGGCCC TTTGTCAGGG CCGTGCCTGG TAGCCGGGTT
S NHXH N     MF   F FH   D B S  E          SHB       B       C N  N
A CPMA C     WO   A NH   S C A  C          AAC       S       V C  L
U IAAE I     OK   U UA   A E U  O          UEE       T       I I  A
6 1233 1     11   1 21   1 1 6  1          631       1       1 1  4

        970         980         990        1000        1010        1020
CCAGGAGTTC GCCGCTCATC TCCTCGACGA CCGTCGCCAT CGTGTCGAAC ACGGTGACCA
B           F           M M P M                  MT          BT T B P
S           N           N M S C                  MA          SS T S P
T           U           L E H R                  EQ          TP H T U
1           1           1 1 1 1                  11          21 1 1 1

        1030        1040        1050        1060
GGTCCCGAAA GGCCTCGGTT TTCATGTCCC GGGAGAATTC
N         S H B           N     SN       ET
L         T A S           L     MC       CS
A         U E A           A     AI       OP
4         1 3 1           3     11       11
```

analysis were with the protein sequence in ORF1 predicted by the FRAME analysis. Within the predicted open reading frame there is an in-frame TTA leucine codon at position 675. As previously mentioned, the distribution of TTA codons in streptomycete genes is limited to genes involved in the regulation of antibiotic production and in the resistance to or export of antibiotics. It is interesting to report that *redZ* belongs to this select group of TTA-containing streptomycete genes. It reaffirms the unusual distribution of TTA-containing genes because *redZ* represents another gene involved in antibiotic production. The possibility that *redZ* mediates the *bldA* dependence of Red biosynthesis is highlighted by the presence of a TTA codon. The TTA codon was present in all three of the Pwb mutants as shown in Figures 4.14, 4.15 , 4.16 and 4.17. In selecting for mutants which produced Red in a *bldA* independent manner, it was expected that mutations would lie within the TTA leucine codon.

It is possible to induce Red biosynthesis in *bldA* mutants under conditions of low phosphate (Guthrie and Chater, 1990). This implies that there are no TTA codons in the Red structural genes. By implication this means that *redZ* is not a structural gene, which raises the exciting possibility that *redZ* performs a regulatory role in Red biosynthesis.

Comparison of the DNA sequences of the 800bp *Eco*RI fragment of the three Pwb mutants and the wild-type Pwb^+ identified no differences. A point mutation was identified in the sequence upstream of the left hand *Eco*RI site in the Pwb-6 mutant. A transition mutation from cytosine to thymine was observed at position 62 in the Pwb-6 mutant (Figure 4.18). The mutation lies in a potential $^-35$ promoter region. The mutation represents a change which

Figure 4.14

Sequence of Pwb$^+$ DNA showing the in-frame TTA codon

The *redZ* gene contains an in-frame TTA codon. The *bldA* gene encodes a leucyl tRNA that recognises the leucine codon TTA. The Pwb mutants were selected by their ability to produce Red in a *bldA* background. It was anticipated that the *bldA* independence of the Pwb mutants was due to mutation of an in-frame TTA codon. Figures 4.14 to 4.17 indicate the presence of the TTA codon in all three Pwb mutants.

Figure 4.14

DNA Sequence of Pwb^{+}showing the in-frame TTA codon.

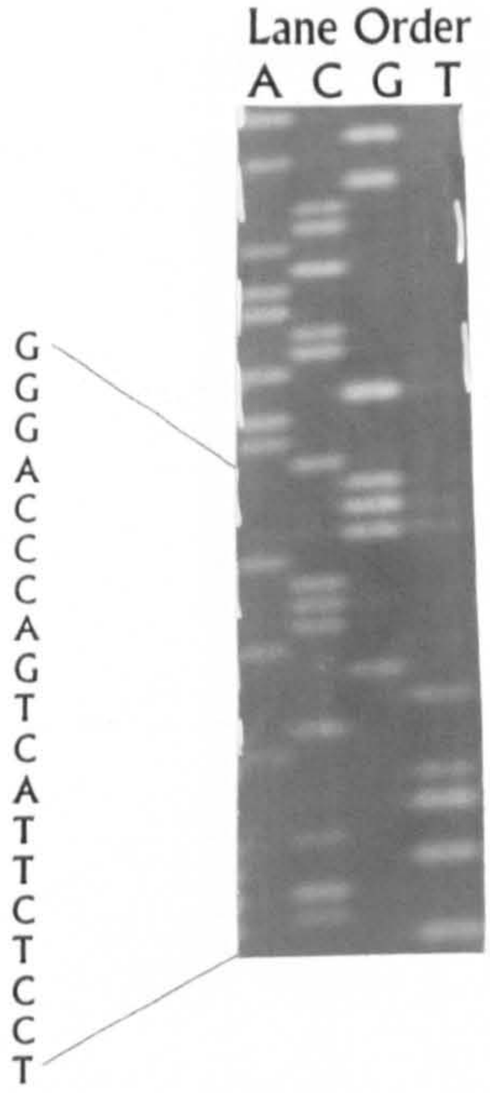


Figure 4.15

DNA Sequence of Pwb-6 DNA showing the in-frame TTA codon

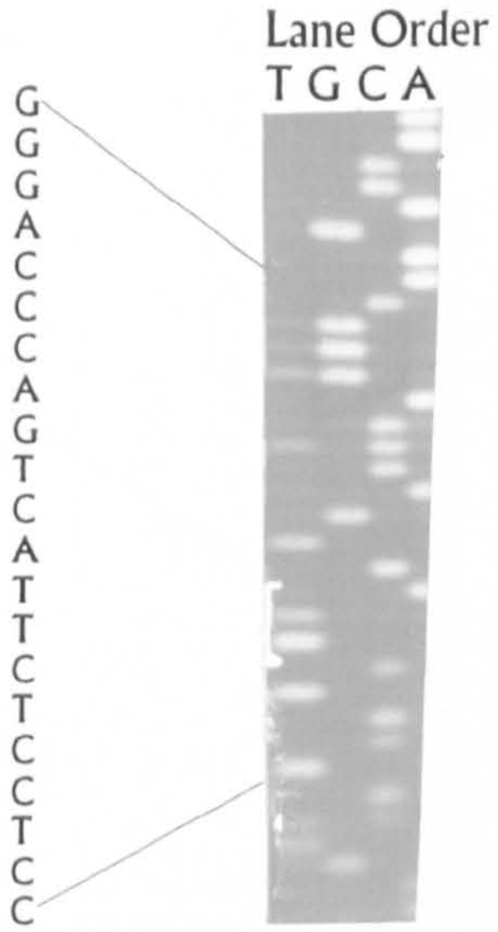


Figure 4.16

DNA Sequence of Pwb-9 DNA showing the in-frame TTA codon

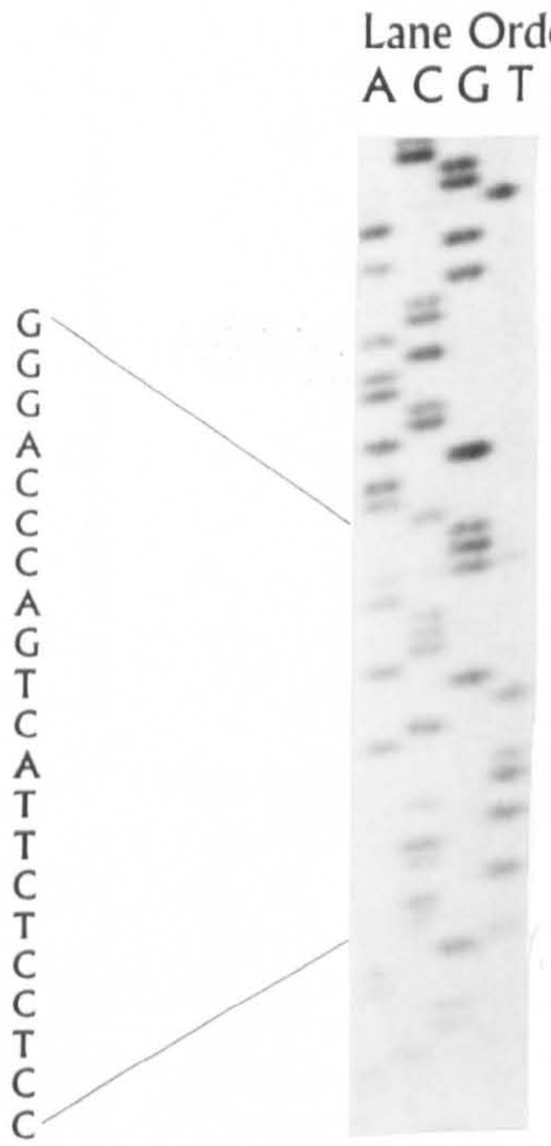
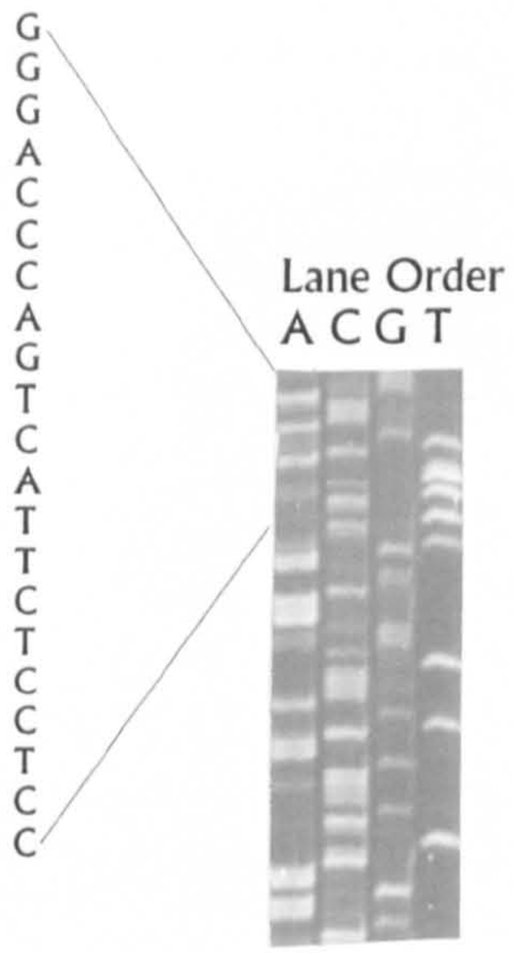


Figure 4.17

DNA Sequence of Pwb-16 DNA showing the in-frame TTA codon



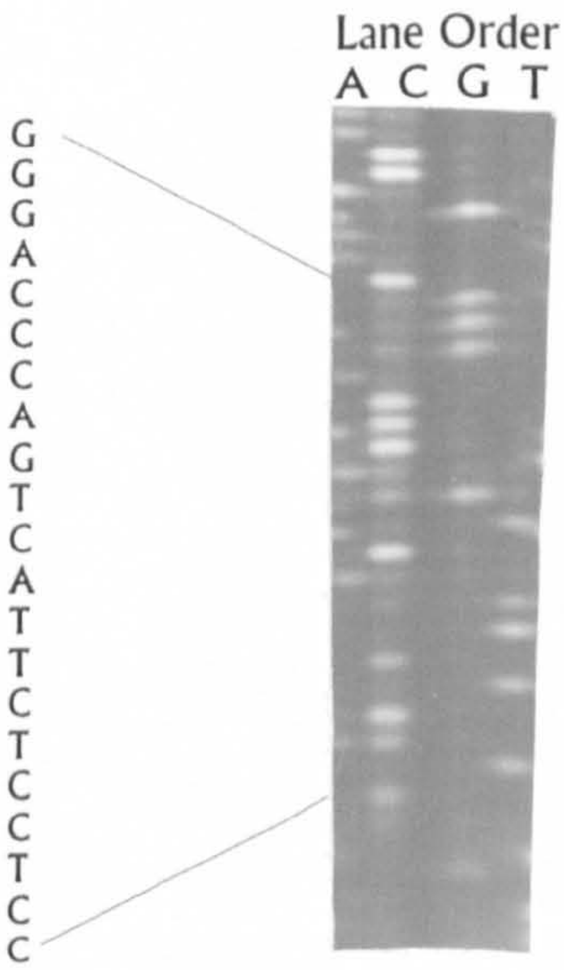


Figure 4.18

DNA Sequence showing the Pwb-6 mutation

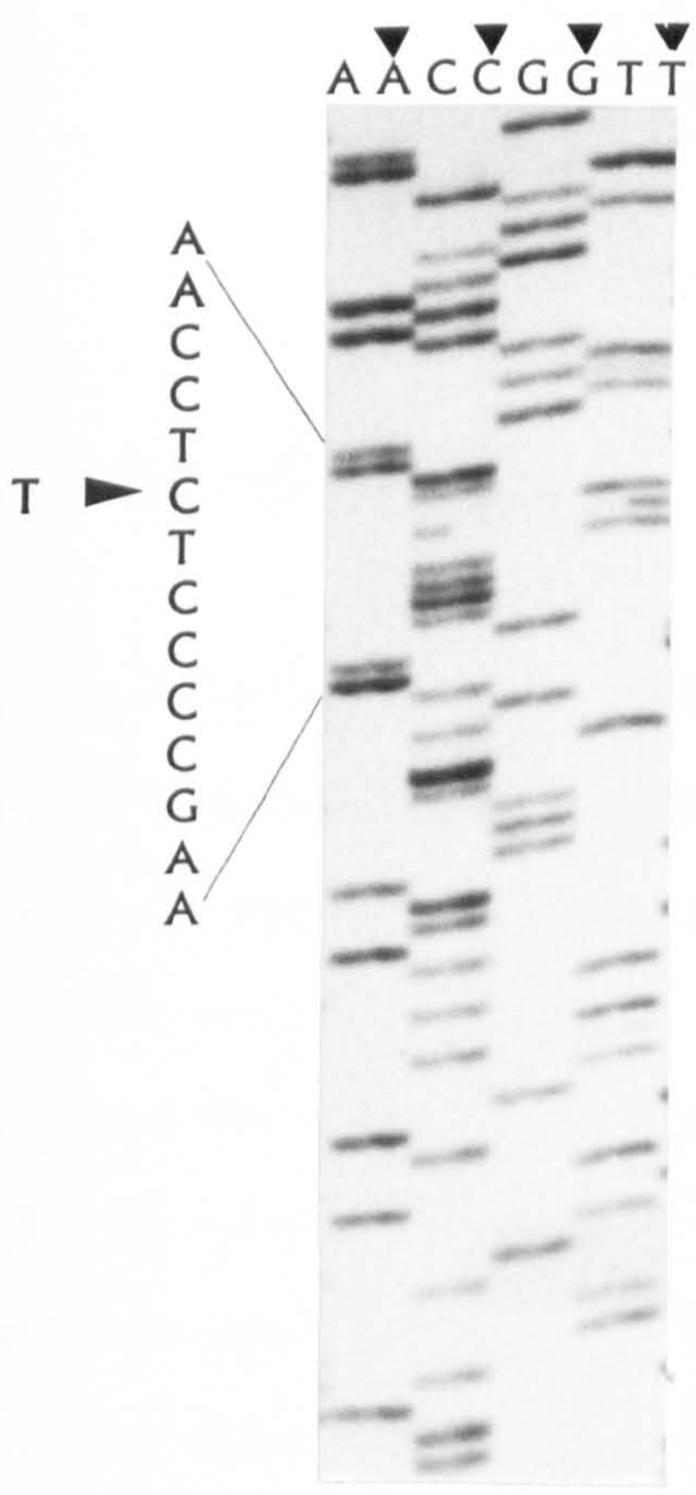
The DNA sequence on the facing page compares the Pwb^{+} and Pwb-6 DNA sequences from nucleotide 25 to 85 as shown in Figure 4.11. The mutation is a transition from a cytosine to a thymidine residue at nucleotide 63. The Pwb-6 mutation was located to a putative -35 promoter region.

```
T C T C C A  Pwb+
|       | |
T T T C C A  Pwb-6
| |     | |
T T G A C A  -35 Consensus sequence
```

(Lewin, 1990)

Figure **4.18**

DNA Sequence showing the Pwb-6 mutation



Pwb-6 DNA sequences are marked ▼
Unmarked sequences represent wildtype DNA

makes the Pwb-6 mutant -35 region more similar to the enteric bacterium major sigma factor promoter -35 consensus sequence as shown below.

```
T C T C C A  Pwb+
|       | |
T T T C C A  Pwb-6
| |     | |
T T G A C A  -35 Consensus sequence
```
(Lewin, 1990)

4.6 Similarity studies

The predicted protein sequence of *redZ* was used to search the databases. The program used to identify similar sequences was BLAST (Altschul *et al.*, 1990). This program identifies local stretches of similarity between the search sequence and those in the database. It does not necessarily give a measure of similarity over the whole length of a sequence. The advantage of searching for conserved sub-sequences is that unconserved regions do not contribute to the measure of similarity. BLAST includes a non-redundant database that derives sequences from the following sources: SwissProt, SwissProt Update, PIR, Genpept, GUpdate and PDB databases. Results of the database search revealed similarity of RedZ to the UhpA-LuxR family of transcriptional activators (Table 4.4, Figure 4.19 and Figure 4.20).

The LuxR family of regulators, in general, form one component of two component regulatory systems. There is usually a sensor kinase that is often membrane bound and a response regulator that is cytoplasmic. The LuxR family form a subclass of the response regulators. Often the sensor kinase and response regulator are located proximally on the chromosome. The sensor

Table 4.4

Similarity of RedZ to response regulators

			N-Terminal Identity		C-Terminal Identity		Overall Identity	
Gene Product	Species of origin	Function	% Identity	Region of analysis	% Identity	Region of analysis	% Identity	Product size
DegU	*Bacillus subtilis*	Secretion of degradative enzymes	19	59-135	44	146-204	20	229
GerE	*Bacillus subtilis*	Formation of spores	-	-	45	143-206	39	74
NarP	*Escherichia coli*		35	4-23	40	146-199	22	225
			20	59-112				
UvrY	*Escherichia coli*	UV damage DNA repair	24	73-205	24	73-205	20	218
BvgA	*Bordetella pertussis*	Virulence	33	1-33	-	-	23	210
			18	51-114				
VsrD	*Burkholderia solanacearum*		28	75-127	30	142-202	-	-
UhpA	*Salmonella typhimurium*	Uptake of hexose phosphates	21	36-126	45	146-191	21.5	196

Table 4.4 continued

			N-Terminal Identity		C-Terminal Identity		Overall Identity	
Gene Product	Species of origin	Function	% Identity	Region of analysis	% Identity	Region of analysis	% Identity	Product size
UhpA	*Escherichia coli*	Uptake of hexose phosphates	21	36-126	43	146-191	-	-
UvrC	*Escherichia coli*	UV damage DNA repair	24	73-205	24	73-205	-	-
ComA	*Bacillus subtilis*	Competence	20	88-156	37	146-205	-	-
CarR	*Erwinia carotovora*	Carbepenem biosynthesis	-	-	40	146-205	13	244
DnrN	*Streptomyces peucetius*	daunorubicin biosynthesis	-	-	45	146-204	25	201
AmfR	*Streptomyces griseus*	aerial mycelium formation	25	79-109	38	146-204	25	170
			33	114-143				
RamR	*Streptomyces coelicolor*	aerial mycelium formation	-	-	40	140-200	23	202
BrpA	*Streptomyces hygroscopicus*	bialophos biosynthesis	-	-	35	145-206	13.5	256

Figure 4.19

Alignment of RedZ to other regulators

The alignment of the predicted RedZ amino acid sequence to other similar sequences is shown on the facing page. Identity is marked in black, conserved changes are highlighted in grey.

Figure 4.19

Alignment of RedZ to other regulators

```
Scafsql  . . . V P S L L D I   E D D E A I R T A L   E L S L T . R Q G H   R V A T . . . A A S   G E D G L K L L R E   Q R P D L I V L D V  53
Escompr  M Q E N Y K I L V V   D D D M R L R A L L   E R Y L T . E Q G F   Q V R S . . . V A N   A E Q M D R L L T R   E S F H L M V L D L  56
 Slcutr  . . . . M R V L V V   E D E Q L L A D A V   A T G L R . R E A M   A V D V . . . V Y D   G A A A L E R I G V   N D Y D V V V L D R  52
Basdegu  . M T K V N I V T I   D D H Q L E R E G V   K R I L D F E P T F   E V V A E G . . D D   G D E A A R I V E H   Y H P D V V T M D I  57
Escuhpa  . . . M I T V A L I   D D H L I V R S G F   A Q L L G L E P D L   Q V V A E F . . G S   G R E A L A G L P G   R G V Q V C M C D I  55
 Scramr  M G E M V R I A V V   H D E K L L R S A L   V Q L L R S D D T L   D V S S H C L D A D   G P E L S A A L P .   . . A D V C V V D G  57
 Scredz  . . M T T R V L V C   C D R V I L G E G I   R A L L E . R H D M   K V Q V E . . . T T   Q R G S L A T A A E   T G P D T L V G V A  54

Scafsql  M L P G I D G F E V   C R R I R . R T D Q   L P I I L L T A R N   D D I D V V V G L E   S G A D D Y V V K P   V Q G R V L D A R T  112
Escompr  M L P G E D G L S I   C R R L R S Q S N P   M P I I M V T A K G   E E V D R I V G L E   I G A D D Y I P K P   F N P R E L L A R T  116
 Slcutr  D L P L V H G D D V   C R K I V E L G M P   T R V L M L T A S G   D V S D R V E G L E   I G A D D Y L P K P   F A F S E L I A R V  112
Basdegu  N M P N V N G V E A   T K Q L V E L Y P E   S K V I I L S I H D   D E N Y V T H A L K   T G A R G Y L L K E   M D A D T L I E A V  117
Escuhpa  S M P D I S G L E L   L S Q L P K . . . G   M A T I M L S V H D   S P A L V E Q A L N   A G A R G E L S K R   C S P D E L I A A V  112
 Scramr  E C L T G P E D A G   A G R L R A R Y G D   . R L V V L A T A K   R P G V L R R A T D   G G A L G L V D K N   A P A H R L I T A V  116
 Scredz  P L F T M D S I D K   L T E L A R L . . .   G K T L L L T K P E   N T H R A F E A L R   V G V R A V L S A E   T S V E E L V H V T  111

Scafsql  R A V L R R G E R E   S T D S A . . . . S   . . . . . F G S   L V I D R S A M T V   T K N G E D L Q L T   P T E L R L L L E L  161
Escompr  R A V L R R Q A N E   L P G A P . . . . S   Q E E A V I A F G K   F K L N L G T R E M   F R E D E P M P L T   S G E F A V L K A L  172
 Slcutr  R A L G R R T S V P   L P . . . . . .     . . . P V L E R A G   I K L D P N R R E V   F R D G K E V Q L A   P K E F A V L E V L  161
Basdegu  K V V A E G G S Y L   H P K V T H N L V N   E F R R L A T S G V   S A H P Q H E V Y P   E I R R P L H I L T   R R E C E V L Q M L  177
Escuhpa  H T V A T G G C Y L   T P D I A I K L A S   G . . . . . .       . . . . . .       . . . . R Q D P L T   K R E R Q V A E K L  149
 Scramr  H T V A R G E R F L   D E T L T V A L L K   G . . . . . .       . . . . . .       . . . . A E M P L T   T R E L G V L T L A  153
 Scredz  R T H T E V N A I I   A P E E A Q E A L T   R Y W R S K A P . .   . . . . . .       . . K N L R P E L T   P R E T E V L L L L  157

Scafsql  S R R P G Q A L S R   Q Q L L R L V W E H   D Y L G D S R L V D   A C V Q R L R A K V   E D V P S S P T L I   R T V R G V G Y R L  221
Escompr  V S H P R E P L S R   D K L M N L A R G R   E Y S A M E R S H D   V Q H S R L R R M V   E E D P A H P R Y I   Q T V W G L G Y V F  232
 Slcutr  M R S E G A V V S A   E Q L L E K A W D E   N T D P F T N V V R   V T V M T L R R K L   G E . . . . P P V I   V T V P G S G Y R I  217
Basdegu  A D G K S N R G I G   E S L . . . . . .     . . F I S E K T V K   N H V S N I L Q K M   N V N D R T Q A V V   V A I K N G W V E M  228
Escuhpa  A Q G M A V K E I A   A E L . . . . . .     . . G L S P K T V H   V H R A N L M E K L   G V S N D V E . L A   R R M F D G W * . .  196
 Scramr  S Q G A P I A E I A   A R L . . . . . .     . . H L S R G T V R   N Y M A T A V R K V   G A R N R V D A I R   I V Q S A G W T . .  202
 Scredz  T Q G K T N T E M A   A T L . . . . . .     . . S V S P T T V R   S H V H R I L R K L   G A A T R A Q A V A   I A Y E S G L L G I  208

Scafsql  D P P Q * . . . . .  225
Escompr  V P D G S K A . . .  239
 Slcutr  . . . . . .          217
Basdegu  R * . . . . . .      229
Escuhpa  . . . . . .          196
 Scramr  . . . . . .          202
 Scredz  C P G Y G T P A R *  217
```

Figure prepared by M.Bibb

Figure 4.20

Alignment of RedZ with response regulators

Alignment of RedZ with other response regulators identified by BLAST analysis is shown. Conserved residues involved in phosphorylation in other response regulators such as UhpA, DnrN and AmfR are highlighted and underlined. As can be seen RedZ contains only contains one of these residues, an aspartate at amino acid 10.

Figure 4.20

Alignment of RedZ with response regulators

```
          1                                                  50
AMFR  ...MTTVLLV HTVPLWRASL ASLLGTGRGI E.LRTAEHGA IRAALSGPSP
RAMR  MGEMVRIAVV HDEKLLRSAL VQLLRSDDTL D.VSSHCLDA DGPELSAALP
DNRN  ..MTIRVVIA EELEMVRAAL VLLVERESDI KVVASTGVGS ELIRTVLEYS
DEGU  .MTKVNIVII DDHQLFREGV KRILDFEPTF EVVAEGDDGD EAARIVEHYH
GERE  .......... .......... .......... .......... ..........
UHPA  ...MITVALI DDHLIVRSGF AQLLGLEPDL QVVAEFGSGR EALAGLPGRG
BVGA  ..MYNKVLII DDHPVLRFAV RVLMEKE.GF EVIGETDNGI DGLKIAREKI
REDZ  ..MTTRVLVC CDRVILGEGI RALLERHDMK VQVETTQRG. .SLATAAETG
      ------V-II DD--LLR--L --LLE-E--- EVV-----G- E---------

          51                                                 100
AMFR  .DVLLTDLDC PGALDVLDEV KTVTAPHGGP CPLAVLTRSD RPSGLRRAYE
RAMR  ADVCVVDGEC LTGPEDAGAG R.LRARYGDR ..LVVLATAK RPGVLRRAFD
DNRN  PDVTVIDVD. MHEREGLLSA MAIQEHLPDC RTL.LTARFA RLATVRQSLA
DEGU  PDVVIMDIN. MPNVNGVEAT KQLVELYPES KVI.ILSIHD DENYVTHALK
GERE  .......... .......... .......... .......... ..........
UHPA  VQVCICDIS. MPDISGLELL SQLPK...GM ATI.VLSVHD SPALVEQALN
BVGA  PNLVVLDIG. IPKLDGLEVI ARLQSLGLPL RVL.VLTGQP PSLFARRCLN
REDZ  PDILV...G. VAPLFTMDSI DKLTELARLG KTL.LLTKPE NTHRAFEALR
      PDV-V-DI-- M----GLE-- --L------- --L-VL---D -----R-AL-

          101                                                150
AMFR  AGALGYIDKY RPVDDLSEVM HKLADGGRHI DESLAFSLLQ V.........
RAMR  GGALGLVDKN APAHRLITAV HTVARGERFL DETLTVALLK G.........
DNRN  SGLGGLVLKS SPPAHLSDAI RSVHEGRRVF DTELTLAAWG N.........
DEGU  TGARGYLLKE MDADTLIEAV KVVAEGGSYL HPKVTHNLVN EFRRLATSGV
GERE  .......... .......... .......... .......... ..........
UHPA  AGARGFLSKR CSPDELIAAV HTVATGGCYL TPDIAVKLAA G.........
BVGA  SGAAGFVCKH ENLHEVINAA KAVMAGYTYF PSTTLSEMRM GDNAKSDSTL
REDZ  VGVRAVLSAE TSVEELVHVI RTITEVNAII APEEAQEALT RYWRSKAP..
      -GA-G-V-K- ---DELI-AV --VAEG--YL -------L-- ----------

          151                                                200
AMFR  .......... ....ADMPLS PRELSVLSMA EGGDTVAGIA GRLHLTPGTV
RAMR  .......... ....AEMPLT TRELGVLTLA SQGAPIAEIA ARLHLSRGTV
DNRN  .......... ....TGCPLT EREMEILQLA AAGDDINEIA VKLSLSPGTV
DEGU  SAHPQHEVYP EIRRPLHILT RRECEVLQML ADGKSNRGIG ESLFISEKTV
GERE  .......MKE KEFQSKPSLT KREREVFELL VQDKTTKEIA SELFISEKTV
UHPA  .......... ....RQDPLT KRERQVAEKL AQGMAVKEIA AELGLSPKTV
BVGA  .......... .....ISVLS NRELTVLQLL AQGMSNKDIA DSMFLSNKTV
REDZ  .......... ..KNLRPELT PRETEVLLLL TQGKTNTEMA ATLSVSPTTV
      ---------- -------PLT -RELEVL-LL AQG----EIA --L-LSPKTV

          201                                          241
AMFR  RNYLAAAIRK SGARNRLDAI RRAKEAGWI. .......... .
RAMR  RNYMATAVRK VGARNRVDAI RIVQSAGWT. .......... .
DNRN  RNYLATTVRK MDARNRLDAV RIASMSGWI. .......... .
DEGU  KNHVSNILQK MNVNDRTQAV VVAIKNGWVE MR........ .
GERE  RNHISNAMQK LGVKGRSQAV VELLRMGELE L......... .
UHPA  HVHRANLLEK LGVSNDVELA HRMFDGW... .......... .
BVGA  STYKTRLLQK LNATSLVELI DLAKRNNLA. .......... .
REDZ  RSHVHRILRK LGAATRAQAV AIAYESGLLG ICPGYGTPAR *
      RN--A--LRK LGA-NR-DAV -IA---GW-- ---------- -
```

kinase phosphorylates the response regulator in reply to an environmental stimulus. Phosphorylation activates the response regulator, which either by DNA interaction activates transcription of specific genes or by protein-protein interaction activates other phenomena. The two components (response regulator and sensor kinase) each have characteristic conserved domains. The response regulator has a N-terminal receiver domain which is involved in activation of the C-terminal DNA binding function, generally via phosphorylation. The C-terminal portion of the response regulators has a helix-turn-helix DNA binding motif and is responsible for transcriptional activation. Regulatory systems are often more elaborate involving other components in addition to the sensor kinase and response regulator. The homology which has been found to *redZ* (ORF1) is with the C-terminal DNA binding motif of this class of regulators (Table 4.4, Figure 4.19 and Figure 4.20).

At the N-terminus of the response regulators there are conserved residues which are believed to form an active site (Stock et al., 1989). The crystal structure of the *E.coli* response regulator CheY has been resolved. Four residues were found to be important for phosphorylation of CheY; Asp12, Asp13, Asp57 and Lys109 (Parkinson and Kofoid, 1992). The Asp12, Asp13 and Asp 57 lie closely together and form an acid pocket which appears to constitute an active site. The site of phosphoryaltion in CheY is Asp57. The conserved residues correspond to Asp8, Asp9, Asp54 and Lys102 of UhpA (Figure 4.20). The predicted sequence of RedZ contains the first aspartate residue corresponding to Asp9 of UhpA, but does not contain the other conserved residues. The absence of an aspartate residue at a position equivalent to Asp54 of UhpA in RedZ is notable. It is predicted that RedZ is not activated by phosphorylation and as such perhaps should not be

considered as a 'response' regulator. This is not unprecedented as GerE, with which RedZ has 39% identity, does not have the N-terminal receiver domain and as such would not be phosphorylated. The identity of RedZ to GerE stretches over the entire length of the GerE protein.

Overall, greatest activity of RedZ to other response regulators is within the region of amino acid residues 146-204 of RedZ. This corresponds to the C-terminal DNA binding region. Even in regulators such as the *Erwinia carotovora* CarR which has a low, 13% overall identity with RedZ, there is 40% identity with the region that encompasses the helix-turn-helix DNA binding domain. A similar situation is seen with BrpA, the activator of bialaphos biosynthesis in *Streptomyces hygroscopicus*. There is low overall identity of 13%, but 35% identity to the region of amino acid 146-206. Perhaps, it is surprising that BrpA is not more similar to RedZ; given that BrpA also lacks conserved residues involved in phosphorylation, and no cognate sensor kinase has been identified yet it is also involved in antibiotic biosynthesis in a *Streptomyces* species.

An increasing number of response regulators and their cognate sensor kinases that have been identified in streptomycetes. Five Streptomycete response regulator genes were identified as homologus to RedZ by the BLAST analysis; *amfR*, *dnrN*, *ramR*, *cutR* as well as the previously mentioned *brpA*.

CutR regulates transcription of a membrane protein involved in copper transport (Tseng and Chen, 1991). The identity of RedZ with CutR was mainly about the N-terminus. The overall identity was low at 16%.

Both AmfR and RamR are involved in aerial mycelium formation and like RedZ contain in-frame TTA codons. AmfR and RamR are believed to have analogus functions in *S.griseus* and *S.lividans* respectively (Ma *et al.*, 1994). The overall similarity of RedZ to RamR and AmfR was 23% and 25% respectively. The AmfR sequence contains the conserved aspartate residue at position 54. This has been shown to be the site of phosphorylation by site directed mutagenesis of the aspartate residue to asparagine (Ueda *et al.*, 1993). The mutated AmfR Asn54 does not induce sporulation.

DnrN is involved in regulation of daunorubicin biosynthesis in *Streptomyces peucetius*. The conserved residues involved in phosphorylation are present, with a conserved change of a glutamate for aspartate at position 10. Site directed mutagenesis of Asp55, the predicted site of phosphorylation, was performed (Otten *et al.*, 1995). The Asp55 was replaced with a glutamate residue and an asparagine residue. Production of daunorubicin (DNR) and an intermediate of DNR, ε-rhodomycinone (RHO) were assayed in a DNR and RHO non-producing *dnrN*⁻strain that carried, on plasmids, either the mutant DnrN Asp55Asn or mutant DnrN Asp55Glu. Suprisingly significant levels of DNR and RHO were detected, 35% for Asp55Asn and 53% for Asp55Glu compared to wild-type levels. Other regulators such as AmfR on substitution of the equivalent aspartate residue resulted in a loss of function, the significance of this is uncertain (Ueda *et al.*, 1993).

4.7 Discussion

The nucleotide sequence and predicted amino acid sequence of *redZ*, a gene present in the Red cluster is reported. The function of RedZ by comparison to DNA sequences in the databases is as an activator of transcription. It is

proposed that RedZ is involved in the regulation of Red biosynthesis. The *redZ* DNA sequence contains an in-frame TTA codon. This suggests that RedZ is responsible for mediating the *bldA* dependence of Red biosynthesis.

RedZ is homologus to members of the UhpA family of response regulators. The RedZ amino acid sequence only contains one of the conserved residues involved in phosphorylation in other response regulators. It is predicted that RedZ is not activated by phosphorylation. It remains to be determined if RedZ requires activation by another means to be fully functional.

The Pwb-6 mutation has been identified as a point mutation within a putative $^{-}$35 promoter region. The mutation clearly does not lie within the coding region. It is anticipated that the Pwb-9 and Pwb-16 mutations will be located in the N-terminal region upstream of the first *Eco*RI site. DNA sequence analysis of the region downstream of the first *Eco*RI site has shown all three Pwb mutants are similar to wild-type. In particular, all three mutants contain an in-frame TTA codon. Mutations within the TTA codon might have been expected to cause *bldA* independent Red production. However this is clearly not the case here. The *bldA* independent Red production of the Pwb mutants might be the result of increased expression of the *redZ* gene. The Pwb-6 mutation makes the predicted $^{-}$35 promoter closer to the consensus sequence. It is anticipated that increased *redZ* expression would result in the Pwb phenotype.

In other two component systems the presence of the response regulator on a multicopy plasmid leads to constitutive high level expression of the target gene. In the case of RedZ, it is hypothesised that increased transcription in the Pwb-6 up-promoter mutant resulted in increased Red production by

comparison to wild-type. Presumably, with increased *redZ* transcripts present, natural supression, i.e. translation by an alternative tRNA to the TTA $tRNA_{leu,}$ occurs at a high enough level to allow RedZ translation and hence Red biosynthesis, assuming *redZ* is essential for Red formation. High levels of spontaneous supression have been identified in streptomycetes (P.Mazodier and J.Davies, personal communication.)

CHAPTER 5

CHARACTERISATION OF RedZ

5.1 Introduction

In the previous chapter the nucleotide sequence of *redZ* and the predicted amino acid sequence of its product RedZ were reported. One of the Pwb mutations that cause Red production in a *bldA* background was identified. The *pwb-6* mutation was shown to lie within a potential -35 promoter region of *redZ*. The mutation increased the similarity of the predicted -35 promoter region of *redZ* to the enteric bacteria major sigma factor -35 promoter region (Figure 4.17).

Homology of the predicted RedZ protein to response regulators of the UhpA-LuxR family was found. Similarity of the predicted *redZ* product to streptomycete regulators RamR, AmfR and DnrN was identified. In addition to the similarity between DnrN and RedZ, similarity between DnrI and RedD has been identified (Stutzman-Engwall *et al.*, 1992).

The *dnrN* and *dnrI* genes were isolated independently on two DNA fragments (dnr_1 and dnr_2) of *Streptomyces peucetius* (Stutzman-Engwall *et al.*, 1992). Both DNA fragments dnr_1 and dnr_2 caused over-production of daunorubicin (DNR) in wild-type strains. In addition, dnr_2 restores DNR resistance to a daunorubicin sensitive strain. Further characterisation of dnr_1 identified the genes *dnrI* and *dnrJ* (Stutzman-Engwall *et al.*, 1992). The genes *dnrN* and *dnrO* were shown to be present on the dnr_2 DNA fragment (Otten *et al.*, 1995). The over-production of DNR by dnr_1 and dnr_2 was found to be conferred by *dnrI* and *dnrN* respectively (Stutzman-Engwall *et al.*, 1992; Otten *et al.*, 1995). The predicted amino acid sequence of *dnrI* shows high similarity to ActIIORF4 and in addition similarity to AfsR and RedD.

Actinorhodin production was restored in an *actII* mutant of *S.coelicolor* by transforming the strain with multicopy *dnrI*.

Given the similarity of *redZ* to *dnrN* and *redD* to *dnrI*, it raises the question of whether the function of the two *red* genes are analogous to the regulatory genes *dnrN* and *dnrI* of daunorubicin biosynthesis. When present in the cell in multicopy *redD* causes massive over-production of Red (Narva and Feitelson, 1990). It is speculated that multicopy $redZ^{+}$ would cause over-production of Red in *S.coelicolor*, in a similar way that *dnrN* in multicopy gives rise to high levels of daunorubicin in *S.peucetius*.

In the absence of the *redZ* gene cloned onto a multicopy vector an alternative approach using clones and a *redZ* mutant strain already available was taken. DNA sequence data obtained of the Pwb-9 *redZ* region does not encompass the entire gene, sequence has not been obtained of the N-terminus. It is anticipated that the Pwb-9 mutation will lie within the promoter region and like the Pwb-6 mutation represent a promoter up mutation. The Pwb-9 strain and the *redZ* gene of Pwb-9 were available. The presence of increased *redZ* transcription due to the presence of the $redZ^{p}$ gene is anticipated. The notation $redZ^{p}$ is used here to designate a *redZ* gene originating from a Pwb mutant so that it can be distinguished from the wild-type *redZ*.

5.2 Isolation of a *redD* lysogen in the Pwb-9 mutant EG111

Inactivation of *redD* by insertional mutagenesis in the Pwb-9 mutant was proposed as another method of determining whether, in the absence of a functional *redD* gene, $redZ^{p}$ can activate Red biosynthesis. This was performed concurrently with the previous experiment. The *redD* disruption

phage ØKC899 was already available (J.White). Lysogens of Pwb-9 with ØKC899 were isolated as previously described (2.4.5) using selection for thiostrepton resistance. Seven ØKC899 lysogens were isolated and all produced Red. From Figure 5.1 it can be seen that two of the lysogens are yellow-orange. On further incubation the two yellow-orange lysogens produced Red (Figure 5.1). The observation was made that the !*redD*i *redZ*^{p} lysogens produced Red at a slower rate than wild-type strains. As can be seen in Figure 5.2 the red pigment produced by the !redDi disrupted lysogens appears more orange by comparison to the pigment produced by the Pwb-9 strain. When streaked to single colonies a number of Red^{-} and sectored colonies were observed (Figure 5.3). On further incubation some of the Red^{-} colonies slowly produced red pigment. Not all Red^{-} colonies and sectors of colonies produced Red on further incubation, a minority remained Red^{-}. Production of Red by a !*redD*i lysogen would suggest that *redZ* is epistatic to *redD*.

The phage ØKC899 contains 700bp of the *redD* region including the translational start site (J.White personal communication). When ØKC899 lysogens were sprayed with catechol a small amount of yellow, 2-hydroxymuconic semialdehyde was seen. This might indicate a small amount of XylE activity. As previously mentioned the reliability of *xylE* as a reporter gene has been questioned (3.3.4). In the phage ØKC902 (!*redX*i) XylE activity was reported (Guthrie and Chater, 1990). Although later sequence data revealed that the DNA insert of this construct did not fuse the transcription of *redX* to *xylE*. Lysogens of ØKC899 were reported to be Red^{-} despite the presence of the *redD* transcriptional start site on the phage (J.White personal communication).

Figure 5.1

Pwb-9 lysogens of ØKC899 (!*redD*i)

Lysogens of Pwb-9 strain were lysogenised with ØKC899 to cause inactivation of *redD*. If $redZ^p$ is epistatic to *redD* then Red biosynthesis in the Pwb-9 !redDi lysogen is not expected. Seven Pwb-9 !redDi lysogens were isolated these are shown on the facing page. As can be seen five of the lysogens are producing red pigment. The two yellow-orange patches did produce red pigment on further incubation.

Figure 5.1

Pwb-9 lysogens of ∅KC899 (!*redD*i)

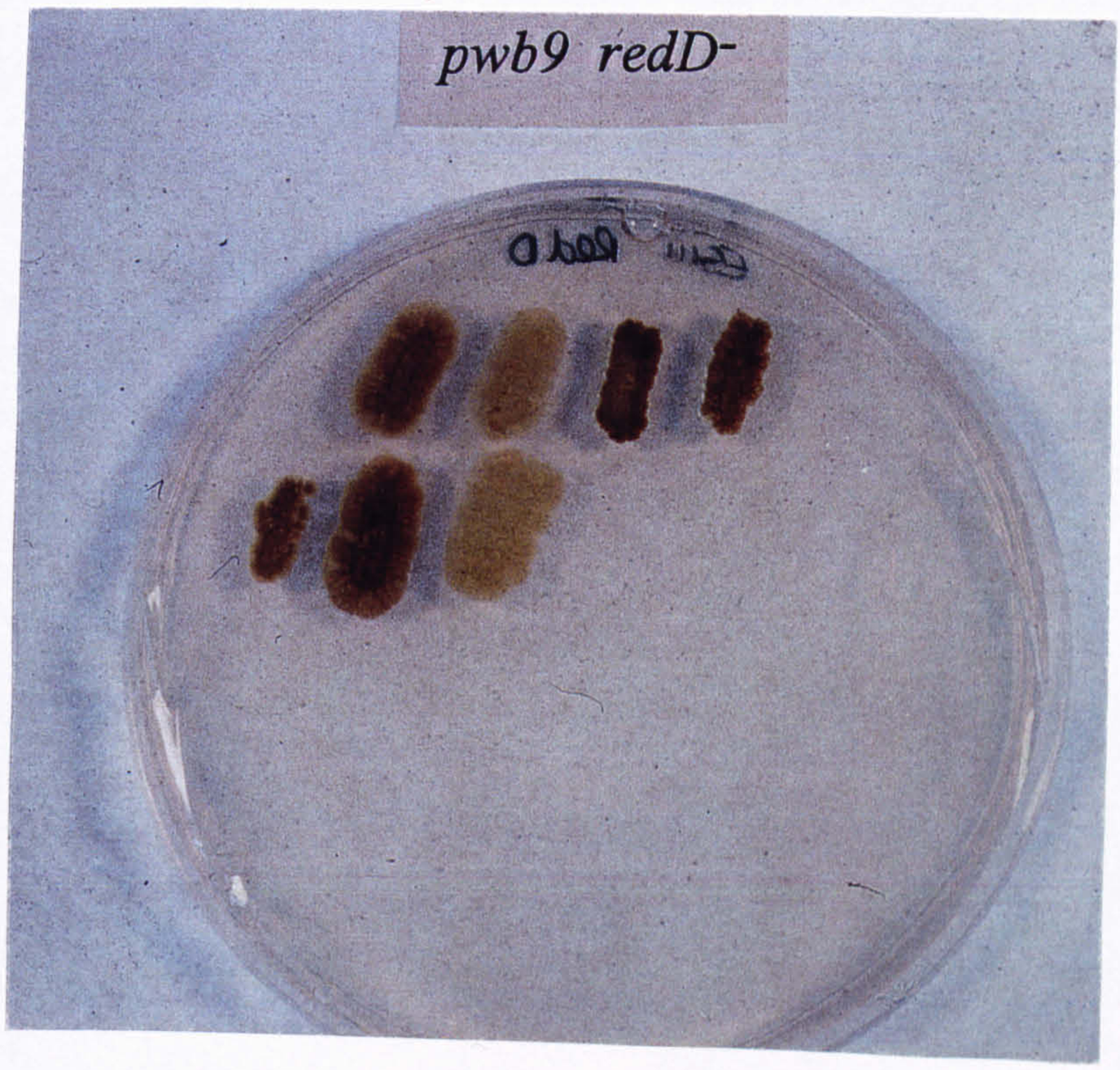


Figure 5.2

Pwb-9 ØKC899 lysogen and progenitor strains Pwb-9 and J1700

The J1700 bldA39 is the parent strain of Pwb-9. Inactivation of *redD* in the Pwb-9 results in red pigment production. As can be seen the red pigment appears more orange than that of the Pwb-9 strain.

Figure 5.2

Pwb-9 ØKC899 lysogen and progenitor strains Pwb-9 and J1700

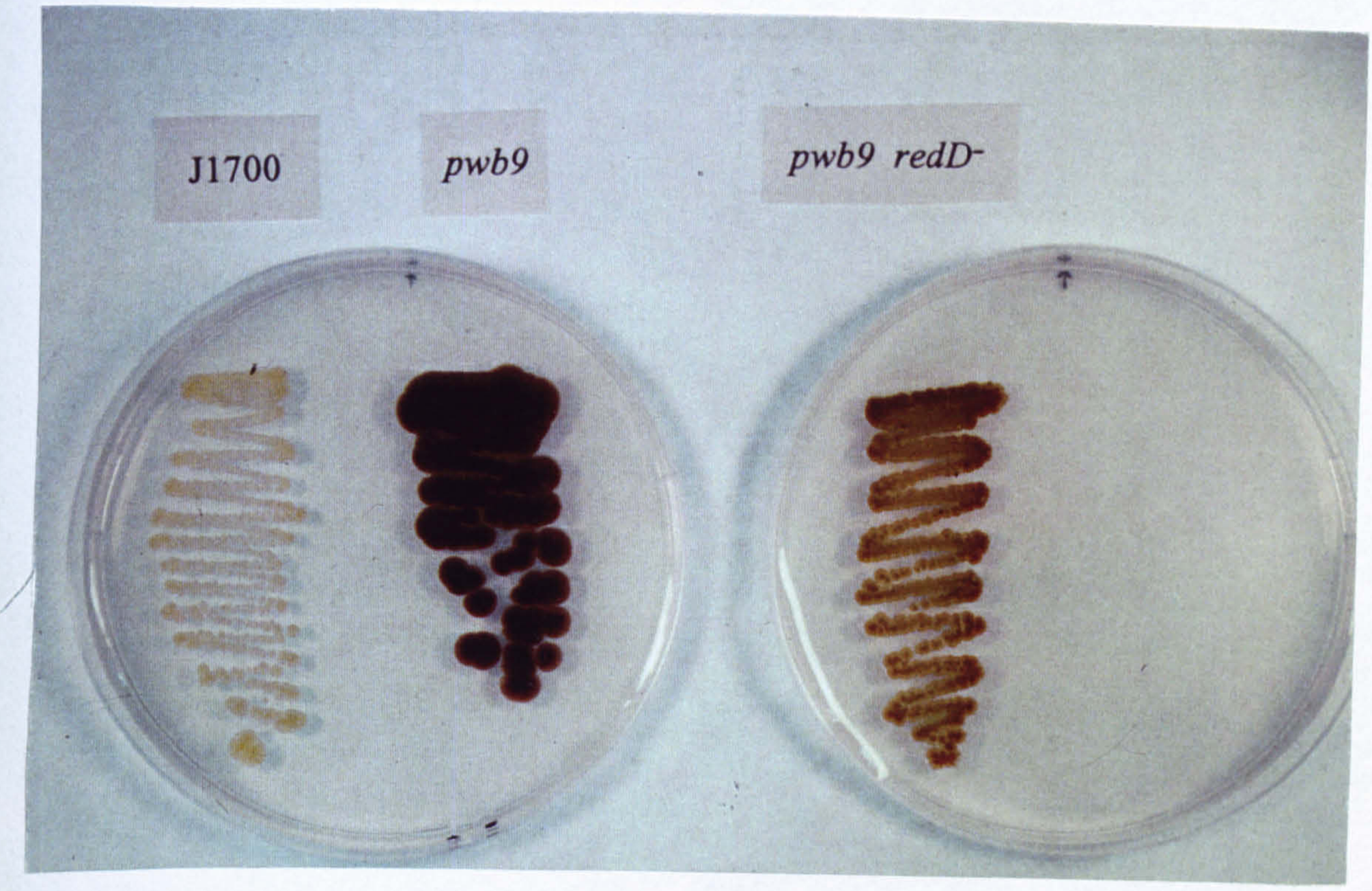


Figure 5.3

Pwb-9 ØKC899 lysogens produce sectored and white colonies

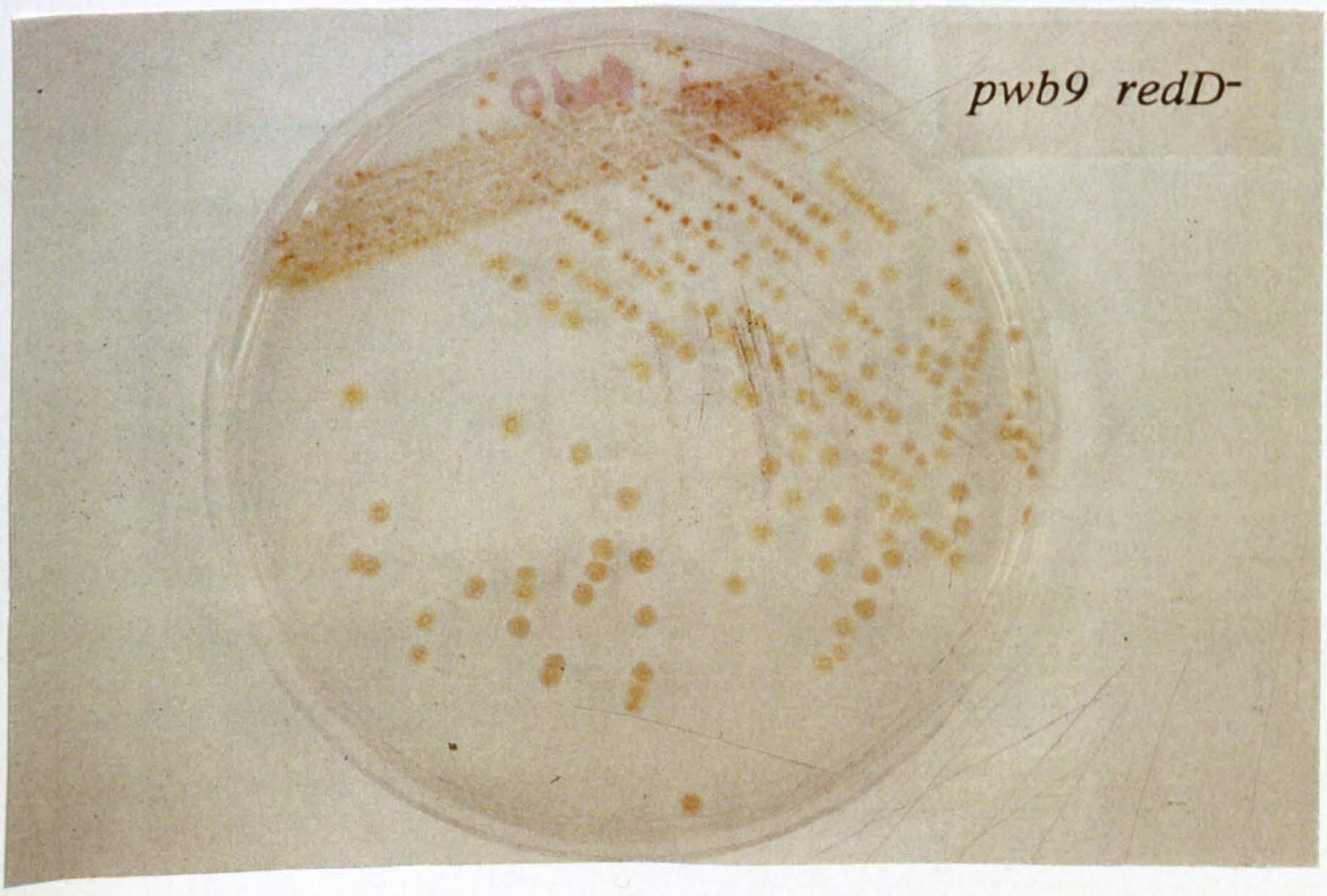
Photographs A and B show white sectors of the Pwb-9 strain inoculated onto an agar plate. Photograph C shows the Pwb-9 strain inoculated onto an agar plate, the white sectors of colonies are not discernable. However it shows the appearance of the colonies, colonies of the progenitor Pwb-9 are red at the same stage of growth as is shown here. Most of the colonies on further incubation gradually became more pigmented, some colonies and sectors of colonies remained white.

Figure 5.3

Pwb-9 ∅KC899 lysogens produce sectored and white colonies

a

b



c

Further discussion of the !*redD*i *redZ*^{p} lysogens assumes that disruption of *redD* occurred in the lysogens. Although it is recognised that Southern hybridisation analysis of the site of integration of the phage into the chromosome is required to prove this assumption.

The occurrence of white sectors in colonies of the lysogens could suggest that a second mutation to abolish Red production is occurring. In the Red producing colonies it was observed that Red accumulation was proceeding at an apparently slower rate than in wild-type strains and certainly was slower compared to the Pwb-9 strain. These observations may also suggest that the production of Red in a *redD*$^{-}$*redZ*^{p} strain is unfavourable. The production of Red appears to cause a metabolic load on the cells or causes selection pressure to abolish Red biosynthesis. Such selection pressure would arise if *redZ* can activate Red biosynthesis but not resistance to Red. In this situation it is implied that a functional *redD* gene is required for the activation of resistance genes.

It was also observed that the red pigment produced in the *redD*$^{-}$*redZ*^{p} appeared more orange than that of the parent strain Pwb-9. One possibility is that not all the Red biosynthetic genes are being activated in the absence of *redD*. The Red pigment produced by *S.coelicolor* A3(2) is made up of a mixture of prodigionines. Undecylprodigiosin is the major constituent with butylcycloheptylprodiginine (BCHP) comprising the next most abundant pigment (Tsao *et al.*, 1974). Undecylprodigiosin is more orange in appearance, while BCHP is pink. It is possible that only undecylprodigiosin biosynthesis is being activated and in the absence of *redD* BCHP production is not induced. Study of *redE* mutants has shown they lack O-methyltranferase activity which catalyses the final methylation step of Red

biosynthesis (Feitelson and Hopwood, 1983). The intermediate produced by *redE* mutants is the yellow pigment norprodigiosin. An increase in the amount of norprodigiosin may have the effect of making the mixture of pigments produced appear more orange. Extraction of the pigments and thin layer chromatography would clarify whether all the pigments that comprise Red were being produced. The appearance of pigment on agar plates can vary when cells are grown on different media. In Figure 5.2 the strains J1700 and Pwb-9 were grown on R5 and the lysogen was grown on R5 supplemented with thiostrepton, it is possible that this accounts for the difference in appearance between Pwb-9 and the !*redD*i*redZ*^{p} lysogen. Quantifying the levels of production of red pigment by the Pwb-9 mutant and by the !redDi*redZ*^{p} lysogens would clarify if the difference in appearance is due to different levels of Red production. This could be acheived by spectrophotometric analysis of pigments extracted from the two sets of cells.

5.3 Isolation of a Pwb lysogen in a *redD* mutant

If *redZ*^{p} is dominant to *redD*, in the absence of a functional *redD* gene activation of Red biosynthesis is expected in a Pwb mutant background. The method of testing this hypothesis was to introduce an extra copy of the *redZ*^{p} DNA into the *pgl*$^{+}$*redD*$^{-}$ mutant JF1. Phages containing Pwb-9 *redZ*^{p} DNA were kindly made available by K.Chater. The phages Ø2 and Ø3 contain 2.4 and 1.7 kb DNA inserts of Pwb-9 chromosomal DNA in the ØC31 derivative ØKC861 (Figure 5.4). It is predicted that lysogens of Ø2 and Ø3 containing *redZ*^{p} will have greater expression of *redZ* than the wild-type. In this way the effect of extra expression of *redZ* could be assessed in a *redD* mutant. Lysogens were selected using thiostrepton as previously described (2.4.5).

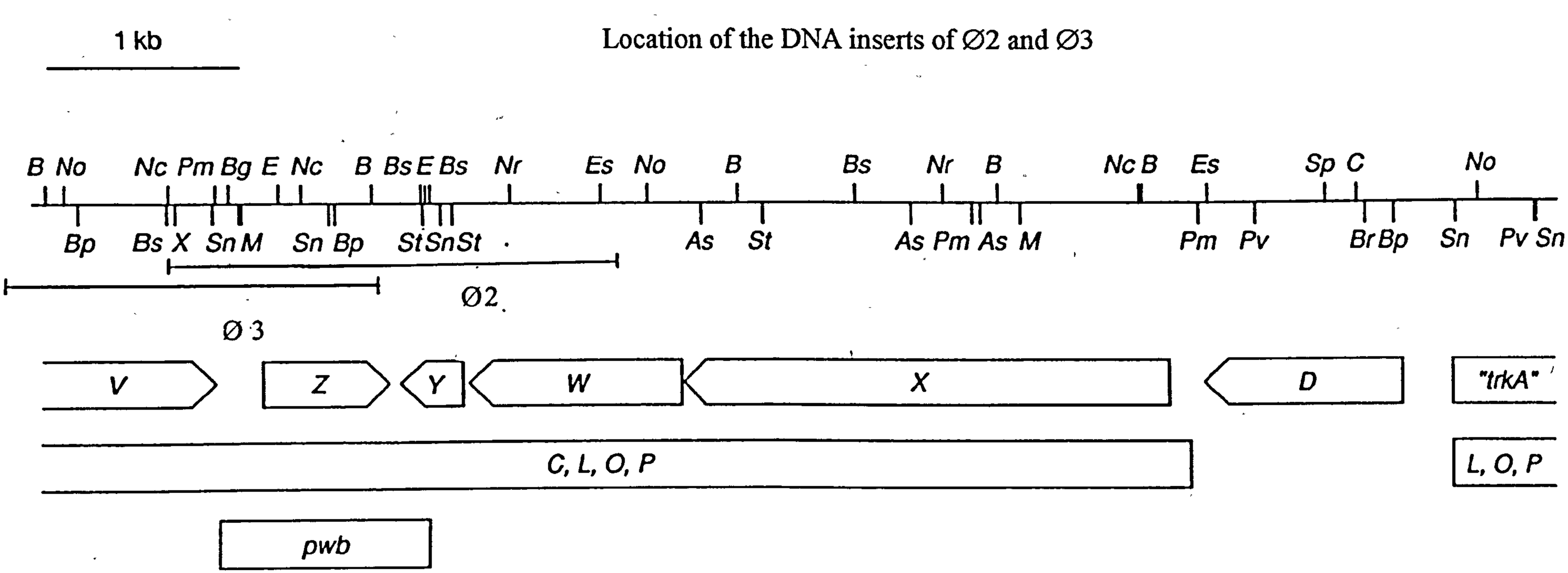
Figure 5.4

The location of the DNA inserts of phages Ø2 and Ø3

The DNA inserts of phages Ø2 and Ø3 are 2.4 and 1.7kb respectively. The phages Ø2 and Ø3 were constructed by inserting chromosomal DNA from the Pwb-9 mutant into the ØC31 derivative ØKC861. This figure shows that the DNA inserted into Ø2 and Ø3 is of the *redZ* region. The DNA insert of Ø3 contains a truncated *redZ*, the *redZ* reading frame continues 44bp after the *Bam*HI site.

Figure 5.4

Location of the DNA inserts of Ø2 and Ø3



As, AscI; B, BamHI; Bg, BglII; Bp, BspEI; Br, BsrGI; Bs, BstEII; C, ClaI; E, EcoRI; Es, EspI; M, MluI; Nc, NcoI; No, NotI; Nr, NruI; Pm, PmlI; Pv, PvuII; Sn, SnoI; Sp, SphI; St, StuI; X, XhoI

Reproduced with permission M.Bibb

One lysogen was isolated with ∅2 and did not produce Red. Lysogens of ∅3 were not isolated. The isolation of only one lysogen is believed to be a reflection of the *pgl*$^+$ status of JF1. It would have been advantageous to have isolated more lysogens, to repeatedly show that the phenotype of *redD*$^-$*redZ*^{p} lysogens was Red$^-$. Southern hybridisation analysis is now required to confirm the site of integration of ∅2 into the chromosome. Another confirmation of this result would be to introduce multicopy *redZ* into JF1 and identify whether Red production is activated.

The localisation of the Pwb-9 mutation and evidence of increased transcription in the *redZ*^{p} as compared to *redZ* has not yet been demonstrated. While acknowledging these reservations, the observation remains that Red biosynthesis was absent in a *redD*$^-$ mutant lysogenised with *redZ*^{p}. This result suggests that activation of Red biosynthesis by *redZ* requires a functional *redD*. Assuming that *redZ* does have a regulatory role, it might be anticipated from this result that *redZ* exerts its regulatory effect on Red biosynthesis via *redD*.

The absence of Red biosynthesis may indicate that *redZ*^{p} is not epistatic to *redD*. An alternative interpretation is that, as in the previous experiment there may be a selection pressure for *redZ*^{p} *redD*$^-$ cells not to produce Red. The lysogen isolated may represent a colony that had already undergone a second site mutation to prevent Red biosynthesis.

This is in conflict with the result of the previous experiment, which showed that Red was not produced by a strain which was *redD*$^-$*redZ*^{p}. In the previous experiment, that *redD*$^-$*redZ*^{p} strain was constructed in a different way, by introducing *redZ*^{p} on a ∅C31 derived phage into a redD$^-$ strain JF1. In order

to resolve the differences in results of the two experiments, it is important to determine the sites of integration of ØKC899 into the Pwb-9 mutant chromosome and Ø2 into the chromosome of redD⁻ mutant JF1. This would confirm that the results observed are due to insertion of the phage into the chromosome in the predicted location.

5.4 Interaction of *redZ* with other regulatory genes

There are a number of genes within *S.coelicolor* A3(2) that have been identified that play a role within the regulatory cascade of secondary metabolism. Regulation of antibiotic biosynthesis occurs at different levels, as identified by the study of various mutations. Mutants which globally block production of all four antibiotics have been isolated such as *absA* mutants (Adamidis *et al.*, 1990). Other mutants have been isolated that abolish production of one or more of the four antibiotics but have reduced or wild type amounts of the other antibiotics. An example of this type of mutant is *afsB* mutants which do not produce Act or Red, but produce normal amounts of CDA and methylenomycin (Horinouchi *et al.*, 1986). A pleiotropic phenotype is also demonstrated by *abaA* mutants, which show loss of Act biosynthesis, almost complete loss of Red production but normal levels of CDA and methylenomycin (Fernandez-Moreno *et al.*, 1992). Some of the mutants under study show different phenotypes depending on the growth medium, for example, *bldA* mutants produce Red when grown on medium with low phosphate concentrations (Guthrie and Chater, 1990). In the case of *absB* mutants a small amount of Red pigment is seen when grown on R5 complex medium (Adamidis and Champness, 1992). In addition to a regulatory function of antibiotic biosynthesis, some mutants isolated have pleiotropic phenotypes that effect A-factor production (e.g. *afsB*) or

sporulation (e.g. *bldA*). It is unclear how all the genes or regulators identified interact to co-ordinate antibiotic production. Some of the genes identified may not be involved in regulation of antibiotic biosynthesis directly, but exert an effect on antibiotic biosynthesis through, for instance, a metabolic disorder disrupting precursor syntheis. A DNA fragment was identified that caused stimulation of Act, Red and A-factor production in *S.lividans*. Two genes were identified from the DNA fragment *afsQ1* and *afsQ2* (Ishizuka *et al.*, 1992). The *afsQ1* gene product showed similarity to the response regulator of two-component signal transduction systems. The *afsQ2* gene product showed similarity to the sensory histidine kinases of two component systems. When *afsQ1* is transformed into *absA* mutants of *S.coelicolor* on a low copy number vector, Act production was observed. However, when *afsQ1* or *afsQ2* genes are disrupted there is no observable effect on antibiotic biosynthesis (Ishizuka et al., 1992).

The *absA* gene product has been shown to be similar to the sensory protein kinases of two-component regulatory sytems. Located closely to *absA* is a gene whose product shows similarity to response regulators of two-component systems. It is possible that cross-talk between response regulator and sensor kinase components of different regulatory systems could occur as has been shown in *E.coli*. This could account for the induction of Act production by *afsQ1* in *S.coelicolor*, as disruption of *afsQ1* has no apparent effect on Act biosynthesis. What is clear, is that there are pathway-specific activator genes like *redD* and *actII*ORF4 which are required for the activation of the antibiotic structural genes.

It is of interest to this study to try and identify what role *redZ* plays in the cascade of regulation of secondary metabolism. Attention was drawn to the

absA locus because DNA sequence identified homology of AbsA to DegS, a sensor kinase of a two-component regulatory system (W.Champness personal communication). This raises the possibility that AbsA could have some interaction with RedZ, a DegU homologue. In *Bacillus subtilis* DegS is the sensor kinase that is believed to phosphorylate DegU (Msadek *et al.*, 1990). The AbsA$^-$ and AbsB$^-$ mutants were isolated by their ability to globally block biosynthesis of all four antibiotics yet still sporulate normally. AbsA and AbsB are located at a distinct locus from the antibiotic clusters of Red, Act and methylenomycin. This implies that AbsA and AbsB are involved in regulation and are not structural components of antibiotic biosynthesis. Multiple copies of *actII*ORF4 can bypass the AbsB block to activate Act biosynthesis. This suggests that AbsB acts at a level of regulation upstream of *actII*ORF4 and may be responsible for activating *actII*ORF4 directly (Adamidis and Champness, 1992). For this reason the interaction of *absB* with *redZ*, a putative pathway regulator, was of interest.

5.4.1 Isolation of Pwb lysogens of *absA* and *absB* mutants

An extra copy of the *redZ*^{p} allele was introduced via a ØC31 derived phage into the *absA* and *absB* mutants, C577 and C120 respectively. The mutants C577 and C120 were received from J.White with permission of W.Champness. Lysogens of *absA* and *absB* mutants of Ø2 and Ø3 were selected using thiostrepton. Ø2 and Ø3 lysogens of J1501, the progenitor strain of C577 and C120, were isolated. Lysogens of *absA* and *absB* with the phages Ø2 but not Ø3 were isolated. Red production was not seen in either the *absA redZ*^{p} or the *absB redZ*^{p} lysogens (Figure 5.5).

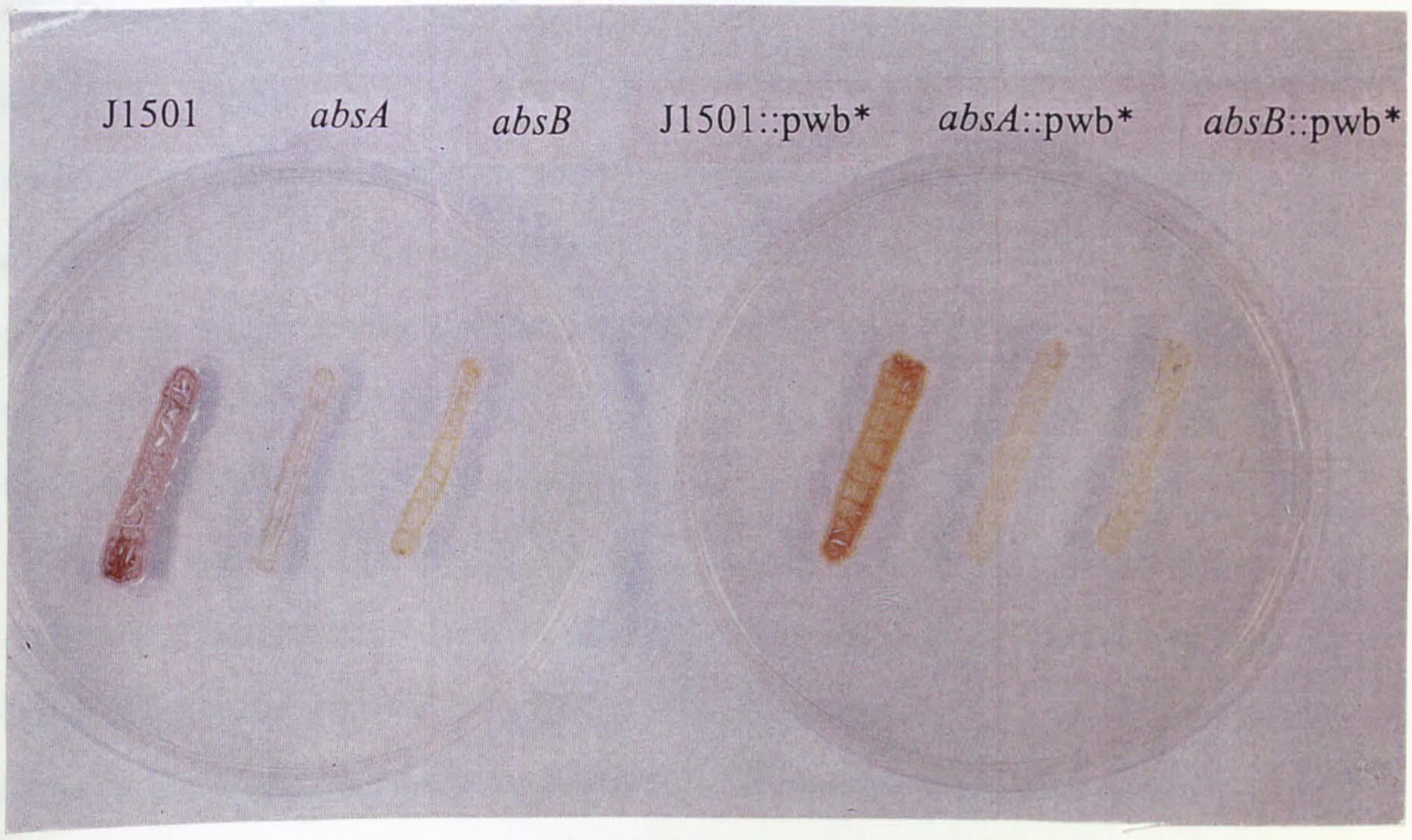
Figure 5.5

Lysogens of AbsA mutants and AbsB mutants with Ø2

The agar plate on the left of the photograph shows the strains J1501, C577 (*absA*) and C120 (*absB*). The agar plate on the right shows the same strains after lysogeny with Ø2. This is indicated as pwb*, to represent the mutant Pwb-9 *redZ* gene previously refered to as $redZ^p$. When a copy of the $redZ^p$ gene is introduced.into the AbsA and AbsB strains Red production is not seen.

Figure 5.5

Lysogens of AbsA mutants and AbsB mutants with Ø2



The absence of Red production in *absA* and *absB* mutants with a copy of *redZ*^{p} implies that there is no protein-protein interaction between *redZ* and *absA* or *absB* to induce Red biosynthesis. It is unlikely that *redZ* is epistatic to *absA* and *absB*, because *redZ* has only been shown to induce Red biosynthesis and does not effect Act, CDA or methylenomycin synthesis. The prospect that there is another level of regulation required to link *redZ* activation to *absA* or *absB* or other global antibiotic regulator, would explain why extra *redZ* transcripts had no effect on Red biosynthesis in *absA* and *absB* mutants. From the predicted amino acid sequence, RedZ does not contain the highly conserved residues involved in phosphorylation in other response regulators (Stock *et al.*, 1989). DNA sequence analysis of the region surrounding *absA* has identified a reponse regulator component which is a strong candidate for the response regulator which would interact with *absA* (W.Champness personal communication). The response regulator and sensor kinase components are often located proximally on the chromosome. It is perhaps not surprising that *redZ*^{p} did not bypass the *absA* mutation, because RedZ lacks the conserved residues phosphorylated in other response regulators. It is also less likely that RedZ and AbsA would interact directly now that a possible response regulator has been found in close proximity to *absA*.

5.5 Discussion

The effect of a predicted promoter-up mutation on *redD*, *absA* and *absB* was investigated. Conflicting results of the effect of *redZ*^{p} in a *redD*$^{-}$ background were obtained. In a *redD*$^{-}$ strain lysogenised with phage containing *redZ*^{p}, no Red biosynthesis was observed. When a Pwb-9 mutant strain was lysogenised with *redD* disruption phage some Red biosynthesis was

seen. The lysogens of the latter experiment also showed an unusual phenotype, Red biosynthesis proceeded at a slower rate than in the parent or wild-type strains and sectored white and red colonies were observed. This might suggest that the biosynthesis of Red is unfavourable in these mutants and may cause a selection pressure to either abolish or reduce Red production. This selection may have already occurred in the single JF1 ∅2 lysogen. It should also be noted that the JF1 ∅2 lysogen retains a wild-type copy of *redZ* in addition to the *redZ*^{p} allele.

A copy of *redZ*^{p} was introduced into the *absA* and *absB* mutants. Red biosynthesis was not observed. AbsA has been identified by DNA sequence homology to be similar to sensor kinase elements of two component regulatory systems. It was proposed that *redZ*, a response regulator homologue, might interact with absA. Red biosynthesis was not observed when *redZ*^{p} was introduced into an *absA* mutant. RedZ lacks the conserved residues involved in phosphorylation in other response regulators and as such might not be expected to interact with a kinase. In addition, DNA sequence analysis has revealed a response regulator homologue in close proximity to *absA* on the chromosome which it is expected to interact with AbsA.

It is of interest to understand where *redZ* lies in the regulatory cascade of Red biosynthesis, assuming that *redZ* does have a regulatory function. By examining whether high or low copy number *redZ* introduced into various regulatory mutants on plasmid vectors and determining Red biosynthesis, it would show whether *redZ* can by pass the regulatory block of the different mutants. To show that *redZ* has an essential role in activating Red biosynthesis, it would be beneficial to construct a *redZ* null mutant. It is anticipated, given the homology of *redZ* to other transcriptional activators,

that *redZ* would have a similar function. If Red biosynthesis was observed in the *redZ* null mutation it would show that *redZ* does not have an essential role in the positive activation of Red biosynthesis.

CHAPTER 6

CONCLUDING REMARKS

6 Concluding remarks and Future work

The overall aim of this study was to investigate the control of undecylprodigiosin biosynthesis in *Streptomyces coelicolor*. Three main approaches were taken; to study the over-production of Red in proline transport mutants, to characterise the Pwb mutations and to try complementing *S.coelicolor* A3(2) Red$^-$ mutants with *Serratia marcescens* prodigiosin DNA. Conclusions and suggestions for future work are discussed.

Over-production of Red in proline transport mutants was observed. In addition to abolition of proline transport a reduction in the activity of proline catabolic enzymes was also reported (Hood *et al.*, 1992). Proline synthesis has been shown to be constitutive. The hypothesis was proposed that proline anabolism and catabolism is in dynamic equilibrium, if this equilibrium is disturbed, as in the proline transport mutants, Red biosynthesis can act as a sink for excess proline. This study has found that in a *put-44* mutant inactivation of Red was not a lethal event. The *put-44* allele is the weakest of the *put* alleles, as such is the most amenable to genetic and biochemical study. In the *put-44* allele only slightly more Red is produced than in the progenitor strain J802. The Put$^-$ mutants which show the least activity of proline catabolic enzymes produce the most Red. It is proposed that in the more extreme *put* alleles such as *put-7*, which produces ten to twenty five times more Red than the parent strain J802 (U. Swoboda, personal communication), inactivation of Red may be a lethal event.

To investigate the effect of increased proline concentration on production of Red, controlled over-production of proline is required. This could be achieved by introducing extra copies of genes of the proline synthetic enzymes

into the wild-type strain. It is anticipated that over-production of proline and, hence, increased production of Red would result. Proline synthesis involves four steps, three are catalysed enzymatically by the ProA, ProB and ProC gene products. The *proA* and *proB* genes have been cloned from *S.coelicolor* A3(2). Attempts to clone the *proC* gene were unsuccessful (Hood *et al.*, 1992). However the enzyme product of *proC* Pyrrolidine-5-carboxylate reductase was purified and amino acid sequence was obtained (D. Smith, personal communication). This sequence is being used in an attempt to clone the gene (R. Hebdon, personal communication).

The question of whether the Red over-production was due to increased stimulation of Red biosynthesis or that biosynthesis was activated earlier was investigated. S1 nuclease mapping was used to detect *red* gene transcripts in a *put-44* mutant and in the parent strain J802. Transcripts of *redD* and *redX* were detected in mid-exponential phase in the *put-44* mutant US224. In the parent strain J802 transcripts of the Red genes were not detected until transition phase. It has been suggested that reduction in growth rate is a factor involved in stimulating antibiotic biosynthesis (Takano *et al.*, 1992). Growth rate is not involved in stimulating Red biosynthesis earlier in the Put$^-$ strain as the Put$^-$ strain US224 has a similar growth rate to the progenitor strain J802. Red biosynthesis but not Act biosynthesis is stimulated, suggesting the effect is specific to Red.

It is of interest to further characterise the Put$^-$ lesion. One strategy would be to isolate Put$^-$ mutants in a clean genetic background i.e. by selecting for Put$^-$ mutants in an *aga*$^+$ strain. The Put$^-$ mutants available at the present time were isolated in an *aga*$^-$ background that causes the inability to grow on five carbon amino acids as sole carbon source (U.Swoboda personal communication). The Put$^-$ mutants used in this study were the progeny of a cross between

PTM44f and J1501. The PTM44f *put-44* mutant was the only mutant to be isolated from J802 directly by selecting for resistance to proline analogues. The PUM mutants isolated as described in section 3.1.1, show a more extreme Put⁻ phenotype than the strains used in this study. The PTM44f mutant has a less extreme Put phenotype only slightly more Red is produced and some activity, although reduced of the proline degradative enzymes was recorded. The mutants US224 and US218 were the most amenable because they are capable of sporulation and growth in liquid medium. In isolating the Put⁻ mutants from the *aga*⁻ strain, J802, further understanding of the *aga*⁻ lesion was unveiled. Owing to the complicated history of the existing Put⁻ mutants it seems appropriate to isolate new Put⁻ mutants in a cleaner genetic background. Isolation of more Put⁻ mutants may facilitate complementation of Put⁻ mutations using a *S. coelicolor* A3(2) DNA library.

The levels of proline within the cell are one mechanism used to control the osmotic potential within many organisms including *Streptomyces* spp (Killham and Firestone, 1984b). In addition to the role that proline has as a precursor for Red, proline has been shown to be a compatible solute (Chapter 1). From DNA sequence analysis of the right hand end of the *red* cluster an incomplete open-reading frame has been identified (M.Bibb, personal communication). The predicted protein sequence of the open-reading frame shows homology to the *trkA* gene of *E.coli*. In *E.coli trkA* encodes a membrane protein which is required for the function of a constitutive, low affinity potassium uptake system (Schlösser *et al.*, 1995). DNA sequence of the *red* cluster past the *trkA* homologue has not been obtained. The DNA clone containing the *red* cluster extends approximately a further 2.2kb past the *trkA* homologue. It has not been determined if another gene required for the biosynthesis of Red is located to the right of the *trkA* homologue. If this were the case it might be anticipated that the *trkA* homologue may have some

interaction with the *red* genes, possibly involved in osmoregulation to account for fluctuating levels of the Red precursor proline. It is also possible that the presence of the *trkA* homologue in close proximity to the *red* genes is co-incidental.

More is known about the chemical pathways and genetics of secondary metabolite biosynthesis than is understood about the supply of precursor molecules. Work of Masurekar and Sosa (1988) provided evidence of a *Streptomyces avermitilis* mutant that over-produced avermectin due to an alteration in precursor transport. This is an analogous situation to the over-production of Red by proline transport mutants of *S. coelicolor* A3(2). An avermectin over-producing mutant of *S. avermitilis* was selected by resistance to threonine analogues. Threonine is a precursor of avermectin, so it was anticipated that the mutation would be involved in threonine biosynthesis. However, over-production of threonine was not found, instead reduced transport of threonine and methionine was demonstrated. The Put$^-$ mutants were selected in a similar manner by resistance to an amino acid analogue. The aim was to investigate the origin of proline incorporated into Red. Red over-production was then observed in these mutants.

The second area of work involved the characterisation of the Pwb mutations. The region of DNA containing the Pwb-6 mutation was sub-cloned and DNA sequence data obtained. An open-reading frame of 864 nucleotides was identified, and was named *redZ*. The predicted protein sequence showed homology to the UhpA-LuxR family of response regulators. Phosphorylation is required for activation of UhpA, DegU and other regulators of this family of proteins. Of four residues important in phosphorylation RedZ contains only one. The residue Asp 54 that is phosphorylated in the UhpA protein, is not

present in RedZ. It is predicted that RedZ does not require phosphorylation for activation.

The *redZ* gene contains an in-frame TTA codon. The leucyl tRNA that recognises the UUA codon is encoded by the *bldA* gene. The Red structural genes do not contain a TTA codon because Red biosynthesis can be induced under conditions of low phosphate in *bldA* mutants (Guthrie and Chater, 1990). The pathway activator *redD* has also been shown not to contain a TTA codon (Narva and Feitelson, 1990). It is proposed that *redZ* mediates the *bldA* dependence of Red biosynthesis.

Only one of the Pwb mutations, Pwb-6 was sequenced in the course of this study. The Pwb-6 mutation lies within a putative promoter region. The mutation makes the -35 promoter region more similar to the enteric bacterium major sigma factor promoter consensus sequence. It is anticipated that the mutation will cause increased transcription from the promoter. This hypothesis can be confirmed by determining the abundance of *redZ* transcripts in a Pwb mutant and the levels of transcripts present in the progenitor of the Pwb mutants, *bldA* cells and levels in wild-type strains.

The proposed increased transcription of *redZ* due to the Pwb-6 mutation results in increased production of Red in the Pwb-6 strain. It is anticipated that the presence of multicopies of *redZ* within the cell would have the same effect on Red biosynthesis. It is predicted that *redZ* is involved in the regulation of Red biosynthesis. Inactivation of *redZ* would determine whether *redZ* is required for activation of Red biosynthesis.

The identification of *redZ*, which is potentially another regulator involved in Red biosynthesis, raises the issue of what interaction, if any, does *redZ*

have with *redD*? Inactivation of *redD* in the Pwb-9 mutant did not abolish Red biosynthesis. Some Red biosynthesis was observed, it appeared less than in the Pwb-9 mutant. The inactivation of *redD* was achieved by using the ØC31 derived, insertional phage ØKC899. When the lysogens of Pwb-9 were streaked to give single colonies white colonies and white sectors of colonies were observed. This was interpreted to imply that there is a selection pressure that favours the abolition of Red biosynthesis. One possible selection pressure that could arise in this situation would be if the *redZ* product was unable to induce transcription of the Red resistance genes whilst it was able to activate Red biosynthesis. In this situation a functional *redD* gene may be required for the activation of resistance genes.

Introduction of an extra copy of the Pwb-9 $redZ^{p}$ gene into a *redD* mutant did not give rise to Red biosynthesis. This could be interpreted to mean that $redZ^{p}$ is not epistatic to *redD*. Another interpretation is that as in the previous experiment, there may be a selection pressure to abolish Red biosynthesis. The lysogen isolated that did not produce Red may have already undergone a second mutation to abolish Red biosynthesis.

The *redD* gene product is a positive activator of Red biosynthesis (Narva and Feitelson, 1990). Transcription of *redD* precedes transcription of the Red structural genes (Takano *et al.*, 1992). DNA binding activity of RedD has not been detected (J.White personal communication). The predicted amino acid sequence of RedZ has a region that shows similarity to the helix-turn-helix DNA binding domain of response regulators of two-component systems. The obvious suggestion is that RedD interacts with RedZ to cause transcriptional activation of the Red structural genes. It is proposed that RedD may require RedZ for activation of transcription of *red* genes.

In the course of identifying a RedD dependent promoter, S1 nuclease mapping experiments were conducted using the *Bam*HI *Eco*RI DNA fragment of the *red* cluster that contains the *redZ* promoter region (J.White personal communication). This provides additional information about the transcription of *redZ*. Using a probe labelled at the *Eco*RI site, a transcript of 165 nucleotides was detected. The translational start site of *redZ* has been predicted by FRAME analysis and is believed to be 100 nucleotides downstream of the transcriptional start site based on this result. Hopwood *et al.* (1986) reported that long mRNA leader sequences are common in streptomycete genes. Given this, a leader sequence for *redZ* of 100 bases is feasible.

In M145 *redZ* transcripts were present in all stages of growth. This suggests that *redZ* may be constitutively transcribed. In the *bldA* strain J1681 *redZ* transcripts were detected. Even more significant is the detection of *redZ* transcripts in a *redD* disrupted lysogen (J.White personal communication). It implies that *redZ* does not rely on *redD* for transcription. This would support the supposition that *redZ* is epistatic to *redD*.

It would be interesting to determine the interaction of *redZ* with other regulators believed to be involved in the induction of secondary metabolism. Attempts were made to determine if *redZ* interacted with the global regulatory loci *absA* and *absB*. The predicted amino acid sequence of AbsA shows homology to sensory protein kinases of two component systems. The similarity of RedZ to the response regulator partner of the sensory kinase in two-component systems prompted this investigation. When *redZ*^{p} was introduced on a ØC31 derived phage into *absA* and *absB* mutants, Red biosynthesis was not observed. The identification of a response regulator partner for AbsA in close proximity on the chromosome and the absence of

conserved residues involved in phosphorylation in RedZ, both support the conclusion that *redZ* does not interact directly with *absA*.

The interaction of *redZ* with other regulatory loci could be assessed by introducing multiple copies of *redZ* into the regulatory mutants and identifying if *redZ* can suppress the mutant phenotypes to cause induction of Red biosynthesis.

APPENDIX 1

Table of Restriction sites within the *Bgl*II-*Eco*RI *redZ* region

	# SITES	FRAGMENTS		FRAGMENT ENDS	
AFL 3 (ACPQGT)					
	1				
	76	984	(92.8)	76	1060
		76	(7.2)	1	76
AHA 2 (GPCGQC)					
	2				
	154	697	(65.8)	363	1060
	363	209	(19.7)	154	363
		154	(14.5)	1	154
ALU 1 (AGCT)					
	1				
	400	660	(62.3)	400	1060
		400	(37.7)	1	400
APA 1 (GGGCCC)					
	1				
	345	715	(67.5)	345	1060
		345	(32.5)	1	345
APA L1 (GTGCAC)					
	2				
	70	529	(49.9)	531	1060
	531	461	(43.5)	70	531
		70	(6.6)	1	70
AVA 1 (CQCGPG)					
	4				
	447	447	(42.2)	1	447
	652	212	(20.0)	836	1048
	836	205	(19.3)	447	652
	1048	184	(17.4)	652	836
		12	(1.1)	1048	1060
AVA 2 (GGRCC)					
	4				
	178	564	(53.2)	178	742
	742	178	(16.8)	1	178
	871	150	(14.2)	871	1021
	1021	129	(12.2)	742	871
		39	(3.7)	1021	1060
BAL 1 (TGGCCA)					
	2				
	325	474	(44.7)	325	799
	799	325	(30.7)	1	325
		261	(24.6)	799	1060
BAM H1 (GGATCC)					
	1				
	757	757	(71.4)	1	757
		303	(28.6)	757	1060

redz

	#	SITES	FRAGMENTS		FRAGMENT ENDS	
BAN 1 (GGQPCC)	2					
		127	906	(85.5)	154	1060
		154	127	(12.0)	1	127
			27	(2.5)	127	154
BAN 2 (GPGCQC)	1					
		345	715	(67.5)	345	1060
			345	(32.5)	1	345
BBV 1 (GCTGC)	2					
		232	447	(42.2)	613	1060
		613	381	(35.9)	232	613
			232	(21.9)	1	232
BCE F1 (ACGGC)	7					
		330	330	(31.1)	1	330
		495	300	(28.3)	495	795
		795	165	(15.6)	330	495
		844	120	(11.3)	940	1060
		877	49	(4.6)	795	844
		924	47	(4.4)	877	924
		940	33	(3.1)	844	877
			16	(1.5)	924	940
BCG 1 (CGANNNNNNTGC)	1					
		554	554	(52.3)	1	554
			506	(47.7)	554	1060
BGL 1 (GCCNNNNNGGC)	2					
		327	493	(46.5)	327	820
		820	327	(30.8)	1	327
			240	(22.6)	820	1060
BIN 1 (GGATC)	3					
		1	756	(71.3)	1	757
		757	302	(28.5)	758	1060
		758	1	(0.1)	757	758
			1	(0.1)	1	1
BSA1 (GGTCTC)	2					
		303	721	(68.0)	339	1060
		339	303	(28.6)	1	303
			36	(3.4)	303	339
BSA A1 (QACGTP)	2					
		171	576	(54.3)	171	747
		747	313	(29.5)	747	1060
			171	(16.1)	1	171

redz

	#	SITES	FRAGMENTS		FRAGMENT ENDS	
BSA J1 (CCNNGG)	14					
		42	269	(25.4)	382	651
		248	206	(19.4)	42	248
		328	113	(10.7)	683	796
		382	111	(10.5)	922	1033
		651	80	(7.5)	248	328
		682	55	(5.2)	867	922
		683	54	(5.1)	328	382
		796	42	(4.0)	1	42
		824	31	(2.9)	836	867
		836	31	(2.9)	651	682
		867	28	(2.6)	796	824
		922	15	(1.4)	1033	1048
		1033	12	(1.1)	1048	1060
		1048	12	(1.1)	824	836
			1	(0.1)	682	683
BSG 1 (GTGCAG)	1					
		294	766	(72.3)	294	1060
			294	(27.7)	1	294
BSM 1 (GAATGC)	2					
		465	465	(43.9)	1	465
		862	397	(37.5)	465	862
			198	(18.7)	862	1060
BSM A1 (GTCTCN)	4					
		34	385	(36.3)	338	723
		302	337	(31.8)	723	1060
		338	268	(25.3)	34	302
		723	36	(3.4)	302	338
			34	(3.2)	1	34
BSP M1 (ACCTGC)	1					
		296	764	(72.1)	296	1060
			296	(27.9)	1	296
BSP M2 (TCCGGA)	1					
		575	575	(54.2)	1	575
			485	(45.8)	575	1060
BSR 1 (ACTGGN)	2					
		527	527	(49.7)	1	527
		770	290	(27.4)	770	1060
			243	(22.9)	527	770
BST E2 (GGTNACC)	1					
		1013	1013	(95.6)	1	1013
			47	(4.4)	1013	1060

redz

	# SITES	FRAGMENTS		FRAGMENT ENDS	
BST N1 (CCRGG)	9				
	42	435	(41.0)	248	683
	191	149	(14.1)	42	191
	224	141	(13.3)	683	824
	248	122	(11.5)	824	946
	683	57	(5.4)	961	1018
	824	42	(4.0)	1018	1060
	946	42	(4.0)	1	42
	961	33	(3.1)	191	224
	1018	24	(2.3)	224	248
		15	(1.4)	946	961
CFR 1 (QGGCCP)	7				
	90	235	(22.2)	90	325
	325	201	(19.0)	505	706
	505	180	(17.0)	325	505
	706	155	(14.6)	905	1060
	799	93	(8.8)	706	799
	875	90	(8.5)	1	90
	905	76	(7.2)	799	875
	.	30	(2.8)	875	905
CFR TEN1 (PCCGGQ)	1				
	157	903	(85.2)	157	1060
		157	(14.8)	1	157
CVI J1 (PGCQ)	23				
	46	158	(14.9)	160	318
	55	87	(8.2)	707	794
	91	87	(8.2)	620	707
	160	87	(8.2)	506	593
	318	79	(7.5)	952	1031
	326	69	(6.5)	91	160
	346	61	(5.8)	445	506
	400	57	(5.4)	819	876
	445	54	(5.1)	346	400
	506	46	(4.3)	1	46
	593	45	(4.2)	400	445
	620	36	(3.4)	55	91
	707	29	(2.7)	1031	1060
	794	27	(2.5)	593	620
	800	25	(2.4)	876	901
	819	20	(1.9)	906	926
	876	20	(1.9)	326	346
	901	19	(1.8)	800	819
	906	13	(1.2)	939	952
	926	13	(1.2)	926	939
	939	9	(0.8)	46	55
	952	8	(0.8)	318	326
	1031	6	(0.6)	794	800
		5	(0.5)	901	906

redz

	#	SITES	FRAGMENTS		FRAGMENT ENDS	
DDE 1 (CTNAG)						
	1					
		83	977	(92.2)	83	1060
			83	(7.8)	1	83
DRD 1 (GACNNNNNNGTC)						
	1					
		351	709	(66.9)	351	1060
			351	(33.1)	1	351
DRD 2 (GAACCA)						
	1					
		544	544	(51.3)	1	544
			516	(48.7)	544	1060
DSA 1 (CCPQGG)						
	4					
		328	414	(39.1)	382	796
		382	328	(30.9)	1	328
		796	138	(13.0)	922	1060
		922	126	(11.9)	796	922
			54	(5.1)	328	382
EAR 1 (CTCTTCNNNN)						
	2					
		375	485	(45.8)	575	1060
		575	375	(35.4)	1	375
			200	(18.9)	375	575
ECI 1 (TCCGCC)						
	2					
		335	426	(40.2)	634	1060
		634	335	(31.6)	1	335
			299	(28.2)	335	634
ECO N1 (CCTNNNNNAGG)						
	2					
		248	681	(64.2)	248	929
		929	248	(23.4)	1	248
			131	(12.4)	929	1060
ECO O109 (PGGNCCQ)						
	4					
		44	575	(54.2)	44	619
		619	279	(26.3)	741	1020
		741	122	(11.5)	619	741
		1020	44	(4.2)	1	44
			40	(3.8)	1020	1060
ECO R1 (GAATTC)						
	2					
		259	796	(75.1)	259	1055
		1055	259	(24.4)	1	259
			5	(0.5)	1055	1060

redz

	#	SITES	FRAGMENTS		FRAGMENT ENDS	
ESP3 1 (GAGACG)						
	1					
		722	722	(68.1)	1	722
			338	(31.9)	722	1060
FAU 1 (CCCGC)						
	4					
		162	690	(65.1)	162	852
		852	162	(15.3)	1	162
		856	144	(13.6)	916	1060
		916	60	(5.7)	856	916
			4	(0.4)	852	856
FNU 4H1 (GCNGC)						
	11					
		92	280	(26.4)	333	613
		232	140	(13.2)	92	232
		278	105	(9.9)	777	882
		316	95	(9.0)	613	708
		333	92	(8.7)	1	92
		613	89	(8.4)	971	1060
		708	72	(6.8)	899	971
		777	69	(6.5)	708	777
		882	46	(4.3)	232	278
		899	38	(3.6)	278	316
		971	17	(1.6)	882	899
			17	(1.6)	316	333
FNU D2 (CGCG)						
	12					
		77	163	(15.4)	77	240
		240	147	(13.9)	315	462
		264	142	(13.4)	918	1060
		315	113	(10.7)	776	889
		462	109	(10.3)	601	710
		480	109	(10.3)	492	601
		492	77	(7.3)	1	77
		601	66	(6.2)	710	776
		710	51	(4.8)	264	315
		776	29	(2.7)	889	918
		889	24	(2.3)	240	264
		918	18	(1.7)	462	480
			12	(1.1)	480	492
FOK 1 (GGATG)						
	3					
		184	668	(63.0)	245	913
		245	184	(17.4)	1	184
		913	147	(13.9)	913	1060
			61	(5.8)	184	245

redz

	# SITES	FRAGMENTS		FRAGMENT ENDS	
GDI 2 (QGGCCG)	5				
	90	415	(39.2)	90	505
	505	201	(19.0)	505	706
	706	169	(15.9)	706	875
	875	155	(14.6)	905	1060
	905	90	(8.5)	1	90
		30	(2.8)	875	905
GSU 1 (CTGGAG)	1				
	273	787	(74.2)	273	1060
		273	(25.8)	1	273
HAE 1 (RGGCCR)	3				
	325	474	(44.7)	325	799
	799	325	(30.7)	1	325
	1030	231	(21.8)	799	1030
		30	(2.8)	1030	1060
HAE 2 (PGCGCQ)	2				
	154	587	(55.4)	473	1060
	473	319	(30.1)	154	473
		154	(14.5)	1	154
HAE 3 (GGCC)	17				
	46	166	(15.7)	160	326
	91	160	(15.1)	346	506
	160	114	(10.8)	506	620
	326	92	(8.7)	939	1031
	346	87	(8.2)	707	794
	506	87	(8.2)	620	707
	620	69	(6.5)	91	160
	707	57	(5.4)	819	876
	794	46	(4.3)	1	46
	800	45	(4.2)	46	91
	819	29	(2.7)	1031	1060
	876	25	(2.4)	876	901
	901	20	(1.9)	906	926
	906	20	(1.9)	326	346
	926	19	(1.8)	800	819
	939	13	(1.2)	926	939
	1031	6	(0.6)	794	800
		5	(0.5)	901	906
HGA 1 (GACGC)	5				
	241	461	(43.5)	599	1060
	364	241	(22.7)	1	241
	428	123	(11.6)	241	364
	481	118	(11.1)	481	599
	599	64	(6.0)	364	428
		53	(5.0)	428	481

redz

	#	SITES	FRAGMENTS		FRAGMENT ENDS	
HGI A1 (GRGCRC)						
	2					
		70	529	(49.9)	531	1060
		531	461	(43.5)	70	531
			70	(6.6)	1	70
HHA 1 (GCGC)						
	12					
		155	155	(14.6)	1	155
		265	153	(14.4)	611	764
		314	149	(14.1)	314	463
		463	141	(13.3)	919	1060
		474	130	(12.3)	789	919
		493	110	(10.4)	155	265
		602	109	(10.3)	493	602
		611	49	(4.6)	265	314
		764	19	(1.8)	474	493
		775	14	(1.3)	775	789
		789	11	(1.0)	764	775
		919	11	(1.0)	463	474
			9	(0.8)	602	611
HINC 2 (GTQPAC)						
	1					
		202	858	(80.9)	202	1060
			202	(19.1)	1	202
HINF 1 (GANTC)						
	3					
		113	395	(37.3)	418	813
		418	305	(28.8)	113	418
		813	247	(23.3)	813	1060
			113	(10.7)	1	113
HPA 2 (CCGG)						
	12					
		132	233	(22.0)	343	576
		158	185	(17.5)	158	343
		343	179	(16.9)	576	755
		576	132	(12.5)	1	132
		755	95	(9.0)	954	1049
		837	82	(7.7)	755	837
		869	46	(4.3)	908	954
		874	32	(3.0)	837	869
		904	30	(2.8)	874	904
		908	26	(2.5)	132	158
		954	11	(1.0)	1049	1060
		1049	5	(0.5)	869	874
			4	(0.4)	904	908
HPH 1 (GGTGA)						
	4					
		39	463	(43.7)	550	1013
		379	340	(32.1)	39	379
		550	171	(16.1)	379	550
		1013	47	(4.4)	1013	1060
			39	(3.7)	1	39

redz

	#	SITES	FRAGMENTS		FRAGMENT ENDS	
MAE 2 (ACGT)						
	5					
		172	312	(29.4)	748	1060
		206	307	(29.0)	206	513
		513	213	(20.1)	535	748
		535	172	(16.2)	1	172
		748	34	(3.2)	172	206
			22	(2.1)	513	535
MAE 3 (GTNAC)						
	2					
		841	841	(79.3)	1	841
		1014	173	(16.3)	841	1014
			46	(4.3)	1014	1060
MBO 2 (GAAGA)						
	5					
		4	434	(40.9)	626	1060
		376	372	(35.1)	4	376
		522	146	(13.8)	376	522
		579	57	(5.4)	522	579
		626	47	(4.4)	579	626
			4	(0.4)	1	4
MCR 1 (CGPQCG)						
	6					
		236	358	(33.8)	517	875
		505	269	(25.4)	236	505
		517	236	(22.3)	1	236
		875	83	(7.8)	905	988
		905	72	(6.8)	988	1060
		988	30	(2.8)	875	905
			12	(1.1)	505	517
MLU 1 (ACGCGT)						
	1					
		76	984	(92.8)	76	1060
			76	(7.2)	1	76
MLY 1 (GASTC)						
	2					
		418	418	(39.4)	1	418
		813	395	(37.3)	418	813
			247	(23.3)	813	1060
MME 1 (TCPAC)						
	6					
		138	465	(43.9)	519	984
		203	191	(18.0)	203	394
		394	138	(13.0)	1	138
		519	125	(11.8)	394	519
		984	65	(6.1)	138	203
		1004	56	(5.3)	1004	1060
			20	(1.9)	984	1004

redz

	#	SITES	FRAGMENTS		FRAGMENT ENDS	
MNL 1 (CCTC)						
	16					
		11	161	(15.2)	821	982
		48	156	(14.7)	99	255
		59	119	(11.2)	255	374
		99	111	(10.5)	471	582
		255	105	(9.9)	716	821
		374	97	(9.2)	374	471
		471	51	(4.8)	982	1033
		582	50	(4.7)	582	632
		632	45	(4.2)	671	716
		651	40	(3.8)	59	99
		668	37	(3.5)	11	48
		671	27	(2.5)	1033	1060
		716	19	(1.8)	632	651
		821	17	(1.6)	651	668
		982	11	(1.0)	48	59
		1033	11	(1.0)	1	11
			3	(0.3)	668	671
MST 1 (TGCGCA)						
	2					
		763	763	(72.0)	1	763
		788	272	(25.7)	788	1060
			25	(2.4)	763	788
MWO 1 (GCNNNNNNNGC)						
	16					
		47	185	(17.5)	47	232
		232	162	(15.3)	604	766
		272	148	(14.0)	912	1060
		314	110	(10.4)	327	437
		327	99	(9.3)	465	564
		437	70	(6.6)	820	890
		465	47	(4.4)	1	47
		564	42	(4.0)	272	314
		585	40	(3.8)	780	820
		602	40	(3.8)	232	272
		604	28	(2.6)	437	465
		766	22	(2.1)	890	912
		780	21	(2.0)	564	585
		820	17	(1.6)	585	602
		890	14	(1.3)	766	780
		912	13	(1.2)	314	327
			2	(0.2)	602	604
NAE 1 (GCCGGC)						
	1					
		157	903	(85.2)	157	1060
			157	(14.8)	1	157
NAR 1 (GGCGCC)						
	1					
		154	906	(85.5)	154	1060
			154	(14.5)	1	154

redz

	# SITES	FRAGMENTS		FRAGMENT ENDS	
NCI 1 (CCSGG)					
	10				
	131	493	(46.5)	343	836
	343	212	(20.0)	131	343
	836	131	(12.4)	1	131
	837	94	(8.9)	954	1048
	868	46	(4.3)	908	954
	903	35	(3.3)	868	903
	908	31	(2.9)	837	868
	954	11	(1.0)	1049	1060
	1048	5	(0.5)	903	908
	1049	1	(0.1)	1048	1049
		1	(0.1)	836	837
NCO 1 (CCATGG)					
	1				
	382	678	(64.0)	382	1060
		382	(36.0)	1	382
NLA 3 (CATG)					
	4				
	96	660	(62.3)	383	1043
	287	191	(18.0)	96	287
	383	96	(9.1)	287	383
	1043	96	(9.1)	1	96
		17	(1.6)	1043	1060
NLA 4 (GGNNCC)					
	9				
	127	275	(25.9)	345	620
	154	200	(18.9)	757	957
	178	167	(15.8)	178	345
	345	127	(12.0)	1	127
	620	122	(11.5)	620	742
	742	64	(6.0)	957	1021
	757	39	(3.7)	1021	1060
	957	27	(2.5)	127	154
	1021	24	(2.3)	154	178
		15	(1.4)	742	757
NSP B2 (CVGCWG)					
	1				
	857	857	(80.8)	1	857
		203	(19.2)	857	1060
PLE 1 (GAGTCNNNN)					
	2				
	414	414	(39.1)	1	414
	813	399	(37.6)	414	813
		247	(23.3)	813	1060
PPU M1 (PGGRCCQ)					
	2				
	741	741	(69.9)	1	741
	1020	279	(26.3)	741	1020
		40	(3.8)	1020	1060

redz

	#	SITES	FRAGMENTS		FRAGMENT ENDS	
PSH A1 (GACNNNNGTC)						
	3					
		195	771	(72.7)	215	986
		215	195	(18.4)	1	195
		986	74	(7.0)	986	1060
			20	(1.9)	195	215
RSA 1 (GTAC)						
	3					
		174	315	(29.7)	174	489
		489	310	(29.2)	750	1060
		750	261	(24.6)	489	750
			174	(16.4)	1	174
RSR 2 (CGGRCCG)						
	1					
		870	870	(82.1)	1	870
			190	(17.9)	870	1060
SAU 3A (GATC)						
	2					
		2	756	(71.3)	2	758
		758	302	(28.5)	758	1060
			2	(0.2)	1	2
SAU 96 (GGNCC)						
	13					
		45	274	(25.8)	346	620
		160	167	(15.8)	178	345
		178	122	(11.5)	620	742
		345	115	(10.8)	45	160
		346	83	(7.8)	938	1021
		620	76	(7.2)	742	818
		742	53	(5.0)	818	871
		818	45	(4.2)	1	45
		871	39	(3.7)	1021	1060
		901	30	(2.8)	871	901
		926	25	(2.4)	901	926
		938	18	(1.7)	160	178
		1021	12	(1.1)	926	938
			1	(0.1)	345	346

redz

	# SITES	FRAGMENTS		FRAGMENT ENDS	
SCR F1 (CCNGG)	19				
	42	340	(32.1)	343	683
	131	141	(13.3)	683	824
	191	95	(9.0)	248	343
	224	89	(8.4)	42	131
	248	60	(5.7)	131	191
	343	57	(5.4)	961	1018
	683	42	(4.0)	1	42
	824	38	(3.6)	908	946
	836	35	(3.3)	868	903
	837	33	(3.1)	191	224
	868	31	(2.9)	837	868
	903	30	(2.8)	1018	1048
	908	24	(2.3)	224	248
	946	12	(1.1)	824	836
	954	11	(1.0)	1049	1060
	961	8	(0.8)	946	954
	1018	7	(0.7)	954	961
	1048	5	(0.5)	903	908
	1049	1	(0.1)	1048	1049
		1	(0.1)	836	837
SDU 1 (G2GC3C)					
	70	529	(49.9)	531	1060
	128	217	(20.5)	128	345
	345	186	(17.5)	345	531
	531	70	(6.6)	1	70
		58	(5.5)	70	128
SFA N1 (GATGC)	4				
	135	438	(41.3)	391	829
	391	256	(24.2)	135	391
	829	148	(14.0)	912	1060
	912	135	(12.7)	1	135
		83	(7.8)	829	912
SMA 1 (CCCGGG)	2				
	836	836	(78.9)	1	836
	1048	212	(20.0)	836	1048
		12	(1.1)	1048	1060
SPL 1 (CGTACG)	1				
	173	887	(83.7)	173	1060
		173	(16.3)	1	173
STU 1 (AGGCCT)	1				
	1030	1030	(97.2)	1	1030
		30	(2.8)	1030	1060

redz

	#	SITES	FRAGMENTS		FRAGMENT ENDS	
STY 1 (CCRRGG)						
	1					
		382	678	(64.0)	382	1060
			382	(36.0)	1	382
TAQ 1 (TCGA)						
	6					
		138	464	(43.8)	520	984
		394	256	(24.2)	138	394
		469	138	(13.0)	1	138
		520	75	(7.1)	394	469
		984	55	(5.2)	1005	1060
		1005	51	(4.8)	469	520
			21	(2.0)	984	1005
TAQ 2 (GACCGA)						
	1					
		516	544	(51.3)	516	1060
			516	(48.7)	1	516
TAQ 20 (CACCCA)						
	1					
		308	752	(70.9)	308	1060
			308	(29.1)	1	308
TFI 1 (GARTC)						
	1					
		113	947	(89.3)	113	1060
			113	(10.7)	1	113
TSP45 1 (GTSAC)						
	1					
		1014	1014	(95.7)	1	1014
			46	(4.3)	1014	1060
TSP E1 (AATT)						
	2					
		260	796	(75.1)	260	1056
		1056	260	(24.5)	1	260
			4	(0.4)	1056	1060
TTH111 1 (GACNNNGTC)						
	2					
		237	779	(73.5)	237	1016
		1016	237	(22.4)	1	237
			44	(4.2)	1016	1060
TTH111 2 (CCAPCA)						
	2					
		271	423	(39.9)	271	694
		694	366	(34.5)	694	1060
			271	(25.6)	1	271
XCM 1 (CCANNNNNNNNNTGG)						
	1					
		694	694	(65.5)	1	694
			366	(34.5)	694	1060

redz

	#	SITES	FRAGMENTS		FRAGMENT ENDS	
XHO 2 (PGATCQ)						
	2					
		1	756	(71.3)	1	757
		757	303	(28.6)	757	1060
			1	(0.1)	*1*	*1*
XMA 3 (CGGCCG)						
	3					
		505	505	(47.6)	1	505
		875	370	(34.9)	505	875
		905	155	(14.6)	905	1060
			30	(2.8)	875	905

The following do not appear:

AAT 2	ACC 1	AFL 2	AGE 1
AHA 3	ALW N1	ASE 1	ASU 2
AVA 3	AVR 2	BBS 1	BCL 1
BGL 2	BSA B1	BSP H1	BSS H2
BST X1	CLA 1	DRA 3	EC047 3
ECO57 1	ECO R5	ESP 1	FSE 1
HGI E2	HIND 3	HPA 1	KPN 1
MAE 1	MFE 1	MSE 1	MST 2
NDE 1	NHE 1	NOT 1	NRU 1
NSI 1	NSP 1	PAC 1	PFL M1
PST 1	PVU 1	PVU 2	SAC 1
SAC 2	SAL 1	SAP 1	SCA 1
SFE 1	SFI 1	SNA 1	SNA B1
SPE 1	SPH 1	SSP 1	XBA 1
XHO 1	XMN 1		

Sequence data of *pwb-6*

```
                        10         20         30         40         50         60
redZ pwb6 1     GGATCTTCTTGAGGTGGAAACCACTTCGTATCAGTCTCTCACCAGGGCCTCGCAAGCCCC

cf3norwich      GGATCTTCTTGAGGTGGAAACCACTTCGTATCAGTCTCTCACCAGGGCCTCGCAAGCCCC



                        70         80         90         100        110        120
redZ pwb6 1     TnTCCAAGTGTGCACACGCGTGCTAAGTTTGGCCGCATGAGGCATATCGGGGGAATCGCA

cf3norwich      TtTCCAAGTGTGCACACGCGTGCTAAGTTTGGCCGCATGAGGCATATCGGGGGAATCGCA



                        130        140        150        160        170        180
redZ pwb6 1     AAGAAAGGTGCCCGGCATCGACTTGCCACCGTCGGCGCCGGCCCGCACGCCACGTACGGT

cf3norwich      AAGAAAGGTGCCCGGCATCGACTTGCCACCGTCGGCGCCGGCCCGCACGCCACGTACGGT



                        190        200        210        220        230        240
redZ pwb6 1     CCCCATCCTTCCTGGACGAAAGTCAACGTATGACGACCCGTGTCCTGGTGTGCTGCGACC

cf3norwich      CCCCATCCTTCCTGGACGAAAGTCAA
cf3hinc1                                CAACGTATGACGACCCGTGTCCTGGTGTGCTGCGACC
cf3hinc2                                CAACGTATGACGACCCGTGTCCTGGTGTGCTGCGACC



                        250        260        270        280        290        300
redZ pwb6 1     GCGTCATCCTGGGCGAGGGAATTCGCGCATTGCTGGAGCGGCACGACATGAAAGTGCAGG

cf3hinc1        GCGTCATCCTGGGCGAGGGAATTCGCGCATTGCTGGAGCGGCACGACATGAAAGTGCAGG
cf3hinc2        GCGTCATCCTGGGCGAGGGAATTCGCGCATTGCTGGAGCGGCACGACATGAAAGTGCAGG
cf3m13132                         GAATTCGCGCATTGCTGGAGCGGCACGACATGAAAGTGCAGG
cf3m13136                         GAATTCGCGCATTGCTGGAGCGGCACGACATGAAAGTGCAGG
cf3na1          <                 GAATTCGCGCATTGCTGGAGCGGCACGACATGAAAGTGCAGG



                        310        320        330        340        350        360
redZ pwb6 1     TGGAGACCACCCAGCGCGGCTCGCTGGCCACGGCGGCGGAGACCGGGCCCGACATTCTGG

cf3hinc1        TGGAGACCACCCAGCtCGGCTCGCTGGCCACGGCGGCGGAGACCGGGCCCGACATTCTGG
cf3hinc2        TGGAGACCACCCAGCGCGGCTCGCTGGCCACGGCGGCGGAGACCGGGCCCGACATTCTGG
cf3m13132       TGGAGACCACCCAGCGCGGCTCGCTGGCCACGGCGGCGGAGACCGGGCCCGACATTCTGG
cf3m13136       TGGAGACCACCCAGCGCGGCTCGCTGGCCACGGCGGCGGAGACCGGGCCCGACATTCTGG
cf3na1          < TGGAGACCACCCAGCGCGGCTCGCTGGCCACGGCGGCGGAGACCGGGCCCGACATTCTGG
cf3nco5         <                                 CCACGGCGGCGGAGACCGGGCCCGACATTCTGG
cf3na2          <                                     CGGAGACCGGGCCCGACATTCTGG
cf3r4501        <                                         GGGCCCGACATTCTGG
```

SEQUENCE DATA OF *pwb-6*

```
                          370        380        390        400        410        420
redZ pwb6 1       TCGGCGTCGCCCCCCCTCTTCACCATGGACAGCATCGACAAGCTCACGGAACTCGCACGAC

cf3m13132         TCGGCGTCGCCCCCCCTCTTCACCATGGACAGCATCGACAAGCTCACGGAACTCGCACGAC
cf3m13136         TCGGCGTCGCCCCCCCTCTTCACCATGGACAGCATCGACAAGCTCACGGAACTCGCACGAC
cf3na1          < TCGGCGTCGCCCCCCCTCTTCACCATGG
cf3nco5         < TCGG GTCGCCCCCCCTCTTCACCATGG
cf3na2          < TCGGgGTCGCCCCCCCTCTTCACCATGG
cf3r4501        < TCGGCGTCGCCCCCCCTCTTCACCATGGACAGCATCGACAAGCTCACGGAACTCGCACGAC
cf3r4502        <                                                 CAAGCTCACGGAACTCGCACGAC
cf3nco1a                                                                        CACGAC
cf3nco1                                                                         AC
```

```
                          430        440        450        460        470        480
redZ pwb6 1       TCGGCAAGACGCTTTTGCTGACCAAGCCCGAGAATACCCACCGCGCATTCGAGGCGCTCC

cf3m13132         TCGGCAAGACGCTTTo CTGACCAAGCC GAGAATACC ACCGCGCATT
cf3m13136         TCGGCAAGACGCTTTTGCTGACCAAGCCCGAGAATACCCACCGCGCATTCGAGG
cf3r4501        < TCGGCAAGACGCTTTTGCTGACCAAGCCCGAGAATACCCACCGCGCATT
cf3r4502        < TCGGCA GACGCTTTTGCTGACCAAGCCCGAGAATACCCACCGCGCATTCGAGGC  TCC
cf3nco1a          TCGGCAAGACGCTTTTGCTGACCAAGCCCGAGAATACCCACCGCGCATTCGAGGCGCTCC
cf3nco1           TCGGCAAGACGCTTTTGCTGACCAAGCCCGAGAATACCCACCGCGCATTCGAGGCGCTCC
cf3nco4                        TTTTGCTGACCAAGCCCGAGAATACCCACCGCGCATTCGAGGCGCTCC
```

```
                          490        500        510        520        530        540
redZ pwb6 1       GCGTCGGAGTACGCGCCGTATTGTCGGCCGAAACGTCGGTCGAAGAACTGGTGCACGTCA

cf3r4502        < GCGTCGGAGTACGCGC GTATTGTCGGCCGAAACGTCGGTCGAAGAACTGGTGCACGTCA
cf3nco1a          GCGTCGGAGTACGCGCCGTATTGTCGGCCGAAACGTCGGTCGAAGAACTGGTGCACGTCA
cf3nco1           GCGTCGGAGTACGCGCCGTATTGTCGGCCGAAACGTCGGTCGAAGAACTGGTGCACGTCA
cf3nco4           GCGTCGGAGTACGCGCCGTATTGTCGGCCGAAACGTCGGTCGAAGAACTGGTGCACGTCA
cf3m1385                                      CCGAAACGTCGGTCGAAGAACTGGTGCACGTCA
cf3nco2                                           ACGTCGGTCGAAGAACTGGTGCACGTCA
```

```
                          550        560        570        580        590        600
redZ pwb6 1       TCAGAACCATCACCGAAGTCAATGCCATAATCGCTCCGGAAGAGGCACAGGAAGCCCTGA

cf3r4502        < TCAGAACCATCACCGAAGTCAATGCCATAATCGCTCC
cf3nco1a          TCAGAACCATCACCGAAGTCAATGCCATAATCGCTCCGGAAGAGGCACAGGAAGCCC
cf3nco1           TCAGAACCATCACCGAAGTCAATGCCATAATCGCTCCGGAAGAGGCACAGGAAGCCCTGA
cf3nco4           TCAGAACCATCACCGAAGTCAATGCCATAATCGCTCCGGAAGAGGCACAGGAAGCCCTGA
cf3m1385          TCAGAACCATCACCGAAGTCAATGCCATAATCGCTCCGGAAGAGGCACAGGAAGCCCTGA
cf3nco2           TCAGAACCATCACCGAAGTCAATGCCATAATCGCTCCGGAAGAGGCACAGGAAGCCCTGA
cf3nco3                 TCACCGAAGTCAATGCCATAATCGCTCCGGAAGAGGCACAGGAAGCCCTGA
cf3b81          <                                 ATAAT GCTC GGAAGAGGCACAGGAAGCCCTGA
cf3b82          <                                     CTCCGGAAGAGGCACAGGAAGCCCTGA
```

SEQUENCE DATA OF *pwb-6*

```
                          610        620        630        640        650        660
redZ pwb6 1     CGCGCTATTGGCGCAGCAAGGCCCCGAAGAACCTCCGCCCCGAACTGACCCCTCGGG AAA

cf3nco1         CGCGCTATTGG
cf3nco4         CGCGCTATTGGCGCAGCAAGGCCCCGAAGAACCTCCGCCCCGAACTGACCCCTCGG  AAA
cf3m1385        CGCGCTATTGGCGCAGCAAGGCCCCGAAGAACCTCCGCCCCGAACTGACCCCTCGGG AAA
cf3nco2         CGCGCTATTGGCGCAGCAAGGCCC GAAGAACCTCCGCCCCGAACTGACCCCTCGGG AAA
cf3nco3         CGCGCTATTGGCGCAGCAAGGCCCCGAAGAACCTCCGCCCCGAACTGACCCCTCGGGcAAA
cf3b81        < CcgGCTATTGGCGCAGCAAGGCCCCGAAGAACCTCCGCCCCGAACTGACCCCTCGGG AAA
cf3b82        < CGCGCTATTGGCGCAGCAAGGCCC GAAGAACCTCCGCCCCGAACTGACCCCTCGG  AAA
cf3s84        <                                                  aCCCTCGG  AA
cf34505                                                             GGG AAA
```

```
                    670        680        690        700        710        720
redZ pwb6 1     CGGAAGTCCTCCTCTTACTGACCCAGGGCAAGACCAACACC GAGATGGCCGCGACCCTCT

cf3nco4         CGGAAGTCCTCCTCTTACTGACCCAGGGCAAGACCAACAC
cf3m1385        CGGAAGTCCTCCTCTTACTGACCCAGGGCAAGACCAACACC GAGATGG
cf3nco2         CGGAAGTCCTCCTCTTACTGACCCAGGGCAAGACCAACACC GAGATGGCC
cf3nco3         CGGAAGTCCTCCTCTTACTGACCCAGGGCAAGACCAACACC.GAGATGGCCGCGACCCTCT
cf3b81        < CGGAAGTCCTCCTCTTACTGACCCAGGGCAAGACCAACACC GAGATGGCCGCGACCCTCT
cf3b82        < CGGAAGTCCTCCTCTTACTGACCCAGGGCAAGACCAACACC GAGATGGCCGCGACCCTCT
cf3s84        <  G AAGTCCTCCTCTTACTGACCCAGG CAAGACCAACACCaGAG TGGCCGCGACCCTCT
cf34505         CGGAAGTCCTCCTCTTACTGACCCAGGGCAAGACCAACACC GAGATGGCCGCGACCCTCT
cf34506              GTCCTCCTCTTACTGACCCAGGGCAAGACCAACACC GAGATGGCCGCGACCCTCT
cf34504                   CCTCTTACTGACCCAGGGCAAGACCAACACC GAGATGGCCGCGACCCTCT
cf34501                              CCCAGGGCAAGACCAACACC GAGATGGCCGCGACCCTCT
cf34502                              CCCAGGGCAAGACCAACACC GAGATGGCCGCGACC TCT
cf32501                                           CACC GAGATGGCCGCGACCCTCT
cf3s82        <                                     CC GAGATGGCCGCGACCCTCT
```

```
                         730        740        750        760        770        780
redZ pwb6 1     CCGTCTCGCCCACCACCGTCAGGTCCCACGTACACCGGATCCTGCGCAAACTGGGCGCGG

cf3b81        < CCGTCTCGCCCACCACCGTCAGGTCCCACGTACACCGGATCC
cf3b82        < CCGTCTCGCCCACCACCGTCAGGTCCCACGTAC
cf3s84        < CCGTCTCGCCCACCACCGTCAGGTCCCACGTACACCGGATCCTGCGCAAACTGGGCGCGG
cf34505         CCGTCTCGCCCACCACCGTCAGGTCCCACGTACACCGGtaCCTGCGCAAACTGG CGC
cf34506         CCGTCTCGCCCACCACCGTCAGGTCCCACGTACACCGGATCCTGCGCAAACTGGGCGCGG
cf34504         CCGTCTCGCCCACCACCGTCAGGTCCCACGTACACCGGATCCTGCGCAAACTGGGCGCGG
cf34501         CCGTCTCGCCCACCACCGTCAGGTCCCACGTACACCGGATCCTGCGCAAACTGGGCGCG
cf34502         CCGTCTCGCCCACCACCGTCAGGTCCCACGTACACCGGATCCTGCGCAAACTGGG
cf32501         CCGTCTCGCCCACCACCGTCAGGTCCCACGTACACCGGATCCTGCGCAAACTGGGCGCGG
cf3s82        < CCGTCTCGCCCACCACCGTCAGGTCCCACGTACACCGGATCCTGCGCAAACTGGGCGCGG
cf3s81        <   GTCTCGCCCACCACCGTCAGGTCCCACGTACACCGGATCCTGCGCAAACTGGGCGCGG
cf3s83        <                   GTCAGGTCCCACGTACACCGGATCCTGCGCAAACTGGGCGCGG
cf3b131                                               GGATCCTGCGCAAACTGGGCGCGG
cf3b132                                               GGATCCTGCGCAAACTGGGCGCGG
```

```
                              790        800        810        820        830        840
redZ pwb6 1     CGACCCGTGCGCAGGCCGTGGCCATCGCCTACGAGTCGGGCCTCCTGGGC ATCTGCCCGG

cf3s84        < CGACCCGTGCGCAGGCCGTGGCcATCGCCTACGAGTCGGGCCTCCTGGGC ATCTGCCCGG
cf34506         CGACCCGTGCGCAGGCCGTGGC ATCGCCTACGAGTCGGGCCTCCTGGGCtATC GCCCGG
cf34504         CGACCCGTGCGCAGGCCGTGGC ATCGCCTACGAGTCGGGCCTCCTGGGC ATCTGCCCGG
cf34501         CGACCCG
cf32501         CGACCCGTGCGCAGGCCGTGGC ATCGCCTACGAGTCGGGCCTCCTGGGC ATCTGCCCGG
cf3s82        < CGACCCGTGCGCAGGCCGTGGCcATCGCCTACGAGTCGGGCCTCCTGGGC ATCTGCCCGG
cf3s81        < CGACCCGTGCGCAGGCCGTGGCcATCGCCTACGAGTCGGGCCTCCTGGGC ATCTGCCCGG
cf3s83        < CGACCCGTGCGCAGGCCGTGGCcATCGCCTACGAGTCGGGCCTCCTGGGC ATCTGCCCGG
cf3b131         CGACCCGTGCGCAGGCCGTGGC ATCGCCTACGAGTCGGGCCTCCTGGGC ATCTGCC GG
cf3b132         CGACCCGTGCGCAGGCCGTGGC ATCGCCTACGAGTCGGGCCTCCTGGGC ATCTGCC GG
cf3m1384      <                                                          GG
```

SEQUENCE DATA OF *pwb-6*

```
                          850        860        870        880        890        900
redZ pwb6 1     GTTACGGCACTCCCGCCCGCTGAATGCCCCGGTCCGGCC GTGCCGCTTCGCGTTCCGTGC

cf3s84        < G
cf34506         GTTACGG
cf34504         GTT
cf32501         GTTACGGCACTCCCGCCCG
cf3s82        < G
cf3s81        < G
cf3s83        < G
cf3b131         GTTACG CACTCCCGCCCGCTGAATGCCCCGGTCCGG
cf3b132         GTTACG CACTCCCGCCCGCTGAATGCCCCGGTCCGGCCcGTGC GCTTCGCGTTCCGTGC
cf3m1384      < GTTACG CACTCCCGCCCagTGAATGCCCCGGTCCGGCC GTGCCG TTCGCGTTCCGTGC
cf32504           CGGCACTCCCGCCCGCTGAATGCCCCGGTCCGGCC GTGCCGCTTCGCGTTCCGTGC
cf32508           CGGCACTCCCGCCCGCTGAATGCCCCGGTCCGGCC GTGCCGCTTCGCGTTCCGTGC
cf32509           CGGCACTCCCGCCCGCTGAATGCCCCGGTCCGGCC GTGCCGCTTCGCGTTCCGTGC
cf32505                 CCCGCCCGCTGAATGCCCCGGTCCGGaCcGTGCCGCTTCGCGTTCCGTGC
cf32507                 CCCGCCCGCTGAATGCCC GGTCCGGCC GTGCCGCTTCGCGTTCCGTGC
cf3m13134     <                   TGAATGCCC GGTC G CC GTGC G T CGCGTTCCGTGC
cf32503                                             C GTGCCGCTTCGCGTTCCGTGC
cf3m1381      <                                           GCTTCGCGTTCCGTGC
cf3m1386      <                                           GCTTCGCGTTCCGTGC
cf3m13133     <                                                GCGTTCCGTGC
```

```
                        910        920          930        940          950
redZ pwb6 1     GGCCC GG CCGGGCATCCCGCG CCAC GGCCCTTTGTCAGGGCCG TG CCT GG

cf3b132         GGCCC GG CCGG CATCCCGCG CCACcGGCCCTTTGTCAGGGCCG TG CCTtGG
cf3m1384      < GGCCC GG CCGGGCATCCCGCG CCAC GGCCCTTTGTCAGGGCCG aG CCT GG
cf32504         GGCCC GG CCGGGCATCCCGCG CCAC GGCCCTTTGTCAGGGCCG TG CCT GG
cf32508         GGCCC GG CCGGGCATCCCGCG CCAC GGCCCTTTGTCAGGGCCG TG CCT GG
cf32509         GGCCC GG CCGGGCATCCCGCG CCAC GGCCCTTTGTCAGGGCCG TG CCT GG
cf32505         GGCCCcGGaCCGGGCATCCCGCGaCCAC GGCCCTTTGTCAGGGCCG TG CCT GG
cf32507         GGCCC GG CCGGGCATCCCGCG CCAC GGCCCTTTGTCAGGGCCG TG CCT GG
cf3m13134     < GGCCC GG CCGGGCATCCCGCG CCAC GGCCCTTTGTCAGGGCCGaTG CCT GG
cf32503         GGCCC GG CCGGGCATCCCGCG CCAC GGCCCTTTGTCAGGGCCG TG CCT GG
cf3m1381      < GGCCC GG CCGGGCATCCCGCG CCAC GGCCCTTTGTCAGGGCCG TG CCT GG
cf3m1386      < GGCCC GG CCGGGCATCCCGCG CCAC GGCCCTTTGTCAGGGCCG TG CCT GG
cf3m13133     < GGCCC GG CCGGGCATCCCGCG CCAC GGCCCTTTGTCAGGGCCG TG CCT GG
cf3m1382      <       GG CCGGGCATCC GCG CCAC GGCCCTTTGTCAGGGCCG TG CCT GG
cf32502                  CCGGGCATCCCGCG CCAC GGCCCTTTGTCAGGGCCGtTGaCCT GG
cf3m1388      <           CGGGCtaCC GCG CCAC GGCCCTTTGTCAGGGCCG TG CCT GG
cf32506                                                                G
```

SEQUENCE DATA OF *pwb-6*

```
                          960        970        980        990        1000       1010
redZ pwb6 1      TAGCCGGGTTCCAGGAGTTCGCCGCTCATCTCCTCGACGACCGTCG CCATCGTGTCGAAC

cf3b132            TAGC GGtT
cf3m1384       <   TAGCCGGGTTCCAGGAGTTCGCCGCTCATCTCCTCGACGACCGTCGaCCATCGTGTCGAAC
cf32504            TAGCCGGGTTCCAGGAGTTCGCCGCTCATCTCCTCGACGACCGTCGcC
cf32508            TAGCCGGGTTCCAGGAGTTCGCCGCTC
cf32509            TAGCCGGGTTCCAGGAGTTCGCCGCT
cf32505            TAGCCGGGTTCCgG
cf32507            TAGCCGGGTTCCAGGAGTTCGCCGCTCA
cf3m13134      <   TAGCCGGGTTCCAGGAGTTCGCCGCTCATCTCCTCGA GACCGTCG CCATCGTGTCGAAC
cf32503            TAGCCGGGTTCCAGGAGTTCGCCGCTCATCTCCTCGACGACCGTCG CCA
cf3m1381       <   TAGCCGGGTTCCAGGAGTTCGCCGCTCATCTCCTCGACGACCGTCG CCATCGTGTCGAAC
cf3m1386       <   TAGCCGGGTTCCAGGAGTTCGCCGCTCATCTCCTCGACGACCGTCG CCATCGTGTCGAAC
cf3m13133      <   TAGCCGGGTTCCAGGAGTTCGCCGCTCATCTCCTCGACGACCGTCG CCATCGTGTCGAAC
cf3m1382       <   TAGCCGGGTTCCAGGAGTTCGCCGCTCATCTCCTCGACGcCCGTCG CCATCGTGTCGAAC
cf32502            TAGCCGGGTTCCAGGAGTTCGCCGCTCATCTCCTCGACGACCGTCG CCATCG GTCGAAC
cf3m1388       <   TAGCCGGGTTCCAGGAGTTCGCCGCTCATCTCCTCGACGACCGTCG CCATCGTGTCGAAC
cf32506            TAGCCGGGTTCCAGGAGTTCGCCGCTCATCTCCTCGACGACCGTCG CCATCGTGTCGAAC
```

```
                          1020       1030       1040       1050       1060
redZ pwb6 1      ACGGTGACCAGGTCCCGAAAGGCCTCGGTTTTCATGTCCCGGGAGAATTC

cf3m1384       <   ACGGT
cf3m13134      <   ACGGTGACCAGGTCCCGAAAGGCCTCGGTTTTCATGTCCCGGGAGAATTC
cf3m1381       <   ACGGTGACCAGGTCCCGAAAGGCCTCGGTTTTCATGTCCCGGGAGAATTC
cf3m1386       <   AC
cf3m13133      <   ACGGTGACCAGGTCCCGAAAGGCCTCGGTTTTCATGTCCCGGGAGAATTC
cf3m1382       <   ACGGTGACCAGGTCCCGAAAGGCCTCGGTTTTCATGTCCCGGGAGAATTC
cf32502            ACGGTGACCAGGTCCCGAAAGGCCTCGGTTTc ATGTCC GGGAGA T C
cf3m1388       <   ACGGTGACCAGGTCCCGAAAGGCCTCGGTTTTCA GTCCCGGGAGAATTC
cf32506            ACGGTGACCAGGTCCCGAAAGGCCTCGGTTTTCATGTCC GGGAGAATTC
```

Sequence data of WT *redZ*

```
                          10         20         30         40         50         60
wt redZ 1          GGATCTTCTTGAGGTGGAAACCACTTCGTATCAGTCTCTCACCAGGGCCTCGCAAGCCCC

cfwtbgl19ex2 <     GGATCTTCTTGAGGTGGAAACCACTTCGTATCAGTCTCTCACCAGGGCCTCGCAAGCCCC
cfwtbgl181         GGATCTTCTTGAGGTGGAAACCACTTCGTATCAGTCTCTCACCAGGGCCTCGCAAGCCCC
cfwtbgl192       < GGATCTTCTTGAGGTGGAAACCACTTCGTATCAGTCTCTCACCAGGGCCTCGCAAGCCCC
cfwtbgl193         GGATCTTCTTGAGGTGGAAACCACTTCGTATCAGTCTCTCACCAGGGCCTCGCAAGCCCC
```

```
                          70         80         90         100        110        120
wt redZ 1          TCTCCAAGTGTGCACACGCGTGCTAAGTTTGGCCGCATGAGGCATATCGGGGGAATCGCA

cfwtbgl19ex2 <     TCTCCAAGTGTGCACACGCGTGCTAAGTTTGGCCGCATGAGGCATATCGGGGGAATCGCA
cfwtbgl181         TCTCCAAGTGTGCACACGCGTGCTAAGTTTGGCCGCATGAGGCATATCGGGGGAATCGCA
cfwtbgl192       < TCTCCAAGTGTGCACACGCGTGCTAAGTTTGGCCGCATGAGGCATATCGGGGGAATCGCA
cfwtbgl193         TCTCCAAGTGTGCACACGCGTGCTAAGTTTGGCCGCATGAG CATATCGGGGGAATCGCA
cfwtbglex2                   GCACACGCGTGCTAAGTTTGGCCGCATGAGGCATATCGGGGGAATCGCA
cfwtbgl191       <                                        CGCATGAGGCATATCGGGG AATCGCA
cfwtbgl182                                                            GGGGGAATCGCA
```

```
                          130        140        150        160        170        180
wt redZ 1          AAGAAAGGTGCCCGGCATCGACTTGCCACCGTCGGCG CCGGCCCGCACGCCACGTACGGT

cfwtbgl19ex2 <     AAGAAAGGTGCCCGGCATCGACTTGCCACCGTCGG
cfwtbgl181         AAGAAAGGTGCCCGGCATCGACTTGCCACCGTCGG
cfwtbgl192       < AAGAAA
cfwtbgl193         AAGAAAGGTGCCCGGCATCGACTTGCCACCGTCGGCG CCGGCCCGCACGCCACGTACGGT
cfwtbglex2         AAGAAAGGTGCCCGGCATCGACTTGCCACCGTCGGCGaCCGGCCCGCACGCCACGTACGGT
cfwtbgl191       < AAGAAAGGTGCCCGGCATCGACTTGCCACCGTCGG G CCGGCCCGCACGCCACGTACGGT
cfwtbgl182         AAGAAAGGTGCCCGGCATCGACTTGCCACCGTCGGCG CCGGCCCGCACGCCACGTACGGT
```

```
                          190        200        210        220        230        240
wt redZ 1          CCCCATCCTTCCTGGACGAAAGTCAACGTATGACGACCCGTG TCCTGGTGTGCTGCGACC

cfwtbgl193         CCCCATCCT
cfwtbglex2         CCCCATCCTTCCTGGACGAAAGTCAACGTATGACGACCCGTGGTCCTGGTGTG
cfwtbgl191       < CCCCATCCTTCCTGGACGAAAGTCAACGTATGACGACCCGTG TCCTGGTGTGCTGCGACC
cfwtbgl182         CCCCATCCTTCCTGGACGAAAGTCAACGTATGACGACCCGTA TCCTGGTGTGCTGCGACC
cfwthinc1                              CAACGTATGACGACCCGTG TCCTGGTGTGCTGCGACC
cfwthinc2                                          CCCGTG TCCTGGTGTGCTGCGACC
```

```
                          250        260        270        280        290        300
wt redZ 1          GCGTCATCCTGGGCGAGGGAATTCGCGCATTGCTGGAGCGGCACGACATGAAAGTGCAGG

cfwtbgl191       < GCGTCATCCTGGGCGAGG AATTC
cfwtbgl182         GCGTCATCCTGGGCGAGGGAATTC
cfwthinc1          GCGTCATCCTGGGCGAGGGAATTCGCGCATTGCTGGAGCGGCACGACATGAAAGTGCAGG
cfwthinc2          GCGTCATCCTGGGCGAGGGAATTCGCGCATTGCTGGAGCGGCACGACATGAAAGTGCAGG
cfwtm132                               GAATTCGCGCATTGCTGGAGCGGCACGACATGAAAGTGCAGG
```

SEQUENCE DATA OF WT *redZ*

```
                      310        320        330        340        350        360
wt redZ 1        TGGAGACCACCCAGCGCGGCTCGCTGGCCACGGCGGCGGAGACCGGGCCCGACATTCTGG

cfwthinc1        TGGAGACCACCCAGCGCGGCTCGCTGGCCACGGCGGCGGAGACCGGGCCCGACATTCTGG
cfwthinc2        TGGAGACCACCCAGC
cfwtm132         TGGAGACCACCCAGCGCGGCTCGCTGGCCACGGCGGCGGAGACCGGGCCCGACATTCTGG


                      370        380        390        400        410        420
wt redZ 1        TCGGCGTCGCCCCCCTCTTCACCATGGACAGCATCGACAAGCTCACGGAACTCGCACGAC

cfwthinc1          TC
cfwtm132           TCGGCGTCGCCCCCCTCTTCACCATGGACAGCATCGACAAGCTCACGGAACTCGCACGAC
cfwtr4501        <                                                CAAGCTCACGGAACTCGCACGAC


                             430        440        450        460        470        480
wt redZ 1          TCGGCAAGACGCTTTTGCTGACCAAGCCCGAGAATACCCACCGCGCATTCGAGGCGCTCC

cfwtm132           TCGGCAAGA
cfwtr4501        < TCGGCA GACGCTTTTGCTGACCAAGCCCGAGAATACCCACCGCGCATTCGAGGC  TCC
cfwtnco1                 GCTTTTGCTGACCAAGCCCGAGAATACCCACCGCGCATTCGAGGCGCTCC


                             490        500        510        520        530        540
wt redZ 1          GCGTCGGAGTACGCGCCGTATTGTCGGCCGAAACGTCGGTCGAAGAACTGGTGCACGTCA

cfwtr4501        < GCGTCGGAGTACGCGC GTATTGTCGGCCGAAACGTCGGTCGAAGAACTGGTGCACGTCA
cfwtnco1           GCGTCGGAGTACGCGCCGTATTGTCGGCCGAAACGTCGGTCGAAGAACTGGTGCACGTCA


                             550        560        570        580        590        600
wt redZ 1          TCAGAACCATCACCGAAGTCAATGCCATAATCGCTCCGGAAGAGGCACAGGAAGCCCTGA

cfwtr4501        < TCAGAACCATCACCGAAGTCAATGCCATAATCGCTCC
cfwtnco1           TCAGAACCATCACCGAAGTCAATGCCATAATCGCTCCGGAAGAGGCACAGGAAGCCCTGA


                             610        620        630        640        650        660
wt redZ 1          CGCGCTATTGGCGCAGCAAGGCCCCGAAGAACCTCCGCCCCGAACTGACCCCTCGGGAAA

cfwtnco1           CGCGCTATTGGCGCAGCAAGGCCCCGAAGAACCTCCGCCCCGAACTGACCCCTCGGGAAA


                             670        680        690        700        710        720
wt redZ 1          CGGAAGTCCTCCTCTTACTGACCCAGGGCAAGACCAACACCGAGATGGCCGCGACCCTCT

cfwtnco1           CGGAAGTCCT
cfwt4501             GAAGTCCTCCTCTTACTGACCCAGGGCAAGACCAACACCGAGATGGCCGCGACCCTCT
cfwt4501a                                CCCAGGGCAAGACCAACACCGAGATGGCCGCGACCCTCT
```

SEQUENCE DATA OF WT *redZ*

```
                    730        740        750        760        770        780
wt redZ 1       CCGTCTCGCCCACCACCGTCAGGTCCCACGTACACCGGATCCTGCGCAAACTGGGCGCGG

cfwt4501        CCGTCTCGCCCACCACCGTCAGGTCCCACGTACACCGGATCCTGCGCAAACTGGGCGCGG
cfwt4501a       CCGTCTCGCCCACCACCGTCAGGTCCCACGTACACCGGATCCTGCGCAAACTGGGCGCGG


                    790        800        810        820        830        840
wt redZ 1       CGACCCGTGCGCAGGCCGTGGC ATCGCCTACGAGTCGGGCCTCCTGGGCATCTGCCCGGG
cfwt4501        CGACCCGTGCGCAGGCCGTGGC ATCGCCTACGAGTCGGGCCTCCTGGGCATCTG
cfwt4501a       CGACCCGTGCGCAGGCCGTGGC ATCGCCTACGAGTCGGGCCTCCTGGGCATCTGCCCGGG


                    850        860        870        880        890        900
wt redZ 1       TTACGGCACTCCCGCCCGCTGAATGCCCCGGTCCGGCCGTGCCGCTTCGCGTTCCGTGCG

cfwt4501a       TTACGGCACTCCCGCCCGCTGAATGCC


                    910        920        930        940        950        960
wt redZ 1       GCCCGGCCGGGCATCCCGCGCCACGGCCCTTTGTCAGGGCCGTGCCTGGTAGCCGGGTTC

cfwt2501         CCCGGCCGGGCATCCCGCGCCACGGCCCTTTGTCAGGGCCGTGCCTGGTAGCCGGGTTC


                    970        980        990        1000       1010       1020
wt redZ 1       CAGGAGTTCGCCGCTCATCTCCTCGACGACCGTCGCCATCGTGTCGAACACGGTGACCAG

cfwt2501        CAGGAGTTCGCCGCTCATCTCCTCGACGACCGTCGCCATCGTGTCGAACACGGTGACCAG


                    1030       1040       1050       1060
wt redZ 1       GTCCCGAAAGGCnTCGGTTTTnATGTCCnGGGAGAATTCGAATCATGGT

pwb3 1          GTCCCGAAAGGCcTCGGTTTTcATGTCCcGGGAGAATTCGAATCATGGT
cfwt2501        GTCCCGAAAGGCcTCGGTTTT ATGTCC GGGAGAATTC
```

REFERENCES

Adamidis T., Riggle P. and Champness W. (1990) Mutations in a new *Streptomyces coelicolor* locus which globally block antibiotic biosynthesis but not sporulation. Journal of Bacteriology **172:** 2962-2969.

Adamidis T. and Champness W. (1992) Genetic analysis of *absB*, A *Streptomyces coelicolor* locus involved in global antibiotic production. Journal of Bacteriology **174:** 4622-4628.

Altschul S.F., Gish W., Miller W., Myers E.W. and Lipman D.J. (1990) Basic Local Alignment Search Tool. Journal of Molecular Biology **215:** 403-410.

Bachman B.J. and Low K.B. (1980) Linkage map of *Escherichia coli* K12. Microbiology Reviews **44:** 1-56.

Bascarán V., Sánchez L., Hardisson C. and Braña A.F. (1991) Stringent response and initiation of secondary metabolism in *Streptomyces clavuligerus*. Journal of General Microbiology **137:** 1625-1634.

Bascarán V., Hardisson C. and Braña A.F. (1992) Proline uptake in *Streptomyces clavuligerus*. FEMS Microbiology Letters **69:** 27-30.

Bibb M.J. and Cohen S.N. (1982) Gene expression in *Streptomyces*: Construction and application of promoter-probe plasmid vectors in *Streptomyces lividans*. Molecular and General Genetics **187:** 265-277.

Brady R.A. and Csonka L.N. (1988) Transcriptional regulation of the *proC* gene of *Salmonella typhimurium*. Journal of Bacteriology **170:** 2379-2380.

Bruton C.J., Guthrie E.P. and Chater K.F. (1991) Phage vectors that allow monitoring of transcription of secondary metabolism genes in *Streptomyces*. Biotechnology **9:** 652-656.

Caballero J.L., Martinez E., Malpartida F. and Hopwood D.A. (1991) Organisation and functions of the *act*VA region of the actinorhodin

gene cluster of *Streptomyces coelicolor*. Molecular and General Genetics **230:** 401-412.

Cairney J., Booth I.R. and Higgins C.F. (1985a) *Salmonella typhimurium proP* gene encodes a transport system for the osmoprotectant betaine. Journal of Bacteriology **164:** 1218-1223.

Cairney J., Booth I.R. and Higgins C.F. (1985b) Osmoregulation of gene expression in *Salmonella typhimurium*: *proU* encodes an osmotically induced betaine transport system. Journal of Bacteriology **164:** 1224-1232.

Cashel M. and Rudd K.E. (1987) The stringent response. *Escherichia coli* and *Salmonella typhimurium* cellular and molecular biology. Ed. F.C.Neidhart. American Society for Microbiology.

Chater K.F. (1972) A morphological and genetic mapping study of white colony mutants of *Streptomyces coelicolor*. Journal of General Microbiology **72:** 9-28.

Chater K.F. and Bruton C.J. (1983) Mutational cloning in *Streptomyces* and the isolation of antibiotic production genes. Gene **26:** 67-78.

Chater K.F. and Bruton C.J. (1985) Resistance, regulatory and production genes for the antibiotic methylenomycin are clustered. EMBO **4:** 1883-1892.

Chater K.F., Bruton C.J, Plaskitt K.A., Buttner M.J., Méndez C. and Helmann J.D. (1989) The developmental fate of *S.coelicolor* hyphae depends upon a gene product homologus with the motility σ factor of *B.subtilis*.

Chater K.F. (1993) Genetics of differentiation in *Streptomyces*. Annual Review of Microbiology **47:** 685-713.

Champness W. (1988) New loci required for *Streptomyces coelicolor* morphological and physiological differentiation. Journal of Bacteriology **170:** 1168-1174.

Choi S.H. and Greenberg E.P. (1992) Genetic dissection of DNA binding and luminescence gene activation by the *Vibrio fischeri* LuxR protein. Journal of Bacteriology **174:** 4064-4069.

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

Coco E.A. Narva K.E. and Feitelson J.S. (1991) New classes of *Streptomyces coelicolor* A3(2) mutants blocked in undecylprodigiosin (Red) biosynthesis. Molecular and General Genetics **227:** 28-32.

Csonka L.N. (1981) Proline over-production results in enhanced osmotolerance in *Salmonella typhimurium*. Molecular and General Genetics **182:** 82-86.

Csonka L.N. and Baich A. (1983) Proline Biosynthesis. Amino acids: Biosynthesis and genetic regulation. Editors Hermann K.M. and Somerville R.L., published by Addison-Wesley.

Csonka L.N. and Hanson A.D. (1991) Procaryotic osmoregulation: Genetics and physiology. Annual Review of Microbiology **45:** 469-606.

Dauenhauer S.A., Hull R.A and Williams R.P. (1984) Cloning and expression in *Escherichia coli* of *Serratia marcescens* genes encoding prodigiosin biosynthesis. Journal of Bacteriology **158:** 1128-1132.

Davis N.K. and Chater K.F. (1990) Spore colour in *Streptomyces coelicolor* A3(2) involves the developmentally regulated synthesis of a compound biosynthetically related to polyketide antibiotics. Molecular Microbiology **4:** 1679-1691.

Davis N.K. and Chater K.F. (1992) The *Streptomyces coelicolor whiB* gene encodes a small transcription factor-like protein dispensable for growth but essential for sporulation. Molecular and General Genetics **232:** 351-358.

Decker H. and Hutchinson C.R. (1993) Transcriptional analysis of the *Streptomyces glaucescens* tetracenomycin C biosynthesis gene cluster. Journal of Bacteriology **175:** 3887-3892.

Distler J., Ebert A., Mansouri K., Pissowotski K., Stockmann M. and Piepersburg W. (1987) Gene cluster for streptomycin biosynthesis in *Streptomyces griseus*: nucleotide sequence of three genes and analysis of transcriptional activity. Nucleic acids research **15:** 8041-8056.

Feitelson J.S. and Hopwood D.A. (1983) Cloning of a *Streptomyces* gene for an O-Methyltransferase involved in antibiotic biosynthesis. Molecular and General Genetics **190:** 394-398.

Feitelson J.S., Malpartida F. and Hopwood D.A. (1985) Genetic and biochemical characterisation of the *red* gene cluster of *Streptomyces coelicolor* A3(2). Journal of General Microbiology **131:** 2431-2441.

Fernández-Moreno M.A., Caballero J.L. Hopwood D.A. and Malpartida F. (1991) The *act* cluster contains regulatory and antibiotic export genes, direct targets for translational control by the *bldA* tRNA gene of streptomyces. Cell **66:** 769-780.

Fernández-Moreno M.A., Martinez E., Boto L., Hopwood D.A. and Malpartida F. (1992a) Nucleotide sequence and deduced functions of a set of co-transcribed genes of *Streptomyces coelicolor* A3(2) including the polyketide synthase for the antibiotic actinorhodin. Journal of Biological Chemistry **267:** 19278-19290.

Fernández-Moreno M.A., Martín-Triana A.J., Martínez E., Niemi J., Kieser H.M., Hopwood D.A. and Malpartida F. (1992b) *abaA*, A new pleiotropic regulatory locus for antibiotic production in *Streptomyces coelicolor*. Journal of Bacteriology **174:** 2958-2967.

Flores M.E. and Sánchez S. (1985) Regulation of erythromycin formation in *Streptomyces erythreus*. FEMS Microbiology Letters **26:** 191-194.

Gormley E.P. and Davies J. (1991) Transfer of plasmid RSF1010 by conjugation from *Escherichia coli* to *Streptomyces lividans* and *Mycobacterium smegmatis*. Journal of Bacteriology **173:** 6705-6708.

Gramajo H.C., Takano E. and Bibb M.J. (1993) Stationary-phase production of the antibiotic actinorhodin in *Streptomyces coelicolor* A3(2) is transcriptionally regulated. Molecular Microbiology **7:**837-845.

Guthrie E.P. and Chater K.F. (1990) The level of a transcript required for production of a *Streptomyces coelicolor* antibiotic is conditionally dependent on a tRNA gene. Journal of Bacteriology **172:** 6189-6193.

Guthrie E.P. and Chater K.F. (1995) in press.

Hara O., Horinouchi S., Uozumi T. and Beppu T. (1983) Genetic analysis of A-factor synthesis in *Streptomyces coelicolor* A3(2) and *Streptomyces griseus*. Journal of General Microbiology **129:** 2939-2944.

Hallam S.E., Malpartida F. and Hopwood D.A. (1988) Nucleotide sequence, transcription and deduced function of a gene involved in polyketide antibiotic synthesis in *Streptomyces coelicolor*. Gene **74:** 305-320.

Helmann J.D. and Chamberlin M.J. (1988) Structure and function of bacterial sigma factors. Annual review of Biochemistry **57:** 839-872.

Herbert D., Phipps P.J. and Strange R.E. (1971) Chemical analysis of microbial cells. Chapter 3 Methods in Microbiology volume 5B: 309-344. Editors J.R. Norris and D.W.Ribbons, Academic Press.

Hobbs G., Frazer C.M., Gardner D.C.J., Flett F. and Oliver S.G. (1990) Pigmented antibiotic production by *Streptomyces coelicolor* A3(2): kinetics and the influence of nutrients. Journal of General Microbiology **136:** 2291-2296.

Hobbs G., Obanye A.I.C., Petty J., Mason C., Barratt E., Gardner D.C.J., Flett F., Smith C.P., Broda P. and Oliver S.G. (1992) An integrated approach to studying regulation of production of the antibiotic

methylenomycin by *Streptomyces coelicolor* A3(2). Journal of Bacteriology **174:** 1487-1494.

Hodgson D.A. (1992) Differentiation in actinomycetes. Prokaryotic structure and function: A new perspective. Society for General Microbiology Symposium Volume **47:** 407-440. Published by Cambridge University Press. Editors S.Mohan, C.Dow and J.A.Cole.

Hong S-K., Kito M., Beppu T. and Horinouchi S. (1991) Phosphorylation of the AfsR product, a global regulatory protein for secondary metabolite formation in *Streptomyces coelicolor* A3(2). Journal of Bacteriology **173:** 2311-2318.

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

Hopwood D.A., Bibb M.J., Chater K.F., Kieser T., Bruton C.J., Kieser H.M., Lydiate D.J., Smith C.P., Ward J.M. and Schrempf H. (1985) Genetic manipulation of streptomyces: A laboratory manual.

Hopwood D.A., Bibb M., Chater K.F., Janssen G.R., Malpartida F. and Smith C.P. (1986) Regulation of gene expression in antibiotic producing Streptomyces. SGM

Hopwood D.A., Chater K.F. and Bibb M.J. Genetics of antibiotic production in *Streptomyces coelicolor* A3(2), a model streptomycete. Genetics and biochemistry of antibiotic production, p 71-107. Editor L.C.Vining, publishers Butterworth-Heinmann USA.

Horinouchi S. and Beppu T. (1992) Autoregulatory factors and communication in Actinomycetes. Annual Review of Microbiology **46:** 377-398.

Horinouchi S. and Beppu T. (1994) A-factor as a microbial hormone that controls cellular differentiation and secondary metabolism in *Streptomyces griseus*. Molecular Microbiology **12:** 859-864.

Horinouchi S., Hara O. and Beppu T. (1983) Cloning of a pleiotropic gene that positively controls biosynthesis of A-factor, actinorhodin and prodigiosin in *Streptomyces coelicolor* A3(2) and *Streptomyces lividans*. Journal of Bacteriology **155:**1238-1248.

Horinouchi S., Kito M., Nishiyama M., Furuya K., Hong S-K. Miyake K. and Beppu T. (1990) Primary structure of AfsR, a global regulatory protein for secondary metabolite formation in Streptomyces coelicolor A3(2). Gene **95:** 49-56.

Horinouchi S., Malpartida F., Hopwood D.A. and Beppu T. (1989) *afsB* stimulates transcription of the actinorhodin biosynthetic pathway in *Streptomyces coelicolor* A3(2) and *Streptomyces lividans*. Molecular and General Genetics **215:** 355-357.

Horinouchi S., Suzuki H. and Beppu T. (1986) Nucleotide sequence of *afsB*, a pleiotropic gene involved in secondary metabolism in *Streptomyces coelicolor* A3(2) and "*Streptomyces lividans*". Journal of Bacteriology **168:** 257-269.

Hutchinson R., Decker H., Madduri K., Otten S.L. and Tang L. (1993) Genetic control of polyketide biosynthesis in the genus *Streptomyces*. Antonie van Leeuwenhoek **64:** 165-176.

Ishizuka H., Horinouchi S., Kieser H.M. Hopwood D.A. and Beppu T. (1992) A putative two-component regulatory system involved in secondary metabolism in *Streptomyces* spp. Journal of Bacteriology **174:** 7585-7594.

Jaurin B. and Cohen S.N. (1984) *Streptomyces lividans* RNA polymerase recognises and uses *Escherichia coli* transcriptional signals. Gene **28:** 83-91.

Jones S., Yu B., Bainton N.J., Birdsall M., Bycroft B.W., Chhabra S.R., Cox A.J., Golby P., Reeves P.J., Stephens S., Winson M.K., Salmond G.P.C., Stewart G.S.A.B and Williams P. (1993) The *lux* autoinducer regulates the production of exoenzyme virulence determinants in *Erwinia carotovora* and *Pseudomonas aeruginosa*. EMBO **12:** 2477-2482.

Kieser H.M., Kieser T and Hopwood D.A. (1992) A combined genetic and physical map of the *Streptomyces coelicolor* A3(2) chromosome. Journal of Bacteriology **174:** 5496-5507.

Killham K. and Firestone M.K. (1984a) Salt stress control of intracellular solutes in *Streptomyces* indigenous to saline soils. Applied and Environmental Microbiology **47:** 301-306.

Killham K. and Firestone M.K. (1984b) Proline transport increased growth efficiency in salt-stressed *Streptomyces griseus*. Applied and Environmental Microbiology **48:** 239-241.

Kinashi H., Shimaji-Murayama M. and Hanafusa T. (1992) Integration of SCP1, a giant linear plasmid, into the *Streptomyces coelicolor* chromosome. Gene 115: 35-41.

Kirby R. and Hopwood D.A. (1977) Genetic determination of methylenomycin synthesis by the SCP1 plasmid of *Streptomyces coelicolor* A3(2). Journal of General Microbiology **98:** 239-252.

Laity C., Chater K.F., Lewis C.G. and Buttner M. (1993) Genetic analysis of the ØC31-specific phage growth limitation (Pgl) system of *Streptomyces coelicolor* A3(2). Molecular Microbiology **7:** 329-336.

Lakey J.H., Lea E.J.A., Rudd B.A., Wright H.M. and Hopwood D.A. (1983) A new channel-forming antibiotic from *Streptomyces coelicolor* A3(2) which requires calcium for its activity. Journal of General Microbiology **129:** 3565-3573.

Lawlor E.J., Baylis H.A. and Chater K.F. (1987) Pleiotropic morphological and antibiotic deficiencies result from mutations in a gene encoding a tRNA-like product in *Streptomyces coelicolor* A3(2). Genes and Development **1:**1305-1310.

Leigh M.J. (1992) Peptide transport in *Streptomyces coelicolor* A3(2). Ph.D Thesis, University of Warwick, England.

Leskiw B.K., Bibb M.J. and Chater K.F. (1991a) The use of a rare codon specifically during development? Molecular Microbiology **5:** 2861-2867.

Leskiw B.K., Lawlor E.J., Fernandez-Abalos J.M. and Chater K.F. (1991b) TTA codons in some genes prevents their expression in a class of developmental, antibiotic-negative, *Streptomyces* mutants. Proceedings of the National Academy of Science USA **88:** 2461-2465.

Leskiw B.K., Mah R., Lawlor E.J. and Chater K.F. (1993) Accumulation of *bldA*-specified tRNA is temporally regulated in *Streptomyces coelicolor* A3(2). Journal of Bacteriology **175:** 1995-2005.

Lewin B. (1990) Genes IV. Oxford University Press.

Lin Y-S., Kieser H.M., Hopwood D.A. and Chen C.W. (1993) The chromosomal DNA of *Streptomyces lividans* 66 is linear. Molecular Microbiology **10:** 923-933.

Lomoskaya N.D., Chater K.F. and Mkrtumain N.M (1980) Genetics and molecular biology of *Streptomyces* bacteriophages. Microbiological Reviews **44:** 206-229.

Ludovice M., Martin J.F., Carrachas P. and Liras P. (1992) Characterisation of the *Streptomyces clavuligerus argC* gene encoding N-acetylglutamyl-phosphate reductase: Expression in *Streptomyces lividans* and effect on clavulanic acid production. Journal of Bacteriology **174:** 4606-4613.

Lydiate D.J., Malpartida F. and Hopwood D.A. (1985) The *Streptomyces* plasmid SCP2*: its function analysis and development into useful cloning vectors. Gene **35:** 223-235.

Ma H. and Kendall K. (1994) Cloning and analysis of a gene cluster from *Streptomyces coelicolor* that causes accelerated aerial mycelium formation in *Streptomyces lividans*. Journal of Bacteriology **176:** 3800-3811.

Mahan M.J. and Csonka L.N. (1983) Genetic analysis of the *proAB* genes of *Salmonella typhimurium*: Physical and genetic analyses of the cloned $proB^+$ $proA^+$ gene of *Escherichia coli* and of a mutant allele that confers proline overproduction and enhanced osmotolerance. Journal of Bacteriology **156:** 1249-1262.

Maloy S. (1987) Proline utilisation operon. *Escherichia coli* and *Salmonella typhimurium* cellular and molecular biology. Ed. F.C.Neidhardt. American Society for Microbiology.

Malpartida F. and Hopwood D.A. (1984) Molecular cloning of the whole biosynthetic pathway of a *Streptomyces* antibiotic and its expression in a heterologous host. Nature **309:** 462-464.

Malpartida F., Niemi J., Navarrete R. and Hopwood D.A. (1990) Cloning and expression in a heterologous host of the complete set of genes for biosynthesis of the *Streptomyces coelicolor* antibiotic undecylprodigiosin. Gene **93:** 91-99.

Marsh J.L., Erfle M. and Wykes E.J. (1984) The pIC plasmid and phage vectors with versatile cloning sites for recombinant selection by insertional inactivation. Gene **32:** 481-485.

Masurekar P. and Sosa M. (1988) A selection technique based on the resistance to the analogues of threonine for the isolation of improved producers of avermectin. Abstracts of the Annual Meeting of the American Society for Microbiology . Abstract O-70, p273.

Maxam A.M. and Gilbert W. (1980) Sequencing and end-labelling DNA with base specific chemical cleavages. Methods in Enzymology **65:** 449-560.

May W. and Formica J.V. (1978) Glutamate induced uptake of proline by *Streptomyces antibioticus*. Journal of Bacteriology **134:** 546-554.

Mazodier P., Petter R and Thompson C. (1989) Intergeneric conjugation between *Escherichia coli* and *Streptomyces* species. Journal of Bacteriology **171:** 3583-3585.

McGowan S., Sebaiha M., Jones S., Yu B, Bainton N., Chan P.F., Bycroft B., Stewart G.S.A.B., Williams P. and Salmond G.P.C. (1995) Carbepenem antibiotic production in *Erwinia carotovora* is regulated by CarR, a homologue of the LuxR transcriptional activator. Microbiology **141:** 541-550.

Merrick M.J. (1976) A morphological and genetic mapping study of bald colony mutants of *Streptomyces coelicolor*. Journal of General Microbiology **96:** 299-315.

Miyake K., Horinouchi S., Yoshida M., Chiba N., Mori K., Nogawa N., Morikawa N. and Beppu T. (1989) Detection and properties of A-factor binding protein from *Streptomyces griseus*. Journal of Bacteriology **171:** 4298-4302.

Miyake K., Kuzuyama T., Horinouchi S. and Beppu T. (1990) The A-factor binding protein of *Streptomyces griseus* negatively controls streptomycin production and sporulation. Journal of Bacteriology **172:** 3003-3008.

Msadek T., Kunst F., Henner D., Klier A., Rapoport G. and Dedonder R. (1990) Signal transduction pathway controlling synthesis of a class of degradative enzymes in *Bacillus subtilis*: Expression of the regulatory genes and analysis of mutations in *degS* and *degU*. Journal of Bacteriology **172:** 824-834.

Murakami T., Anzai H., Imai S., Satoh A., Nagaoka K. and Thompson C.J. (1986) The bialaphos biosynthetic genes of *Streptomyces hygroscopicus*: Molecular cloning and characterisation of the gene cluster. Molecular and General Genetics **205:** 42-50.

Narva K.E. and Feitelson J.S. (1990) Nucleotide sequence and transcriptional analysis of the *redD* locus of *Streptomyces coelicolor* A3(2). Journal of Bacteriology **172:** 326-333.

Ninfa A.J., Gottlin Ninfa E., Lupas A.N., Stock A., Magasanik B. and Stock J. (1988) Crosstalk between bacterial chemotaxis signal transduction proteins and regulators of transcription of the Ntr regulon: Evidence that nitrogen assimilation and chemotaxis are controlled by a common phosphotransfer mechanism. Proceedings of the National Academy of Science USA **85:** 5492-5496.

Novák J., Hájek P., Rezanka T. and Vanek Z. (1992) Nitrogen regulation of fatty acids and avermectins biosynthesis in *Streptomyces avermitilis*. FEMS Microbiology Letters **93:** 57-62.

Ochi K. (1986) Occurrence of the stringent response in *Streptomyces* sp. and its significance for the initiation of morphological and physiological differentiation. Journal of General Microbiology **132:** 2621-2631.

Omura S., Tanaka Y., Mamada H. and Masuma R. (1984) Effect of ammonium ion, inorganic phosphate and amino acids on the biosynthesis of protylonolide, a precursor of tylosin aglycone. Journal of Antibiotics **37:** 494-502.

Omura S. and Tanaka Y. (1986) Biosynthesis of tylosin and its regulation by ammonium and phosphate. Regulation of secondary metabolite formation, Editors H. Kleinhauf, H. von Doren, H. Dornauer and G. Nesemann. Publishers VCH.

Otten S.L., Ferguson J. and Hutchinson C.R. (1995) Regulation of Daunorubicin production in *Streptomyces peucetius* by the dnrR2 locus. Journal of Bacteriology **177:** 1216-1224.

Paget M.S.B., Hintermann G. and Smith C.P. (1994) Construction and application of Streptomycete promoter probe vectors which employ the *Streptomyces glaucescens* tyrosinase-encoding gene as reporter. Gene 146: 105-110.

Parkinson J.S. and Kofoid E.C. (1992) Communication modules in bacterial signalling proteins. Annual Review of Genetics **26:** 71-112.

Passador L., Cook J.M., Gambello M.J., Rust L. and Iglewski B. (1993) Expression of *Pseudomonas aeruginosa* virulence genes requires cell-to-cell communication. Science **260:** 1127-1130.

Piper K.R., Beck von Bodman S., Farrand S.K. (1993) Conjugation factor of *Agrobacterium tumefaciens* regulates Ti plasmid transfer by autoinduction. Nature **362:** 448-450.

Rudd B.A.M. and Hopwood D.A. (1979) Genetics of actinorhodin biosynthesis by *S.coelicolor* A3(2). Journal of General Microbiology **114:** 35-43.

Rudd B.A.M and Hopwood D.A. (1980) A pigmented mycelial antibiotic in *Streptomyces coelicolor*: Control by a chromosomal gene cluster. Journal of General Microbiology **119:** 333-340.

Salmond G.P.C., Bycroft B.W., Stewart G.S.A.B. and Williams P. (1995) The bacterial 'enigma': cracking the code of cell-cell communication. Molecular Microbiology **16:** 615-624.

Sambrook J., Fritsch E.F. and Maniatis T. (1989) Molecular cloning: A laboratory manual. Second Edition, Cold Spring Harbour Press.

Sanger F., Nicklen S. and Coulsen A.R. (1977) Sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Science U.S.A. **74:** 5463-5468.

Schlösser A., Meldorf M., Stumpe S., Bakker E.P. and Epstein W. (1995) TrkH and its homologue, TrkG, determine the specificity and kinetics of cation transport by the Trk system of *Escherichia coli*. Journal of Bacteriology **177:** 1908-1910.

Schottel J.L., Bibb M.J. and Cohen S.N. (1981) Cloning and expression in *Streptomyces lividans* of antibiotic resistance genes derived from *Escherichia coli*. Journal of Bacteriology **146:** 360-368.

Schrempf H. and Goebel W. (1977) Characterisation of a plasmid from *Streptomyces coelicolor* A3(2). Journal of Bacteriology **131:** 251-258.

Smith D.D.S., Wood N.J. and Hodgson D.A. (1995) Interaction between primary and secondary metabolism in *Streptomyces coelicolor* A3(2): Role of Pyrroline-5-carboxylate dehydrogenase. Microbiology **141:** 1739-1744.

Smith M.C.M., Ingram C.J., Owen C.E. and Wood N.T. (1992) Gene expression in the *Streptomyces* temperate phage ∅C31. Gene **115:** 43-48.

Stein D. and Cohen S.N. (1989) Cloned regulatory gene of *Streptomyces lividans* can suppress the pigment deficiency phenotype of different developmental mutants. Journal of Bacteriology **171:** 2258-2261.

Stock J.B., Ninfa A.J. and Stock A.M. (1989) Protein phosphorylation and regulation of adaptive responses in bacteria. Microbiological reviews **53:** 450-490.

Strauch E., Takano E., Baylis H.A. and Bibb, M..J. (1991) The stringent response in *Streptomyces coelicolor* A3(2). Molecular Microbiology **5:**289-298.

Stutzman-Engwall K.J., Otten S.L. and Hutchinson C.R. (1992) Regulation of secondary metabolism in *Streptomyces* spp. and overproduction of

daunorubicin in *Streptomyces peucetius*. Journal of Bacteriology **174:** 144-154.

Takano E., Gramajo H.C., Strauch E., Andres A., White J. and Bibb M. (1992) Transcriptional regulation of the *redD* transcriptional activator gene accounts for growth-phase-dependent production of the antibiotic undecylprodigiosin in *Streptomyces coelicolor* A3(2). Molecular Microbiology **6:** 2797-2804.

Townsend D.E. and Wilkinson B.J. (1992) Proline transport in *Staphlococcus aureus*: a high affinity system and a low affinity system involved in osmoregulation. Journal of Bacteriology **174:** 2702-2710.

Tsao S.W., Rudd B.A.M., He X.G., Chang C.J and Floss H.G. (1985) Identification of a red pigment from *Streptomyces coelicolor* A3(2) as a mixture of prodigiosin derivatives. Journal of Antibiotics **38:** 128-131.

Tseng H.C. and Chen C.W. (1991) A cloned *ompR*-like gene of *Streptomyces lividans* 66 suppresses defective *melC1*, a putative copper-transfer gene. Molecular Microbiology **5:** 1187-1196.

Ueda K., Miyake K., Horinouchi S. and Beppu T. (1993) A gene cluster involved in aerial mycelium formation in *Streptomyces griseus* encodes proteins similar to the response regulators of two-component regulatory systems and membrane translocators. Journal of Bacteriology **175:** 2006-2016.

Ukti S.A., Lash T.D. and Wilkinson B.J. (1995) Proline betaine is a highly effective osmoprotectant for *Staphlococcus aureus*. Archives of Microbiology **163:** 138-142.

Verbruggen v., Villaroel R. and Montagu M.V. (1993) Osmoregulation of a Pyrroline-5-carboxylate reductase gene in *Arabidopsis thaliana*. Plant Physiology **103:** 771-781.

Vivian A. (1971) Genetic control of fertility in *Streptomyces coelicolor* A3(2): Plasmid involvement in the interconversion of UF and IF strains. Journal of General Microbiology **69:** 353-364.

Vögtli M., Chang P-C. and Cohen S.N. (1992) afsR2: a previously undetected gene encoding a 63-amino acid protein that stimulates antibiotic production in *Streptomyces lividans*. Molecular Microbiology **14:** 643-653.

Vujaklija D., Horinouchi S. and Beppu T. (1993) Detection of an A-factor responsive protein that binds to the upstream activation sequence of *strR*, a regulatory gene for streptomycin biosynthesis in *Streptomyces griseus*. Journal of Bacteriology 175: 2652-2661.

Wanner B.L. (1992) Is cross regulation by phosphorylation of two-component response regulator proteins important in Bacteria? Journal of Bacteriology **174:** 2053-2058.

Wasserman H.H., Shaw C. and Sykes R.J. (1974) The biosynthesis of metacycloprodigiosin and undecylprodigiosin. Tetrahedron Letters **33:** 2787-2790.

Williams R.P. and Quadri S.M.H. (1980) The pigment of *Serratia*, p31-75. The genus *Serratia*. Editors A. von Graevenitz and S.J.Rubin, published by CRC Press.

Williams S.T., Goodfellow M., Wellington E.M.H., Vickers J.C., Alderson G., Sneath P.H.A. and Sackin M.J. (1983) Numerical classification of *Streptomyces* and related genera. Journal of General Microbiology **129:** 1743-1813.

Yanisch-Perron C., Vieira J. and Messing J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:** 103-119.

Zhang L., Murphy P.J., Kerr A. and Tate M.E. (1993) *Agrobacterium* conjugation and gene regulation by *N*-acyl-L-homoserine lactones. Nature **362:** 446-450.