Jefferies, Johanna M.C. (1998) *Functional studies of the otrB gene from Streptomyces rimosus.* PhD thesis.

http://theses.gla.ac.uk/2079/

Copyright and moral rights for this thesis are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Glasgow Theses Service
http://theses.gla.ac.uk/
theses@gla.ac.uk

# Functional studies of the *otrB* gene from *Streptomyces rimosus*

by

JOHANNA. M. C. JEFFERIES

A thesis submitted for the degree of Doctor of Philosophy

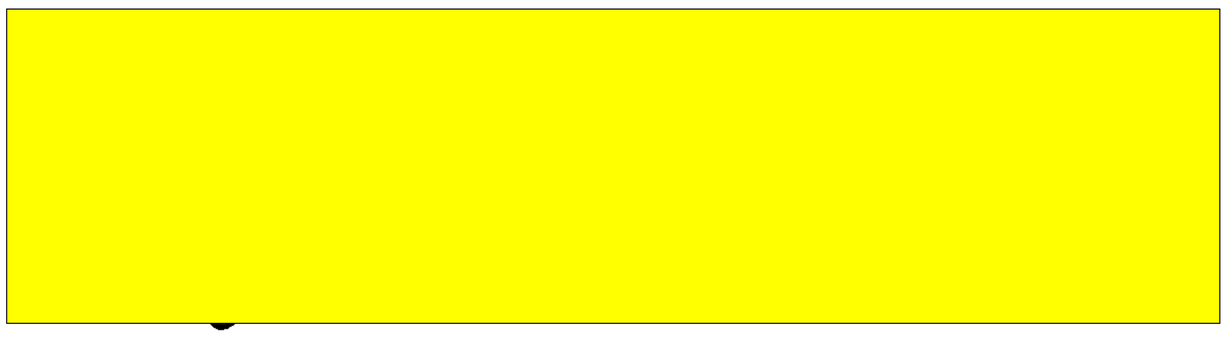
Divsion of Molecular Genetics,
Institute of Biomedical and Life Sciences
University of Glasgow

April 1998

The research reported in this thesis is my own original work except where otherwise stated and has not been submitted for any other degree

Johanna. M. C. Jefferies

April 1998



# PAGE NUMBERING AS ORIGINAL

**Dedicated to my family and to Rab with thanks for all their support.**

## ACKNOWLEDGEMENTS

This work was supported by a grant from the BBSRC.

Many thanks to my supervisor, Iain Hunter for all the guidance and encouragement given to me whilst working for and writing up this thesis and also for putting up with my constant worrying and moaning.

I would like to thank Marion, Sue,Vera and Anne for being great technicians and not nagging me too much about the state of my bench or the absence of a labcoat.

Cheers to all other "Streppies" past and present, especially Sue, Al, Natasa. P, Petko and Teresa for listening to me talk to myself and swear at inanimate objects, and for the invaluable tips and discussions.

Thanks also goes to everyone who has made my life more fun during the last few years at Glasgow and at Strathclyde, Mo, Ed, Pete, Tara, Sue (again!) thanks for the friendships (and all the nights out!).

Thanks also goes to Roth for helping me with sequencing and to Chris and everyone in Peter Henderson's Lab in Leeds for giving me a hand with the membrane preps. I would also like to thank Nigel Watson for letting me have time off to finish this thesis!

Finally but absolutely not leastly, a special thanks to Rab for being the one who has put up with me most of all, during the highs and more importantly the lows of being a Ph.D. student; Cheers pal!

## ABSTRACT

Oxytetracyline, (OTC), is a secondary metabolite antibiotic produced by the actinomycete *Streptomyces rimosus*. All of the structural genes for OTC synthesis are clustered on the *S. rimosus* chromosome and flanked by resistance determinants, *otrA* and *otrB*. *OtrA* mediates resistance through non-covalent modification of the bacterial ribosome whilst *otrB*, the subject of this thesis acts by efflux of the drug from the cell.

At the outset of this thesis *otrB* had previously been cloned and sequenced. During the course of the work some discrepancies with previously published data were found, parts of the gene were then re-sequenced and a revised *otrB* sequence submitted to the NCBI database.

The OtrB protein is a transmembrane protein belonging the Major facilitator Superfamily (MFS). Sequence comparison with other members of the family shows OtrB to be part of a subfamily of antibiotic transporters and multi drug resistance proteins made up from 14 transmembrane helices (6+2+6) arrangement. OtrB contains many conserved motifs typical of the subfamily and of the MFS.

*otrB* was cloned into *E. coli* and expressed. The functional activity of the cloned gene was assessed by growth on various concentrations of OTC. Substrate specificity was investigated using TET and CTC in the growth media.

Isolation of OtrB from *E. coli* as a polyhistidine fusion was attempted, reasons why this was not successful are discussed.

Tetracycline exporters from Gram-negative bacteria generally contain 12 transmembrane helices (6+6). The relevance of the two "extra" helices present in OtrB was investigated by the construction and expression in *E. coli*, of a deletion mutant in which the two putative central helices were absent.

# CONTENTS


| LIST OF CONTENTS | PAGE |
|---|---|
| Title | i |
| Copyright | ii |
| Dedication | iii |
| Acknowledgements | iv |
| Abstract | v |
| List of contents | vi |
| List of figures | xii |
| List of tables | xiv |
| Abbreviations | xv |
| **Chapter 1 Introduction** | |
| 1.1 General Introduction | 2 |
| 1.2 Secondary metabolites | 2 |
| 1.2.1 Function & evolution of secondary metabolites | 2 |
| 1.2.2 Control of Secondary metabolism | 5 |
| 1.2.2.1 Transcriptional regulation | 6 |
| 1.2.2.2 A-factor | 6 |
| 1.2.2.3 A-factor in *S. griseus* | 6 |
| 1.2.2.4 A-factor receptor protein | 6 |
| 1.2.2.5A-factor-dependent binding protein | 7 |
| 1.2.2.6 StrR | 8 |
| 1.2.2.7 A model for the regulation of Streptomycin biosynthesis | 8 |
| 1.2.2.8 A-factor in *S. coelicolor* | 10 |
| 1.2.2.9 AbsA and AbsB | 10 |
| 1.2.2.10 AfsR (AfsB) | 10 |
| 1.2.2.11 AfsK | 11 |
| 1.2.2.12 AfsR2 | 11 |
| 1.2.2.13 AfsQ1-Q2 | 12 |

1.2.2.14 AbaA 13
1.2.2.15 Sigma factors 13
1.2.2.16 $\sigma^{WhiG}$ 13
1.2.2.17 *hrd* σ factor 14
1.2.2.18 σF 16
1.2.3. Post-transcriptional regulation 17
1.2.3.1 Regulation by the use of a rare tRNA 17
1.2.3.2 Regulation by ppGpp and pppGpp 20
1.2.3.3 (p)ppGpp and antibiotic biosynthesis 21
1.2.3.4 (p)ppGpp and morphological differentiation 22
1.3 Polyketides 22
1.3.1 What are polyketides? 22
1.3.2 Polyketide biosynthesis 25
1.3.2.1 Polyketide synthase 25
1.3.2.2 The Acyl Carrier Protein (ACP) 30
1.4 The OTC cluster 30
1.5 Antibiotic resistance 39
1.5.1 Modification of drug molecules 39
1.5.2 Modification of target sites 40
1.5.3 Synthesis of novel resistant cellular compounds 41
1.5.4 Export of the drug 41
1.6 Tetracyclines and Tetracycline resistance 42
1.6.1 Tetracyclines 42
1.6.2 Resistance to tetracyclines 45
1.6.2.1Tetracycline efflux 45
1.6.2.2 Gram-negative TetA proteins, structure function studies 47
1.6.2.3 Gram-positive Tet proteins 50
1.6.2.4 Tetracycline resistance by ribosomal modification 51
1.6.2.5 Tetracycline modification 51
1.7 The Major Facilitator Superfamily (MFS) 52

**Chapter 2 Materials & methods** 59

2.1 Introduction 60

2.2 Bacterial strains and vectors 60

2.2.1 Chemicals & Biochemicals 63

2.2.2 Enzymes & Kits 64

2.3 Standard media & microbiological techniques 64

2.3.1 Media used for the growth of *E. coli* 64

2.3.2 Media used for the growth of *Streptomyces* 65

2.3.3 Sterilisation of media 65

2.3.4 Antibiotics & indicators 66

2.3.5 Growth of *Streptomyces* mycelia 66

2.3.5.1 Growth of mycelia in liquid 66

2.3.5.2 Harvesting of mycelia 66

2.3.6 Growth of *E. coli* in liquid media 66

2.3.6.1 Expression of cloned genes in *E. coli* using IPTG 67

2.3.7 Preparation of *Streptomyces* spores 67

2.3.7.1 Preparation and preservation of spore suspensions 67

2.3.7.2 Spore counts 68

2.3.8 Preservation of *E. coli* strains 68

2.3.9 Introduction of plasmid DNA into *E coli* 68

2.3.9.1 Preparation of competent cells 68

2.3.9.2 Transformation procedures 69

2.3.9.3 Selection of pUC derived recombinants 70

2.3.10 Introduction of plasmid into *Streptomyces* 71

2.3.10.1 Protoplast method 71

2.3.10.2 Preparation of protoplasts 71

2.3.10.3 Transformation of protoplasts 72

2.3.10.4 Regeneration of transformed protoplasts 72

2.4 General DNA methods 72

2.4.1 Commonly used buffers 72

2.4.2 Preparation of plasmid DNA 73

2.4.2.1 Reagents for the isolation of plasmid DNA 73
2.4.2.2. Preparation of plasmid DNA from *E. coli* 73
2.4.2.3 Preparation of plasmid DNA by the Keiser method 74
2.4.3 Precipitation of DNA from solution 75
2.4.4 Spectrophotometric measurement of nucleic acids 75
2.4.5 Digestion of DNA with restriction enzymes 75
2.4.6 Ligation of DNA fragments 76
2.4.7 Removal of 5' phosphate from linearised DNA 76
2.4.8 Filling in of overhanging bases resulting from restriction digest 76
2.4.9 Addition of linkers to blunt ended DNA fragments 77
2.4.10 Removal of excess linkers 77
2.4.11 Agarose gel electrophoresis 77
2.4.11.1 Preparation of agarose gels 77
2.4.11.2 Staining of agarose gels 78
2.4.11.3 Photography of agarose gels 78
2.4.12 Recovery of DNA fragments from agarose gels 78
2.4.13 DNA sequencing 79
2.4.13.1 Preparation of plasmid template for sequencing 79
2.4.13.2 Preparation of sequencing primers 79
2.4.14 Polymerase Chain Reaction 79
2.4.14.1 PCR primers 80
2.5 General Protein methods 81
2.5.1 Preparation of crude protein extracts 81
2.5.2 Preparation of membrane fractions from *E coli* 81
2.5.2.1 Cell disruption 81
2.5.2.2 Solubilisation of membranes 81
2.5.3 Isolation of HisTagged fusion protein 82
2.5.3.1 Preparation of resin 82
2.5.3.2 Binding of protein to resin 82
2.5.3.3 Column purification of fusion protein 82
2.5.4 Protein assays 83
2.5.4.1 BCA protein assay 83

2.5.4.2 Biorad DC protein assay 83
2.5.5 Denaturing polyacrylamide gel electrophoresis (SDS PAGE) 83
2.5.5.1 Reagents for SDS PAGE 83
2.5.5.2 Apparatus for SDS PAGE 84
2.5.5.3 Protocol for SDS PAGE 84
2.5.6 Staining of polyacrylamide gels 85
2.5.7 Western blotting 85
2.5.7.1 Transfer buffer 85
2.5.7.2 Apparatus for transfer 86
2.5.7.3 Protocol for transfer 86
2.5.8 Immunodetection 87
2.5.8.1 Reagents for immunodetection 87
2.5.8.2 Protocol for immunodetection 87
2.6 Computer methods 88

**Chapter 3 The *otrb* gene and gene product** 89

3.1 Introduction` 90
3.2 Aims of the Chapter 91
3.3 Re-sequencing of parts of *otrB* 91
3.3.1 Sequencing strategy 91
3.3.2 Analysis of sequence data 95
3.4 Comparison of sequence data for otrB 96
3.4.1 Outcome of initial sequence analysis 99
3.5 Second stage sequencing strategy 100
3.5.1 Primer design 100
3.5.2 Sequencing results 100
3.5.3 BIBBPROGAC analysis 101
3.6 Peptide analysis 101
3.6.1 Hydropathy profile of OtrB 109
3.6.2 Major Facilitator Superfamily (MFS) 109
3.6.3 Conserved motifs 118
3.6.4 Relationship to other MFS members 121

3.7 Discussion 125
3.7.1 Discussion of sequencing results 125
3.7.2 Discussion of OtrB within the Major Facilitator Superfamily 125
3.8 Future work 128

**Chapter 4, Expression of *otrB* in *Streptomyces* and *E. coli*** 130
4.1 Introduction 131
4.2 Aims of the Chapter 133
4.3 Initial expression of *otrB* in *Streptomyces* 133
4.3.1 Plasmid constructs containing *otrB* 133
4.3.2 Expression of *otrB* in *S. rimosus* strain 15883s 134
4.3.3 Expression of *otrB* in *S. albus* 134
4.4 Expression of *otrB* in *E. coli* 135
4.4.1 The T7 expression system 135
4.4.2 Construction of expression plasmid pJJ4 136
4.4.3 Expression of pJJ4 and pJJ5 137
4.5 Discussion 143
4.5.1 Discussion of OTC resistance levels conferred by pJJ4 143
4.5.2 Discussion of OTC resistance levels conferred by pJJ5 144
4.5.3 Discussion of specificity tests 145

**Chapter 5; Construction of otrB-polyhistidine fusions and deletion of two central transmembrane helices** 147
5.1 Introduction 148
5.2 Construction of OtrB fusion plasmids 150
5.2.1 Construction of intermediate construct, pJJ7 150
5.2.2 Construction of pJJ8 and pJJ9 153
5.2.3 Construction of pJJ18 153
5.3 Levels of OTC resistance 153
5.3.1 OTC resistance conferred by pJJ8 and pJJ9 153
5.3.2 OTC resistance conferred by pJJ18 155
5.4 Attempts to isolate HisTagged fusion product 158

5.5 Attempts to detect the otrB fusion protein using immunodetection 163
5.6 Discussion 165
5.6.1 Discussion of pJJ8 and pJJ9 results 165
5.6.2 Discussion of pJJ18 results 165
5.7 Deletion of the two putative central helices 166
5.8 Construction of truncated *otrB* gene 166
5.8.1 Construction of pJJ10 and pJJ11 166
5.8.2 Design of PCR primers 167
5.9 Construction of pJJ12, pJJ14 and pJJ17 166
5.10 Expression and resistance results 172
5.10.1 Expression of pJJ14 and pJJ17 172
5.11 Discussion of OTC resistance conferred by truncated OtrB 176
5.12 Lack of detection of OtrB-His 178
5.13 Further sequencing of pJJ8 179
5.13.1 Discussion of sequencing results 179
5.14 Future work 182

**Chapter 6 General Discussion** 183
6.1 Introduction 184
6.2 Expression of *otrB* in *Streptomyces* 184
6.3 Expression of *otrB* in *E. coli* 185
6.4 Construction, Expression in *E. coli* and attempted purification of a HisTagged OtrB fusion protein 186
6.5 Construction and expression in *E. coli* of a partial deletion of *otrB* 187
6.6 Sequencing and protein structure analysis 187
6.7 Future Work 188

Bibliography 189

# List of Figures


**Fig 1.1** The life cycle of *S. coelicolor* 3

**Fig 1.2** Model for Streptomycin biosynthesis in *S. coelicolor* 9

**Fig 1.3** Heirarchy of sporulation controlling genes in *S. coelicolor* 18

**Fig.1.4** Polyketide metabolites 24

**Fig 1.5** Fatty acid biosynthesis 27

**Fig 1.6** Organisation of PKS genes 29

**Fig 1.7** OTC biosynthesis pathway 33

**Fig 1.8** Organisation of OTC cluster 36

**Fig 1.9** Tetracycline and analogues of tetracycline 44

**Fig 1.10** Two dimensional model of TET resistance protein 49

**Fig 1.11** Early phylogenetic relationships between drug resistance proteins 55

**Fig 1.12** Phylogenetic tree for the MFS 56

**Fig 1.13** Phylogenetic trees for the 12 and 14 TMH sub-families 57

**Fig 3.1a** PCR and cloning strategy for *otrB* 93

**Fig 3.1b** Plan of work for Chapter 3 94

**Fig 3.2** Sequence comparison of pJJ6 sequence with pJJIBI sequence 96

**Fig 3.3** Sequence comparison pJJ6 sequence with *tet347* sequence 97

**Fig 3.4** Sequence comparison of pJJ6 sequence with MacGregor-Pryde sequence 98

**Fig 3.5** Design of *otrB* sequencing primers 102

**Fig 3.6** BIBBPROGAC analysis of revised *otrB* sequence 104

**Fig 3.7** Revised *otrB* nucleotide sequence and amino acid translation 105

**Fig 3.8** Hydropathy plot for OtrB 108

**Fig 3.9** OtrB putative transmembrane regions 110

**Fig 3.10** PILEUP of OtrB with 8 other Streptomyces drug exporters 111

**Fig 3.11** PILEUP of OtrB with 2 TET efflux proteins from Gram-positive bacteria 114

**Fig 3.12** PILEUP of OtrB with 6 TET efflux proteins from Gram-negative bacteria 116

**Fig 4.1** pIJ702, pGLW12 & pGLW81 132

**Fig 4.2** Construction of pJJ1-pJJ4 138

**Fig 4.3** Construction of pJJ5 140

**Fig 4.4** OTC resistance profiles of pJJ4 141

**Fig 4.5** OTC resistance profiles of pJJ5 142

**Fig 5.1** Construction of pJJ7 151

**Fig 5.2** Construction of pJJ8 and pJJ9 153

**Fig 5.3** Construction of pJJ18 155

**Fig 5.4** OTC resistance profiles of pJJ8 and pJJ9 156

**Fig 5.5** OTC resistance profiles of pJJ18 157

**Fig 5.6** PAGE of whole membrane fractions of BL21(DE3)/pJJ8 159

**Fig 5.7** PAGE of column eluates stained with Coomassie blue 160

**Fig 5.8** PAGE of column eluates stained with silver 161

**Fig 5.9** Western blot of column eluates probed with Anti-His IgG using fluoresecence 164

**Fig 5.10** Positioning of PCR primers for OtrB deletion mutation 169

**Fig 5.11** Sequence of mutated central cytoplasmic loop of OtrB 170

**Fig 5.12** Construction of pJJ12 and pJJ13 173

**Fig 5.13** Construction of pJJ14 and pJJ17 174

**Fig 5.14** Screening for pJJ17 by restriction digest 175

**Fig 5.15** Sequence of polypeptide tag 181

## List of Tables

**Table 2.1** Bacterial strains 60

**Table 2.2** Plasmids and Vectors 61

**Table 3.1** References and accession numbers of proteins used in PILEUP analyses of OtrB 122

**Table 3.2** Results of TblastN search for *otrB* homologues 124

**Table 4.1** Results of OTC resistance tests for *otrB* in two *Streptomyces* sp. 135

**Table 4.2** Results of *otrB* specificity tests in BL21(DE3) pLysS 143

**Table 5.1** OTC resistance of pJJ14/pJJ17 162

**Table 5.2** Codon preference for 67 Streptomycete genes 171

**Table 5.3** Results of pJJ14/17 OTC resistance tests 172

## Abbreviations

| | |
|---|---|
| ATP | Adenosine 5' triphosphate |
| bp | base pairs |
| CIAP | calf intestinal alkaline phosphatase |
| cfu | colony forming units |
| C-(terminal) | carboxy-(terminal) |
| CTC | chlortetracycline |
| Da | dalton |
| DDM | dodecylmaltoside |
| DMSO | dimethylsulphoxide |
| DTT | dithiothreitol |
| dNTP | deoxynucleotide 5' triphosphate |
| dATP | deoxyadenosine 5' triphosphate |
| dCTP | deoxycytidine 5' triphosphate |
| dGTP | deoxyguanosine 5' triphosphate |
| dTTP | deoxythymidine 5' triphosphate |
| DNA | deoxyribonucleic acid |
| $dH_2O$ | distilled water |
| E | RNA polymerase core enzyme |
| EDTA | ethylene diaminetetra-acetic-acid (disodium salt) |
| EtBr | ethidium bromide |
| EtOH | ethanol |
| g | gram (or centrifugal force equivalent to gravitational acceleration |
| IPTG | isopropylthio-β-D-galactoside |
| hrs | hours |
| kb | kilobases |
| l | litres |
| M | molar |
| MFS | major facillitator superfamily |
| ml | millilitres |

| | |
|---|---|
| mins | minutes |
| N-(terminal) | amino-(terminal) |
| OD | optical density |
| ORF | open reading frame |
| OTC | oxytetracycline |
| PEG | polyethylene glycol |
| PCR | polymerase chain reaction |
| PMF | proton motive force |
| RNA | ribonucleic acid |
| rpm | revolutions per minute |
| rRNA | ribosomal ribonucleic acid |
| SD | standard deviation |
| SDS | sodium dodecylsulphate |
| TET | tetracycline |
| TEMED | NNN-N'-tetramethylethylenediame |
| TMH | transmembrane helix |
| Tris | tris(hydroxymethyl) methylamine |
| tRNA | transfer ribonucleic acid |
| U | units |
| UV | ultraviolet |
| V | volts |
| v/v | volume for volume |
| w/v | weight for volume |
| x-gal | 5-bromo-4-chloro-3indolyl-β-galactoside |

# CHAPTER 1

## Introduction

# Chapter 1

# Introduction

## 1.1 General Introduction to the Streptomycetes

*Streptomyces* are a group of Gram-positive, aerobic, filamentous soil bacteria. Unlike many bacteria they exhibit a complex life cycle. The *Streptomyces* exist in the soil as part of a large community of organisms including fungi and other bacterial species which recycle complex nutrients such as lignin and celluloses from dead plant matter. These soil organisms usually exist in a state of nutrient limitation.

This type of competitive existence has lead to various organisms evolving adaptations to enhance their competitiveness over the others which share the same nutrient sources; the complex morphological differentiation of the *Streptomyces* is one such example, (Fig 1.1). The filamentous growth of these organisms bestows upon them a competitive advantage by enabling the bacteria to come into contact with fresh sources of nutrients as they grow. When all reachable nutrient sources become depleted, sporulation is triggered, resulting in dispersion to new locations where a more nutritionally-rich environment will hopefully be found. In this environment, spores germinate and new hyphae can start to develop once again.

## 1.2 Secondary metabolites

### 1.2.1 Function & evolution of secondary metabolites

*Streptomyces* spp., like many Actinomycetes produce a large number of natural products which are inessential for growth. These are, by this definition, secondary metabolites. They are usually produced after vegetative growth, in conjunction with differentiation. In 1986 it was estimated that over 60% of the known bioactive compounds at that time were produced by *Streptomyces* species (Omura, 1986) although, with the high throughput screening of products from other natural sources taking place today, this figure may have decreased slightly in the 1990s. There is one school of thought which holds that the ability of the *Streptomyces* to produce so many bioactive secondary metabolites confers a selective advantage upon them. This argument around selective pressure has three main points as follows:

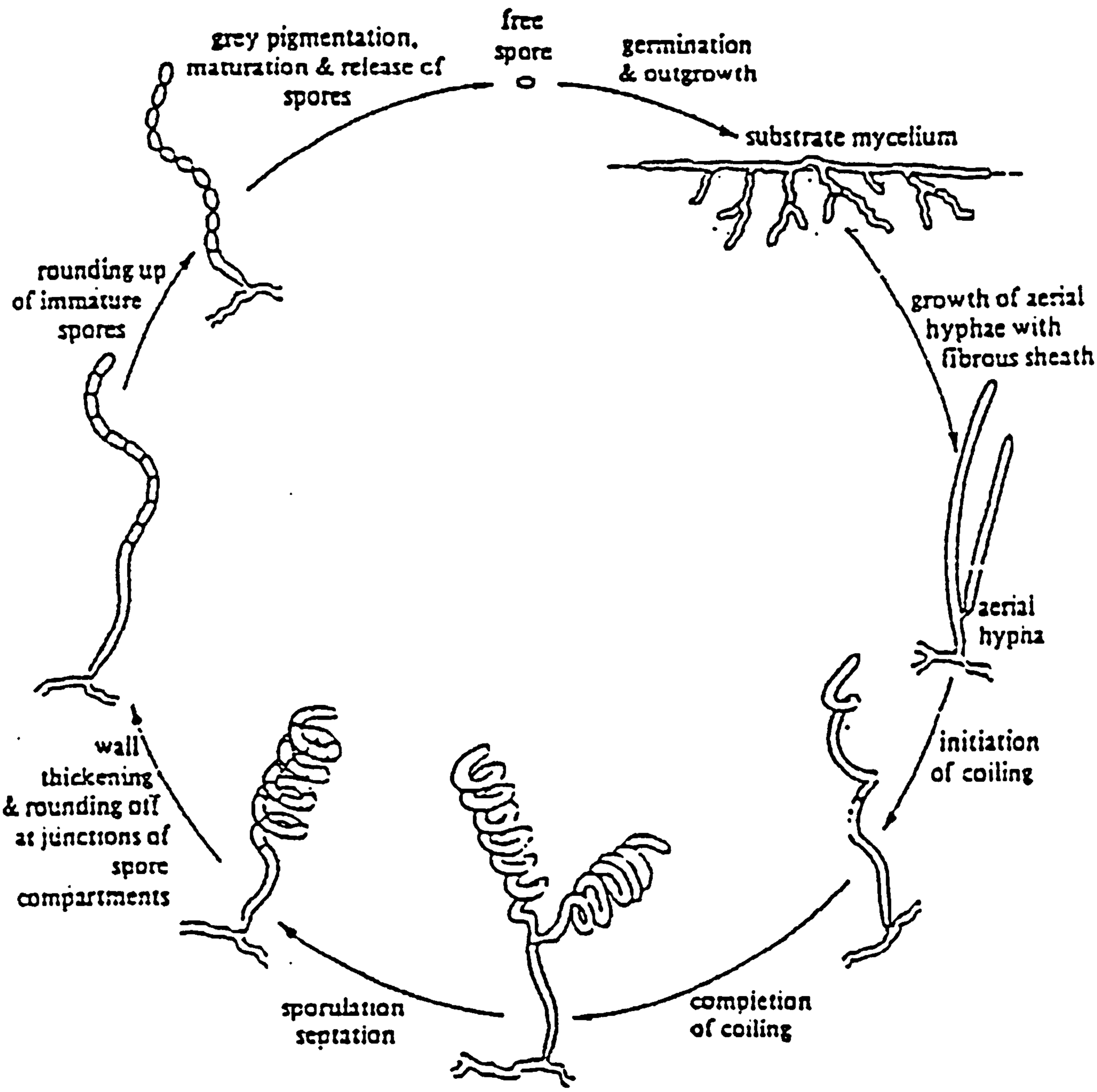



Fig 1.1

Life cycle of *S. coelicolor*, adapted from Chater & Merrick, 1979.

1) Secondary metabolism may have evolved as a method of self defence for organisms which can not elicit an immune response ( Stone & Williams, 1992).
2) If the metabolites produced by the secondary metabolism conferred no selective advantage upon the host, their biosynthesis pathways would not be so complex and energetically costly to the host (Katz & Demain, 1977).
3) As so many secondary metabolites are bioactive and inhibit the growth of other microbes, they should be involved in competition between microbes, animals and plants in their natural soil habitat (Demain, 1980).

Another school of thought is the argument that, despite bestowing selectively advantageous properties on the host cell, secondary metabolism did not arise primarily for this purpose. Instead, its purpose was to act as a "shunt" to divert intermediates of primary metabolism and prevent their build up to toxic levels during the switch from growth to stationary phase (Bu'Lock, 1980).

There is also a third theory on the evolution of secondary metabolites, which is that the structure evolved chemically before the complex pathways for their biosynthesis had developed. The low weight chemical compounds are thought by some to have played important roles early on in the evolution of biochemical reactions, effecting and moderating condensation reactions by chemical and structural interactions with receptor sites on simple early macromolecules i.e., secondary metabolites are suggested to have evolved as pre-polypeptide catalysts which were replaced as the reactions became ever more complex and proteins became involved (Davies, 1990). These low molecular weight "effectors" now interact with the original receptor sites in antagonistic fashion, thus explaining their antibiotic activity (Davies *et al*, 1992). There is some evidence for this theory, condensation of amino acids to form oligo and polypeptides is now undertaken by translation on the ribosome. However, the theory states that originally, such condensations were catalysed by secondary metabolites. It is certainly the case that most well known inhibitors of protein synthesis are extremely specific in their binding to ribosomal target sites, (Cundliffe, 1987) and the complexes formed by these interactions are at least as stable as many enzyme-substrate interactions.

Such an explanation holds well for antibiotics which inhibit protein synthesis However, modification of the antibiotic or it's transport into and out of the cell must

have evolved at a later time when the biosynthetic pathways that result in secondary metabolite formation evolved. For this theory to be true there must have, at one time, been a switch in the method of antibiotic synthesis from a non-enzymatic reaction to a method of biosynthesis similar to that which we see today involving several proteins. When and why then, did this change occur? Obviously after the antibiotics had given their role as ribosomal regulators over to ribosomal proteins but there must have been some impetus to develop a whole new biosynthetic pathway (Stone & Williams, 1992).

The view subscribed to by the majority of Streptomycete biologists seems to be the first. It is believed that secondary metabolism exists to protect the host during those points of the life cycle when its metabolic nutrients could be scavenged by other microbes in the vicinity. The theory is that the connection between morphological differentiation and the onset of secondary metabolite production is not one of chance- indeed the genetic controls for the two processes show important correlations (Hopwood, 1986).

### 1.2.2 Control of secondary metabolism

The discovery that N-acyl-homoserine-lactones (HSLs) act as a cell to cell signals amongst some Gram-negative bacteria in the phenomenon known as "Quorum sensing" (for a review see Swift *et al*, 1996) means that we should look again at the functions of antibiotics and the secondary metabolism pathways. Quorum sensing has been defined as, "a cell to cell communication device to regulate transcription of multiple target genes in concert with cell density" (Milton *et al*, 1997). In Gram-positive bacteria all of the "pheremones" so far identified have been either gamma-butyrolactones or post-translationally modified peptides (Ji *et al* 1995; Wirth *et al*, 1996). HSLs and their counterparts are now known to regulate diverse physiological processes in a variety of bacteria, including secondary metabolite production in the plant pathogen *Erwinia carotovora* (Bainton *et al*, 1992, Wirth *et al*, 1996).

In Streptomycetes, A-factor which is a structure containing gamma-butyrolactone has been shown to act as a microbial hormone in a similar way to the HSL's. A-factor was originally discovered by Khokhlov (Khokhlov *et al* 1973, 1967,) as a compound that could effect the sporulation of a *bld* mutant of *S. griseus* and will be further

discussed below. The regulation of secondary metabolism and morphological differentiation in the Streptomycetes is complex and occurs at multiple levels, as one might expect, in order to effect the many changes which occur as the organisms enter stationary phase.

Regulation occurs at the levels of transcription, translation and at the post-translational level.

### 1.2.2.1 Transcriptional regulation:

### 1.2.2.2 A-factor

The A-factor mentioned briefly above is a low molecular weight metabolite which acts as a microbial hormone in *S. griseus*. It is also found in other Streptomycetes but its function in these organisms is unclear. A-factor was first found in *S. griseus* and *S. bikiniensis* by Khokhlov *et al* (1973). It is essential for streptomycin biosynthesis, streptommycin resistance and spore formation in these organisms (Hara & Bepu, 1982). A-factor negative mutants of *S. griseus* were obtained by curing treatments which suggests that an extra chromosomal element is involved in A-factor synthesis (Hara & Beppu, 1982). It is now known that the A-factor determinant, *AfsA* is located on a giant circular plasmid, (1500kb) in *S. griseus* (Horinouchi & Beppu, 1994). In *S. coelicolor* this is not the case, A-factor is stable to curing treatments and has been shown to be situated in the chromosome (Hara *et al*, 1983).

### 1.2.2.3 A-factor in *S. griseus*

In *S. griseus*, A-factor performs it's regulatory function by binding to a specific receptor protein which, in the absence of A-factor, acts as a transcriptional repressor for genes involved in morphological and physiological differentiation. In the presence of A-factor, the receptor protein transfers the signal to the upstream activation sequence of the *strR* gene.

### 1.2.2.4 A-factor receptor protein

the A-factor receptor protein or (A-factor binding protein) of *S. griseus* negatively controls both streptomycin production and sporulation (Miyake *et al*, 1990). A mutant was isolated which did not require A-factor to produce streptomycin under normal

conditions It was concluded that this strain, *S. griseus* 2247 is deficient in the A-factor binding protein (Hara & Beppu, 1982). As strain 2247 was not deficient for A-factor but could produce streptomycin independently of it, a hypothesis was formulated. This states that the A-factor binding protein in the wild type *S. griseus* acts as a repressor-regulator for the streptomycin biosynthetic cluster and that that the repression is released when it binds with A-factor. In order to prove this, A-factor negative mutants of *S. griseus* isolated by Horinouchi *et al*, (1984) were further mutated. By looking for the pseudoreversion of streptomycin production, strains which were also deficient in the A-factor binding protein should be identified. Three such mutants were isolated. On examination the mutants retained their A-factor negative phenotype as expected, they were able to sporulate normally, unlike the parent strain, and displayed streptomycin resistance which was not observed for the parent strain. During the course of these investigations it was observed that the mutants produced streptomycin at an earlier stage of growth than the wild type *S. griseus*, after two days rather than after three, the cell masses of the two strains were almost the same. The amount of drug produced by the mutants was about 10 times higher than the wild type yields. It was suggested that the early production could be due to the triggering of streptomycin synthesis and sporulation being linked to the intracellular concentration of A-factor. As the binding factor is absent, the concentration of free A-factor will build up to the critical level earlier on in the life cycle.

**1.2.2.5 A-factor dependent binding protein (AdpP).**

A 41kDa protein has been shown to bind to the upstream activation sequence of the *strR* regulatory gene (Vujaklija *et al,* 1993). During the course of these experiments Vujaklija and co-workers found that three additional proteins also bound specifically to the region near the upstream activating sequence of *strR*, the precise functions of these proteins are not known but it is suggested that they may play regulatory roles in modulating A-factor/UAS binding. The A-factor dependent binding protein (AdpP), concerns only Streptomycin production, not aerial mycelium formation because transformants of wild type *S. griseus* containing the *strR* promoter on a high copy number plasmid producing little or no Streptomycin (Sm) but still form abundant spores. This is in accordance with the hypothesis that AdpP acts as an activator type

regulator, the lack of Sm biosynthesis being due to titration effect. (Horinouchi & Beppu, 1994).

### 1.2.2.6 strR

*strR* encodes an activator protein for genes in the streptomycin biosynthesis cluster. There is more than a 10 fold increase in *strR* transcription in the presence of A-factor. Work done by Distler *et al*, (1987) identified three mRNAs that were only detectable in the presence of A-factor; *strR* a regulatory gene, *aphD* encoding Sm-6-phosphtransferase and *strB* encoding aminocyclitol amidinotransferase. The group of Horinouchi and Beppu then conducted experiments to determine which of these three promoters is directly affected by A-factor. This work showed that only the *strR* promoter responded directly to A-factor and that the transcript encoding *aphD* is produced mainly by read through from the *strR* promoter (Vujaklija *et al*, 1991). As *aphD* encodes the streptomycin resistance determinant ,this makes sense.

Cotranscription ensures that resistance is induced just before streptomycin biosynthesis begins. The StrR protein is then thought to act as a transcription factor for other genes in the Sm biosynthetic cluster. StrR has been shown to be a DNA binding protein that functions by a mechanism that is not yet clear (Distler *et al*, 1992).

### 1.2.2.7 A model for the regulation of Sm biosynthesis

Taking into account the available data, a model for the way in which A-factor regulates streptomycin biosynthesis in *S. griseus* has been presented in a review by Horinouchi & Beppu (1994).

At an early stage in growth, intracellular concentration of A-factor is low. At this time the A-factor receptor protein represses the transcription of an (as yet) unidentified gene, gene X. The product of this gene should be involved in Sm biosynthesis and aerial mycelium formation since both phenotypes are triggered by the joint action of A-factor and A-factor receptor protein. When A-factor reaches a critical level this acts as a signal for Sm biosynthesis by binding to the A-factor receptor protein which is then released from gene X , allowing expression of the gene X product. This protein product then induces expression of AdpA, the transcription factor for StrR. Once StrR

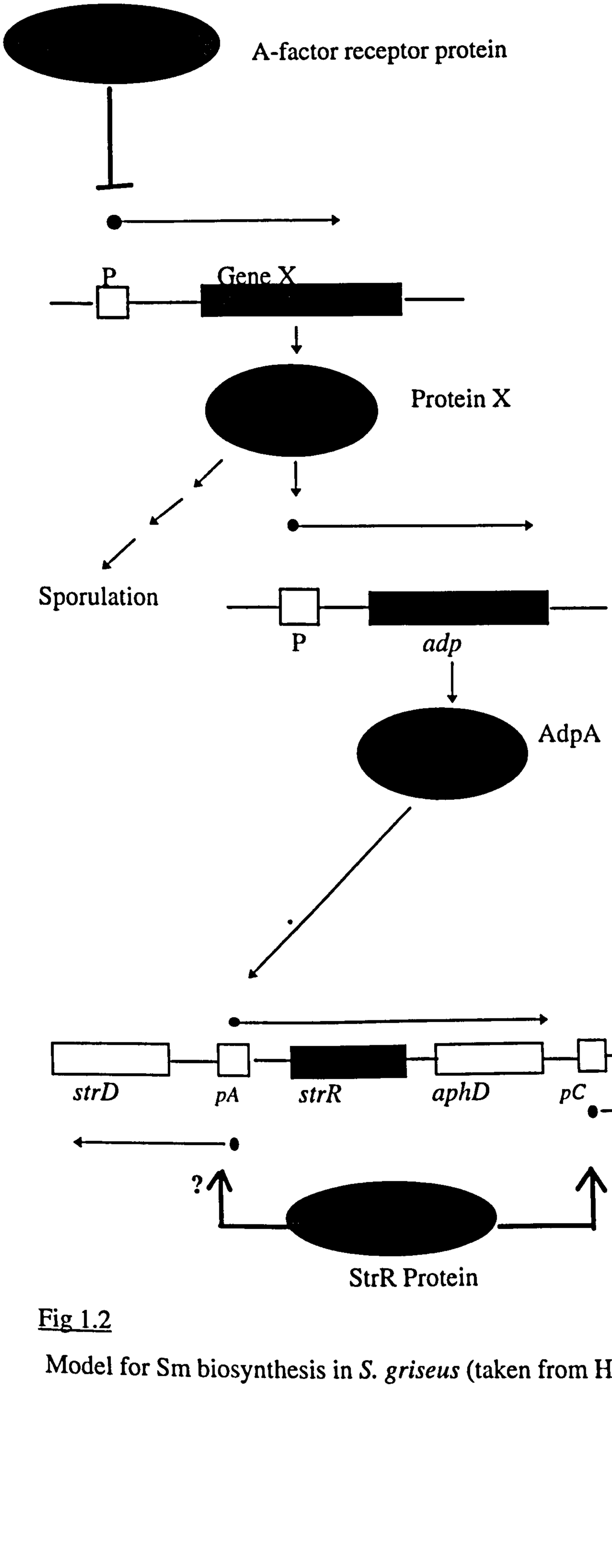



Fig 1.2

Model for Sm biosynthesis in *S. griseus* (taken from Horinouchi & Beppu, 1994)

is produced it can act as a transcription factor for other Sm genes including *strB* and *strD* (A flow diagram representing these processes is presented in Fig. 1.2).

### 1.2.2.8 A-factor in *S. coelicolor*

A-factor in plays no known regulatory or hormonal function in *S. coelicolor*. It is one of five secondary metabolites produced by the organism. Secondary metabolite production is controlled and regulated by a large inter-connected network of genes which is now beginning to be understood.

### 1.2.2.9 AbsA and AbsB

Mutants of *S. coelicolor* which are totally blocked in synthesis of all four antibiotics, *absA* and *absB* mutants, were isolated (Adamidis *et al* 1990). The global control of antibiotic synthesis in this organism is regulated by these two loci. By studying the genes and gene products which suppress these mutations, information about the hierarchical nature of the regulatory network can be gleaned.

### 1.2.2.10 AfsR (AfsB)

A DNA fragment from *S. coelicolor* which controlled streptomycin biosynthesis was isolated . The gene *afsB* was identified because it complemented an *afsB* mutation on the chromosome of *S. lividans.* (Horinouchi & Beppu, 1983). However, this complementation was later shown not to be a direct consequence of the cloned gene. Expression of the gene plèiotropically and positively controls, the production of three metabolites, A-factor, actinorhodin and undecylprodigiosin. These three metabolites are unrelated with respect to structure. The AfsR protein is important in it's own right, disruption of protein function resulting in a loss of pigment production. When *afsR* is cloned on a plasmid into *S. lividans*, which contains the *act* genes but does not produce actinorhodin under normal physiological conditions, actinorhodin synthesis is induced (Horinouchi *et al* 1984). This further strengthens the case for its function as a regulator. *afsR* was cloned and sequenced (Horinouchi *et al*, 1990).The deduced AfsR protein is 993 amino acids and contains A-and B-type ATP binding domains in the N-terminal portion and two DNA binding concensus sequences in the C-terminal part. When either of the two ATP binding regions was mutated, loss of pigment formation

in *S. lividans* was observed. A similar phenotype was observed on disruption of the chromosomal *afsR* gene in *S. coelicolor* indicating that AfsR is required for antibiotic synthesis. The presence of the ATP binding motifs led researchers to assay the protein for autophosphorylation. Phosphorylation did not occur in a system using radiolabeled ATP and $MgCl_2$ but was observed when AfsR was incubated with the *S. coelicolor* cell lysate. Labelling was proportional to the amount of AfsR protein (Hong *et al*, 1991). A similar AfsR-phosphorylating activity was observed in the lysate of *S. lividans*. It was suggested by Hong *et al*, (1991) that AfsR may be part of a two component regulatory system, widespread in many prokaryotes.

### 1.2.2.11 AfsK

In fact this is not the case, AfsR was found to be phosphorylated by a eukaryotic type protein kinase, AfsK (Matsumoto *et al*, 1994).

Kinase activity was located in the membrane fraction of *S. coelicolor* when a plasmid containing *afsR* and its downstream region was present. This suggested that *afsK* was situated downstream of *afsR* and that AfsK was membrane associated. The peptide sequence of AfsK shows remarkable similarity to eukaryotic serine/threonine protein kinases. Although AfsK phosphorylates AfsR, it is probably not the only protein to do so. Disruption of the chromosomal *afsK* gene results in reduction but not cessation of actinorhodin production and phosphorylated AfsR can still be detected in the disruptant (Matumoto *et al*, 1994).

### 1.2.2.12 AfsR2

Another gene at the *afsR* locus has been documented for *S. lividans* (Vogtli *et al*, 1994). *afsR2* is situated just downstream of *afsR* and also affects the production of actinorhodin, specifically by activating transcription of the *actII*-ORF4 regulatory gene of the actinorhodin biosynthetic cluster. AfsR2 also stimulates undecylprodigiosin production. Even though overproduction of AfsR2 increases *actII*-ORF4 mRNA concentration, it is not required for expression of this gene since actinorhodin is produced in *afsR* mutants of *S. lividans*. It is suggested by Vogtli *et al*, (1994) that some of the stimulatory effects on production of actinorhodin previously attributed to *afsR* are in fact caused by *afsR2* as the DNA fragment used by Horinouchi *et al* in the

*afsR* studies contained some downstream DNA now known to encode AfsR2. The mode of action of AfsR2 is not clear. There is similarity between a segment of AfsR2 and a domain found in most σ factors but AfsR2 is unlikely to function as a σ factor by itself as the most highly conserved region amongst other σ factors is not present and the AfsR2 protein is considerably smaller than most σ factors.

### 1.2.2.13 AfsQ1-Q2

AfsQ1 is another phosphorylated protein that plays a role in regulating antibiotic synthesis in *S. coelicolor*. The gene has been cloned and characterised. (Ishizuka *et al* 1992) as a response regulator typical of the prokaryotic two component regulatory systems previously mentioned. A second gene *afsQ2* downstream of *afsQ1* encodes a sensory histidine kinase which is thought to detect environmental signals and transduce such signals to AfsQ1 by phosphorylation. The two genes are transcribed polycistronically. A fragment containing these two genes is capable of conferring actinorhodin production on *S. lividans* 66HH21, an A-factor deficient mutant. When *afsQ1* is truncated, no actinorhodin biosynthesis is observed. In a strain of *S. coelicolor* in which the *afsQ1* and 2 genes were disrupted, growth, sporulation and production of actinorhodin, undecylprodigiosin and A-factor were comparable with the wild type, indicating that *afsQ1* and Q2 are not essential for these phenotypes in *S. coelicolor.*

Like the *afsR* gene, *afsQ1* was found to complement *absA* mutants. These are mutants that result in a global block of antibiotic synthesis in *S. coelicolor* (Adamidis & Champness, 1992). This supports the hypothesis that the afsQ1/Q2 system is involved in regulation of secondary metabolite production and is part of a network of regulatory genes which work in concert with some genes and independently of others in the network. It is thought that other sensor/regulator systems may be present in *S. coelicolor* (Ishizuka *et al*, 1992).

### 1.2.2.14 AbaA

There is another pleiotropic regulatory locus present in *S. coelicolor*, which like the *abs* and *afs* genes acts independently of sporulation. *abaA* was identified as a segment of DNA from *S. coelicolor* that increases actinorhodin synthesis when cloned in *S. lividans* (Fernandez-Moreno *et al*, 1992a). When the chromosomal *abaA* gene is disrupted by prophage insertion, production of actinorhodin is abolished and production of CDA and prodigiosin is greatly reduced while methylenomycin production is not significantly affected. These phenotypes are very similar to those of *afs* mutants but the two loci have different positions on the genetic map (Fernandez-Moreno *et al*, 1992b). The *abaA* DNA sequence contains at least two ORFs; ORF A and ORF B. ORF B is implicated in control of antibiotic production. No insertion mutants of ORFA were obtained so a similar role for it cannot be ruled out.

### 1.2.2.15 Sigma factors

### 1.2.2.16 $\sigma^{whiG}$

Multiple forms of RNA polymerase exist in *Streptomyces coelicolor*. At least nine sigma factors have been identified and six of these play a role in regulating differentiation and secondary metabolism.

Most of the work concerning *Streptomyces* sigma factors involved in differentiation has been carried out using *S. coelicolor* and *S. griseus*. The first gene identified as encoding a sigma factor was the *whiG* gene of *S. coelicolor* (Chater *et al*, 1989). Mutations at the *whi* loci are so called because they block the formation of spores in the aerial mycelium which therefore stay white even with prolonged incubation (Hopwood *et al*, 1970). The *whiG* mutants are blocked at an early stage in this process. They display very long straight hyphae that lack septation, which is characteristic of sporulation in the wild type.

When *whiG* is expressed at high copy number in *S. coelicolor*, increased sporulation is observed (Mendez & Chater, 1987). This suggests that the *whiG* product positively regulates sporulation. Studies of the nucleotide sequence of the gene indicate that *whiG* encodes a sigma factor, closely related to $\sigma^{28}$ from *Bacillus subtilis* which regulates motility, chemotaxis and autolysis (Helmann *et al*, 1988). The highest similarity was found in the regions which encode the amino acids that make contact

with the DNA. As $\sigma^{28}$ mutants of *B. subtilis* also form long filaments, it is possible that both factors are involved in transcription of autolysin genes.

Further experiments showed that $\sigma^{whiG}$ is essential for sporulation in *S. coelicolor* but is not required for vegetative growth (Chater *et al*, 1989) and that a *whiG* homologue is present in most streptomycetes. The *S. aureofaciens* homologue, *rpoZ*, has been cloned (Kormanec *et al*, 1994) Its gene product shows extensive similarity to $\sigma^{whiG}$, and $\sigma^{28}$. Similarities have also been observed in the regions upstream of these genes, implying that the two genes may be regulated in similar ways.

### 1.2.2.17 *hrd* σ factors

Four further sigma factor genes were cloned independently by two groups, the Buttner group at the John Innes Centre in Norwich, UK ( Buttner *et al*, 1990) and the group of Tanaka *et al* in Tokyo (Tanaka *et al*, 1991). The *hrd* genes, *hrdA hrdB, hrdC* and *hrdD,* were cloned using an oligonucleotide probe based on conserved residues from *E. coli* and *B. subtilis* sigma factors (Tanaka *et al*, 1988). The same genes were cloned by the Buttner group using a different probe, the *rpoD* gene from *Myxococcus xanthus.* As *hrdB* deletion mutants could not be obtained, (Buttner *et al*, 1990) it was deduced that HrdB must be the principal vegetative sigma factor in *S. coelicolor*. This was backed up by sequencing data which showed HrdB to have high homology with the *rpoD* gene products of *B. subtilis* and *E. coli* which are responsible for regulation of housekeeping genes in those organisms. Mutants of *hrdC* and *hrdD* were not visibly different in phenotype to the wild type *S. coelicolor* (Buttner *et al*, 1990) and hence these genes were deduced to be non-essential.

Transcriptional analysis of *hrdB, hdrC* and *hrdD* genes was carried out to provide more information about their expression. For *hrdD,* two transcripts were identified: one with 5' end 365bp upstream of the putative start codon and the other beginning 117 bp upstream, indicating that there are tandem promoters for *hrdD* in *S. coelicolor*. No *hrdC* transcripts were seen. It has been suggested following the construction of double and triple *hrd* mutants that *hrdC* and *hrdD* may have identical or overlapping promoter specificities (Buttner & Lewis, 1992). As the double and triple mutants of *hrdA, C* and *D* are not able to be distinguished phenotypically from the wild type, it could be that all the promoters acted upon by these proteins can also

be recognised by the *hrdB* gene product which is essential for growth and differentiation of *S. coelicolor*. It is now known that $\sigma^{hrdD}$ is responsible for transcription of the *reed* and *actII*-ORF4 genes of *S. coelicolor in vitro* (Fujii *et al*, 1996). These genes encode regulators required for biosynthesis of undecylprodigiosin and actinorhodin. There is a possibility that these promoters may be recognised *in vivo* by both $\sigma^{hrdD}$ and $\sigma^{hrdB}$. Homologues of the *hrd* genes have been found in other Streptomycetes including *S. lividans* and *S. griseus* (Buttner *et al*, 1990) and *S. aureofaciens* (Kormanec & Farkasovski, 1983) but only *hrdB* is present in all of the species so far. This further backs up the hypothesis that *hrdB* encodes the principal vegetative sigma factor in *Streptomyces* spp. The work of Kormanec and Farkasovski has shown that the existence of tandem promoters, like that observed for *hrdD* in *S. coelicolor*, is more widespread. Four *hrd* genes have been identified in *S. aureofaciens*, *hrd A, B, D* and *E*. It was discovered that transcription from *hrd* promoters depended on the developmental stage. *hrdA, B* and *D* each have at least two promoters with the downstream promoter being weakest in all three cases. *hrdB* is transcribed from both promoters at all developmental stages, both of the *hrdD* promoters are active only in the vegetative stage and *hrdA* is transcribed from its tandem promoters at the time corresponding to aerial mycelium formation. No transcripts were identified for *hrdE*.

Even though *hrdA* was transcribed only at the formation of aerial mycelium, there was no visible phenotypic difference between a *hrdA* mutant and the wild type *S aureofaciens*.

The internal promoter of the essential *hrdB* is of interest, tandem promoters were not identified in the *S. coelicolor hrdB* gene. Why should there be two gene products if this is a housekeeping sigma factor? It is suggested that the weaker internal *hrdB*-P1 promoter may be able to respond to specific physiological conditions (Kormanec & Farkasowski, 1993). This phenomenon of two polypeptides encoded by the same sequence is known in Streptomyces, the sporulation gene of *S. griseus* encodes two peptides that differ at their N-termini and are generated by temporal control of accumulation of the transcript (Babcock & Kendrick, 1990).

### 1.2.2.18 $\sigma^F$

More recently another sigma factor concerned with morphological differentiation has been identified (Potuckova *et al*, 1995). *sigF* was cloned from *S. aureofaciens* using an oligonucleotide probe derived from a conserved region of the $\sigma^{70}$ family. The *sigF* mutant did not correspond to any of the existing *whi* loci. There are homologues to *sigF* in other Streptomycetes. Sequencing of the *S. coelicolor sigF* showed that it encodes a peptide with high similarity to other sigma factors. Disruptants of *sigF* were constructed in *S. coelicolor*. These mutants were normal for vegetative growth but later displayed a *whi*-like phenotype. Microscopic analysis revealed that this phenotype was not the documented *whi* phenotype as some spores were produced but they were small and deformed compared to wild type. The walls of the mutant spores were thinner and sensitive to detergents, unlike the wild type. These observations lead to the hypothesis that $\sigma^F$ controls a regulon of late spore-specific genes. (Potuckova *et al*, 1995). There is now enough information to formulate a hypothesis stating that morphological differentiation in the streptomycetes is controlled by a sigma factor cascade. Work to elucidate the hierarchy of this cascade is now being carried out. $\sigma^{whiG}$ is epistatic to the *hrd* sigma factors and does not depend on $\sigma^F$ or any of the six known *whi* genes (*whi A, B, G, H, I and J*) required for sporulation septum formation (Kelemen *et al*, 1996). This puts it at the top of the regulatory cascade that controls spore chain development in *S. coelicolor*. Transcription of *sigF* depends on all of these early *whi* genes including *whiG* (Keleman *et al*, 1995. The dependence of $\sigma^{SigF}$ upon WhiH and WhiJ may not be absolute, as only very weak signals were detected in S1 nuclease protection analysis (Keleman *et al*, 1995). Transcription of *sigF* is developmentally regulated as transcripts are detected only at the time of sporulation-septum formation. No transcripts were detected in the vegetative phase, during aerial mycelium development or in later stages of spore maturation. Transcription of *whiG* carries on after the *sigF* transcripts cease to be present. *whiG* transcripts were present throughout all developmental phases as were those of *hrdB*, the housekeeping sigma factor, that was used as a positive control. As transcripts of *whiG* are present in the vegetative stage, before any differentiation has occurred, it is suggested that regulation of WhiG activity is at the post translational level. WhiG is related to another sigma factor known to be regulated this way, $\sigma^{FliA}$ from *S. typhimurium*. This is involved in

the late stages of flagellum biosynthesis. Another protein, FlgM, binds $\sigma^{FliA}$ until the basal body hook is complete. At this point, FlgM is exported through the hook, freeing the sigma factor. The *esp* locus of *S. coelicolor* whose mutants display ectopic sporulation is suspected of encoding a FlgM homologue in this organism (Keleman *et al*, 1995).

Although transcription of *sigF* depends on $\sigma^{whiG}$, the transcription of the *sigF* gene is not carried out directly by the E $\sigma^{whiG}$ polymerase holoenzyme. The promoter region of *sigF* shows no similarities to other genes transcribed by $\sigma^{whiG}$. The hierarchy elucidated so far has $\sigma^{whiG}$ at the top and $\sigma^{F}$ controlling the late events. The positions of the other *whi* genes are not yet determined but it is known that transcription of *whiH* is directly by E $\sigma^{whiG}$ (Whatling *et al*, in preparation (cited in Keleman *et al*, 1996)) and that *whiB* transcription does not depend on *whiG*.

Fig 1.3 shows the hierarchy of sporulation-controlling genes so far deduced in *S. coelicolor* (Keleman *et al*, 1996).

### 1.2.3 Post -transcriptional regulation

#### 1.2.3.1 Regulation by use of a rare tRNA

*bld* mutants of *S. coelicolor* were first characterised as being defective in aerial mycelium formation and sporulation (Merrick, 1976). *bldA* is one such mutant. This mutation also results in the total block of production of four antibiotics normally produced by the strain: actinorhodin, undecylprodigiosin, methylenomycin and the calcium dependent antibiotic (CDA). The phenotype of *bldA* mutants is carbon source dependent. When grown on minimal agar containing glucose or cellobiose as sole carbon source the mutant phenotype is seen but when grown with maltose or melibiose as the carbon source, *bldA* mutants produce aerial mycelia and sporulate as wild type. However the production of antibiotics is not restored. There are however a class of *bld* mutants known as the *pwb* mutants which can (by second site suppressor activity) reinstate the production of antibiotics from the *red* or *act* clusters but not both (White & Bibb, 1997, Guthrie *et al*, 1998). The *bldA* gene encodes a leucine tRNA for the codon, UUA which is rare in the G+C rich DNA of *Streptomyces* spp (Lawlor *et al*, 1987).

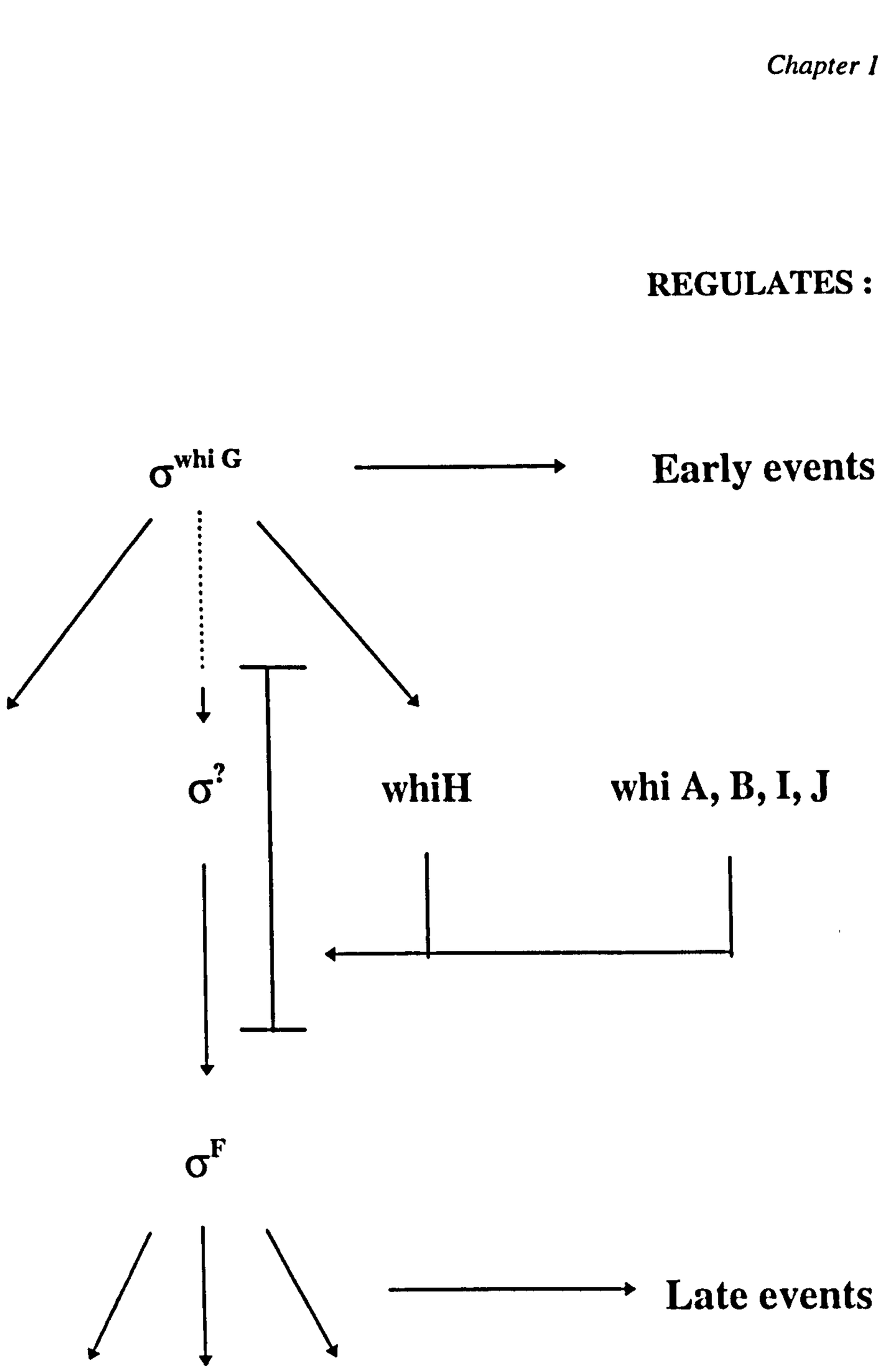



Fig 1.3

Heirarchy of genes controlling *S. coelicolor* sporulation. Solid arrows indicate direct biochemical interaction, broken arrows indicate genetic dependence which could be direct or indirect (taken from Keleman *et al*, 1996).

Expression of several genes that contain a TTA codon is lost or reduced in *bldA* mutants (Leskiw *et al*, 1991) including the transcriptional activator for actinorhodin biosynthesis (Fernandez- Moreno *et al*, 1991). This verifies that the *bldA* gene product is the principal route by which UUA codons are translated in *S. coelicolor*. In wild type *S. coelicolor*, the *bldA* encoded tRNA accumulates at the time of aerial mycelium development- It is just prior to production of the secondary metabolites.

Deletion of *bldA* had no obvious effects on vegetative growth but interfered with mycelium formation and sporulation (Leskiw *et al*, 1991). It was originally thought that accumulation of the *bldA* encoded tRNA was temporally regulated as it was present at low levels during the early stages of the life cycle and was present in increased abundance in older cultures. It is now thought that this is not the case and S1 experiments indicate that accumulation is not temporally regulated (Leskwi *et al*, 1991). Developmental regulation of translation of UUA codons may be at a number of different levels: processing of the *bldA* RNA transcript, its charging of the tRNA with the leucine residue, or possibly its access to the ribosome. It has also been shown that some genes containing UUA codons can be translated at an extremely low level in *bldA* mutants. Two genes coding for antibiotic modifying (resistance) enzymes (Leslaw *et al*, 1991) were able to be translated in a *bldA* mutant, *hyg* encoding hygromycin phosphotransferase and *aad* encoding aureomycin adenyl transferase. These genes contain one and two TTAs respectively (Hollingshead *et al*, 1985) and were able to be translated in *bldA* mutants (Leskwi *et al*, 1991) and confer resistance on the hosts. Expression may be due to insertion of another amino acid in place of leucine, if the codon/anticodon interaction is weak.

Data compiled from coding regions from 100 *Streptomyces* genes (Bibb *et al*, 1984) have shown that TTA codons are rare and mostly present in genes required for morphological and physiological differentiation. It may be that the presence of TTA codons confers some adaptive benefit but such benefits are not obvious. The lack of TTA codons in genes associated with vegetative growth may just be due to the shedding of inefficient codons through evolution.

Actinorhodin and undecylprodigiosin are not produced by *bldA* mutants, the *actII* -ORF4 gene which regulates actinorhodin production in *S. coelicolor* contains TTA codons but the *reed* gene, which plays a similar role in the *red* synthetic cluster

does not. It has therefore been suggested that a second pathway specific regulatory gene, that contains a TTA codon is also required for *red* activation (Guthrie & Chater, 1990).

### 1.2.3.2 Regulation by ppGpp and pppGpp.

The stringent response is coupled to amino acid depletion which causes the intracellular accumulation of RNA to cease, the accumulation of guanosine 5'-diphosphate 3' diphosphate (ppGpp) and guanosine 5'-triphoshpate 3' diphosphate, (pppGpp), is believed to be responsible for the onset of the stringent response (Gallant, 1979) and to be the initiating signal for secondary metabolism and differentiation in several *Streptomyces* strains (Ochi *et al*, 1986b, 1987a). It was observed by Ochi in *Streptomyces* MA406-A-1, that after the intracellular accumulation of ppGpp and pppGpp that follows nutritional shift down, initiation of antibiotic synthesis occurred. This work also showed that ppGpp inhibits the synthesis of GTP and so is indirectly involved in morphological differentiation. The stringent response and the accumulation of (p)ppGpp has been studied very extensively in *E. coli*, where the polyphosphorylated nucleotides are synthesised by at least two possible routes.

The main route is by (p)ppGpp synthetase activity. This enzyme is encoded by the *relA* gene (*relA* mutants do not exhibit the stringent response so are said to display a relaxed phenotype). Under conditions of amino acid deprivation, the synthetase acts on ribosomes when binding of uncharged tRNAs to the receptor site is occurring. This reaction involves transfer of pyrophosphate from ATP to GTP or GDP. The resulting transient accumulation of (p)ppGpp leads to complex regulatory reprogramming resulting in reduction in rate of transcription of stable mRNAs and increased transcription of certain amino acid operons. *relC* mutants have defective ribosomal (p)ppGpp synthesis. The *relC* gene encodes the L11 protein in the 50S ribosomal subunit, the same subunit implicated in the binding of the RelA protein (Parker *et al*, 1976). A *relC* mutant of *S. coelicolor* has been isolated (Ochi *et al*, 1990) it is deficient not only for the stringent response, but also for production of aerial mycelium actinorhodin and undecylprodigiosin.

The minor route by which (p)ppGpp is synthesised in *E. coli* utilises the product of the *spoT* gene, the ribosome independent (p)ppGpp synthetase II (Sarrubi *et al*, 1989). The

activity of this gene product was deduced due to the lack of (p)ppGpp accumulation in *relA spoT* double mutants (Xiao *et al*, 1991), compared with the accumulation of the molecule(s) in strains deleted for *relA* (Metzger *et al*, 1989).

### 1.2.3.3 (p)ppGpp and antibiotic synthesis

That (p)ppGpp may be the signal that links the onset of secondary metabolism to physiological differentiation is an attractive theory, and although the work of Ochi has provided information that links the accumulation of (p)ppGpp with initiation of secondary metabolism (Ochi, 1986a, 1986b), no such direct link was established by others (Strauch *et al*, 1991; Bascaran *et al*, 1991; Takano *et al*, 1992). Work by Martinez-Costa *et al* (1996) in the laboratory of Malpartida, has identified *relA/spoT* homologues in *S. coelicolor* and *S. lividans*. Further studies using the *S. coelicolor* gene showed that it was important in activation of secondary metabolism as it could initiate actinorhodin production when transformed into *S. lividans* by causing increased transcription of the *act*II ORF4 gene which acts pleiotropically. Deletion and insertional mutants of *S. coelicolor* were made to investigate further the importance of the *relA/spoT* homologue. The insertion mutants unexpectedly showed no apparent phenotypic differences from the parent strain. The deletion mutant on the other hand exhibited slowed growth, poor sporulation and abolition of actinorhodin production. The production of undecylprodigiosin and CDA was not affected by the mutation. The lack of actinorhodin production in the deletion mutant is due to transcription of *act*II *ORF4* being at a basal level. Production was restored when *actII* ORF4 DNA was present on a plasmid or when the plasmid containing the *relA* homologue was present. In tests for (p)ppGpp from ribosomes, the deletion mutants proved negative.

It is somewhat surprising that undecylprodigiosin production was unaffected by deletion of the *relA* homologue, as the regulator for this cluster (*redD*), shows marked similarity with the *act*II-ORF4 gene product. A regulatory function for *relA* in both antibiotic pathways cannot be ruled out as extra copies of the gene increases production of both antibiotics in the parent strain and in the *relA* deletion mutant (Martinez-Costa *et al*, 1996).

The Martinez-Costa *relA/spoT* deletion strain shows some likeness *to relA* mutants of *S. coelicolor* isolated by Ochi *et al*, (1990). The Ochi mutant was however affected in

its ability to produce undecylprodigiosin, provoking the thought that there may be alternative pathways for (p)ppGpp production in *S. coelicolor*.

The possible link between accumulation of (p)ppGpp and initiation of antibiotic synthesis is not yet definitely proved or disproved. It is suggested by Bascaran *et al*, (1991) that sensing of growth rate may be a critical factor in regulating the start point of secondary metabolite biosynthesis.

### 1.2.3.4 (p)ppGpp and morphological differentiation

As well as linking (p)ppGpp accumulation with antibiotic biosynthesis, the work of Ochi has implicated (p)ppGpp indirectly with the onset of aerial mycelium formation . In *B. subtilis,* sporulation is the result of a decrease in GTP caused by the stringent response and the concomitant production of antibiotics is a consequence of the stringent response (Ochi & Oshawa, 1984; Ochi, 1985). The same phenomena have been observed for *Streptomyces* MA406-A-1 and other *Streptomyces* spp (Ochi, 1986a). Studies on GTP synthesis utilise decoyinine, a specific inhibitor of GMP (and indirectly GTP) synthesis. Four *Streptomyces* strains studied by Ochi developed aerial mycelium on the addition of decoynine. These findings are in agreement with those for *B. subtilis* which underwent sporulation but did not produce antibiotics when treated with decoynine.

The stringent response is thought to affect GTP levels in the cell. *S. griseus* can be induced to sporulate by transferring cells from a rich to a nutritionally poor medium (Kendrick & Ensign, 1983). *Rel* mutants produce much fewer spores than wild type cells when treated in this way (Ochi, 1987b). Differences in the ppGpp levels between the two strains were apparent: *rel* mutants accumulated 10 times less of the molecule than the wild type which also showed large decreases in GTP, not seen in the *rel* strain. It was suggested that the GTP levels undergo this transient dramatic decrease because severe amino acid deficit leads to increased numbers of uncharged tRNA's and hence to ppGpp synthesis in wild type, but not *rel,* cells.

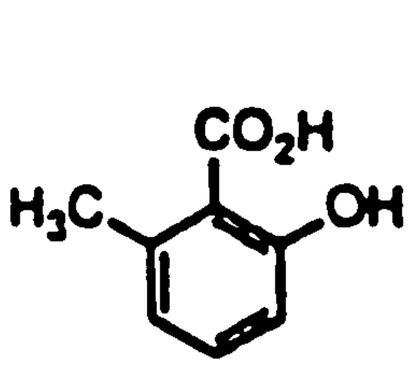
ppGpp in *E. coli* is a strong inhibitor of IMP hydrogenase, the first enzyme in the purine biosynthesis pathway. ppGpp also acts in this way in *S. griseus* (Ochi, 1987a). Reduction in levels of GTP could also result directly from nutritional shift down, nitrogen deficiency may curtail the purine biosynthesis pathway resulting in

lower levels of GMP and GTP. Studies on *Streptomyces clavuligerus* (Bascaran *et al*, 1991) have demonstrated the existence of a burst of ppGpp synthesis triggered by the stringent response in that organism.

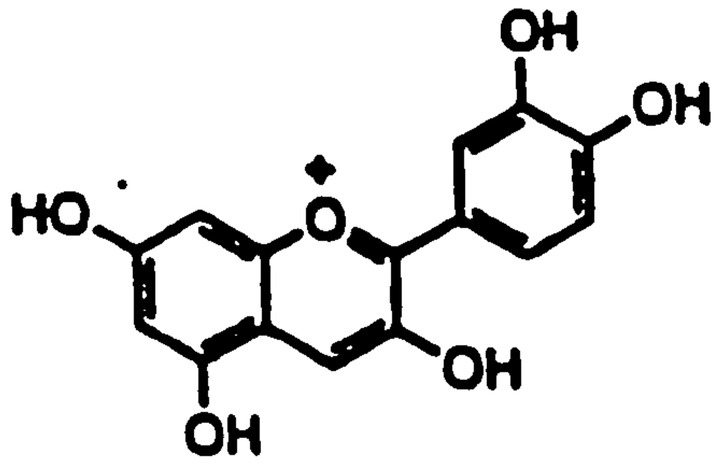
## 1.3 Polyketide Biosynthesis

### 1.3.1 What are Polyketides?

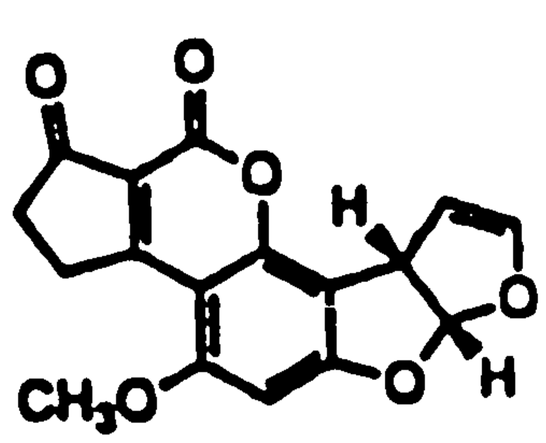
Polyketides are a large group of molecules, all of which are synthesised in nature as secondary metabolites from small carboxylic acids by a process which is very similar to that used by primary metabolism to synthesise fatty acids. Polyketide derived compounds are common in both prokaryotes and eukaryotes, being metabolites of bacteria, fungi and plants (Fig 1.4). Although the structures of these compounds vary in size and complexity, they all share one common factor:- their pattern of biosynthesis, which is discussed in section 1.3.2. Polyketide compounds play a large number of roles, including antibiotics (bacteria) flavenoids (plants) and mycotoxins (fungi). Their chemical diversity stems from the programming of the PKS and to the huge number of possible modifications that can occur after chain assembly, such as aromatic, ether or macrolide ring formation, and addition of different functional groups. Due to the fact that secondary metabolites are only produced under certain environmental and physiological conditions, knowledge relating to them is rather incomplete and anthropocentric. The most well studied examples are those which either help or hinder man, the most obvious example being the vast array of antibiotics, antitumor and antifungal compounds. Also well studied are the polyketide derived colours of flowering plants and the highly poisonous and carcinogenic aflatoxins produced by fungi.



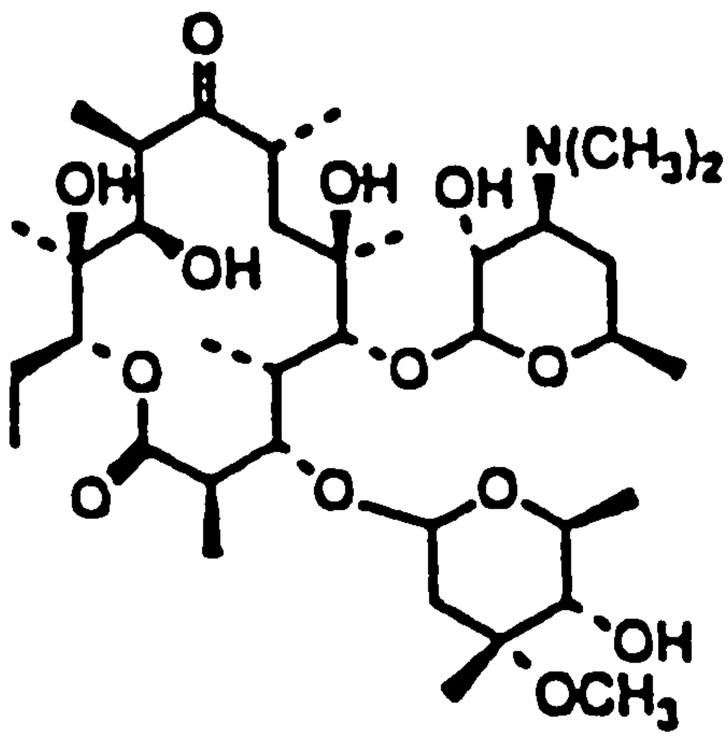

6-MSA



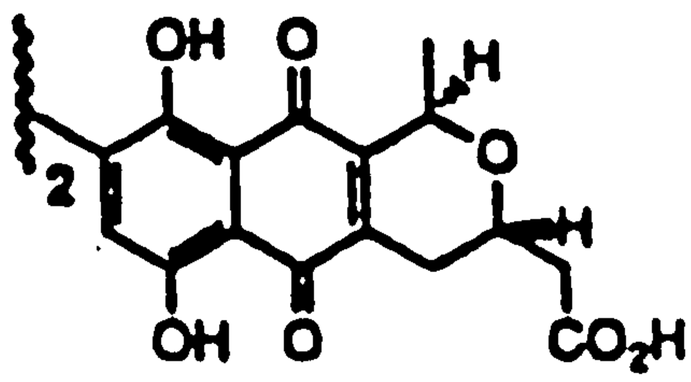

Cyanidin



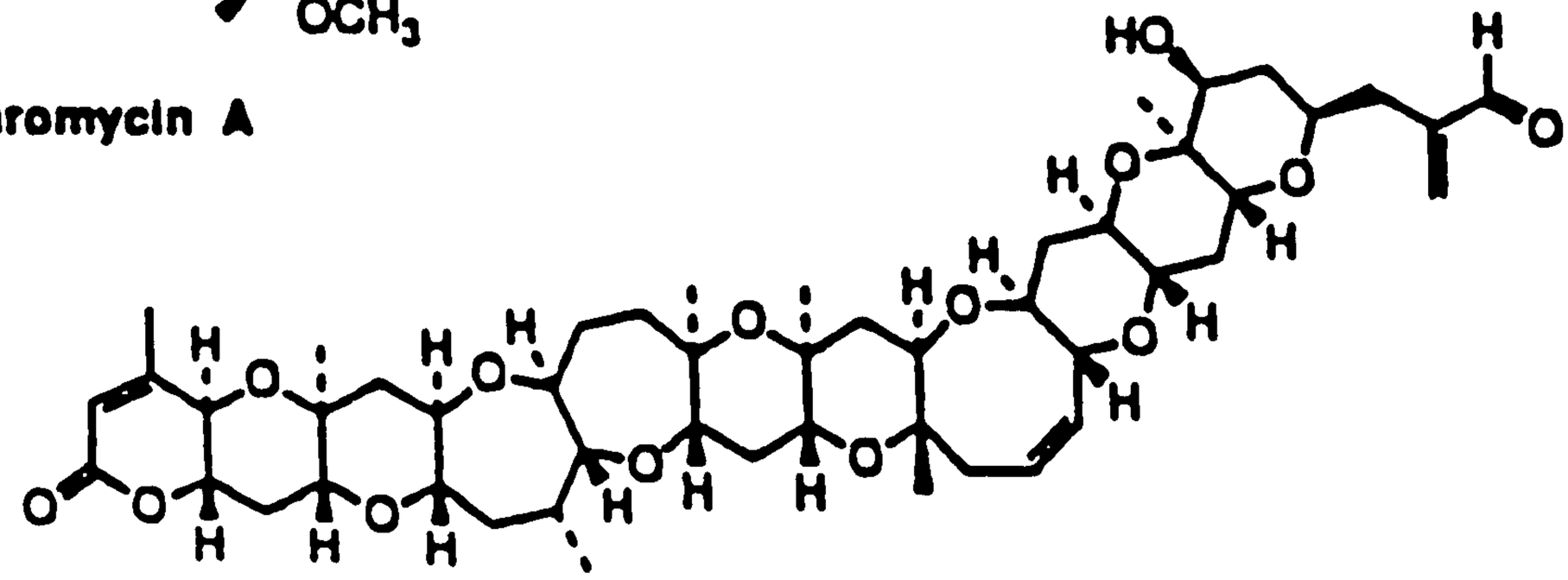

Aflatoxin B




Erythromycin A




Actinorhodin




Brevetoxin B

Fig 1.4

Structures of 6-methylsalicylic acid, a flavone (cyanidin); aflatoxin B; erythromycin A; actinorhodin and brevitoxin B. (Taken from Hopwood & Sherman, (1990).

### 1.3.2 Polyketide Biosynthesis

Polyketide biosynthesis occurs in a similar manner to that by which the fatty acids are synthesised in primary metabolism, in which 2-carbon extender units are added to a starter acyl group via a cycle of condensation, keto reduction, dehydration and eno-reduction. Fatty acid biosynthesis is catalysed by enzyme systems called fatty acid synthases (FAS).. There are 7 steps in fatty acid synthesis (Fig 1.5) Reactions 1 and 2 are priming reactions in which the condensation reaction precursors, malonyl Co-A and an acetyl group linked to acetyl carrier protein (ACP) by a thioester bond, are loaded onto the condensing enzyme, β ketoacyl-synthase. Reaction 3 is the condensation reaction whereby malonyl ACP is decarboxylated, the resulting carbanion then attacks the acetyl-thioester bond to form β ketoacyl-ACP. Reaction 4 is a reduction step catalysed by β ketoacyl-ACP reductase. Reaction 5 is a deydration step and reaction 6 is the second reduction step, this time an enoyl-reduction. Reactions 2-6 are then repeated six times to give a final product of 16 carbon chain length, palmitoyl-ACP. The ACP moiety is removed by palmitoyl thioesterase in the seventh and final step to form palmitic acid.

#### 1.3.2.1 Polyketide Synthase (PKS)

The PKS systems are categorised as three types, Type I PKS's are large multifunctional proteins in which all the enzymatic functions are carried out by various domains of a single polypeptide. The products of Type I PKS range from simple, 6-methylsalcylic acid produced by *Penicillium*, (Wang *et al*, 1990) to more complex, e.g. erythromycin from *Saccarospora erythraea*. (Cortes *et al*, 1990). In contrast to Type I, Type II PKS's consist of individual enzymes, each one a monofunctional unit. Several of the aromatic antibiotics are made by this type of system e.g. actinorhodin from *S. coelicolor*, tetracenomycin from *S. glaucesens* and importantly for this study, oxytetracycline from *S. rimosus* (Hopwood & Sherman, 1990). Both the Type I and II PKSs are similar to FAS in that they act in conjunction with acyl carrier protein (ACP) to catalyse addition of the building units.

Type III PKS's are key enzymes for the biosynthesis of plant metabolites. Interestingly these enzymes are unlike PKS Type I and II as they do not utilise (ACP)

but act directly on coenzyme-A (Hahlbrock, 1981). Chalcone synthase from parsley is an example of such a system.

In the biosynthesis of polyketides, the full length carbon chain is generated by addition of appropriate extender units onto a starting acyl unit as in fatty acid synthesis (Fig 1.5). The starter for polyketide biosynthesis does not have to be acetyl CoA. Oxytetracycline biosynthesis begins with malonamyl CoA (Thomas & Williams, 1983), daunorubicin starts with propionyl coA (Kitamura *et al*, 1981). A variety of extender units are also used in the biosynthesis of polyketides although most use malonyl CoA. In erythromycin biosynthesis however, methylmalonyl CoA is used as an alternative extender unit (Omura & Tanaka, 1984). When the chain reaches its final length, it is stabilised by cyclase reactions to form aromatic or macrolactone rings. During the biosynthesis of fatty acids, the FAS has only a limited number of "choices" to make. The final chain length must be decided and in a few versions, a choice of starter unit. Also, in bacteria where unsaturated fatty acids are common, unsaturation results from the loss of the eno-reduction step at one or two points in the chain. (In mammalian fatty acid synthesis, double bonds are introduced by specific desaturases once the chain has been completed). It is possible that components of the FAS and PKS systems could be shared between the two systems. The finding that the malonyl transferase gene for *S. coelicolor* lies 2.8 mega bases away from the *act* cluster and next to an ORF whose product resembles ketoacylsynthase II of *E. coli* gives weight to this hypothesis (Revill *et al*,1994). Malonyltransferase may indeed be shared between the FAS and PKS clusters of *S. coelicolor*. Polyketide synthesis can be thought of as an extended version of fatty acid synthesis involving a much larger range of "choices" by the PKS. Reflecting this, the PKS must be more "highly programmable" (Hopwood & Sherman, 1990). The unique feature of the PKS's is their ability to chose different starter units for biosynthesis. Fidelity of PKS programming is not absolute. *S. cinnamonensis* produces the polyketide metabolite monesin A as its major polyether metabolite. During the synthesis of this molecule, 37 choices are required to be made by the PKS (Hopwood & Sherman, 1990). Another molecule, monesin B is also produced by the organism albeit in low quantities. Monesin B production is thought to be due to incorporation of propionate instead of butyrate as the fifth extender unit (Gorman et al, 1968).

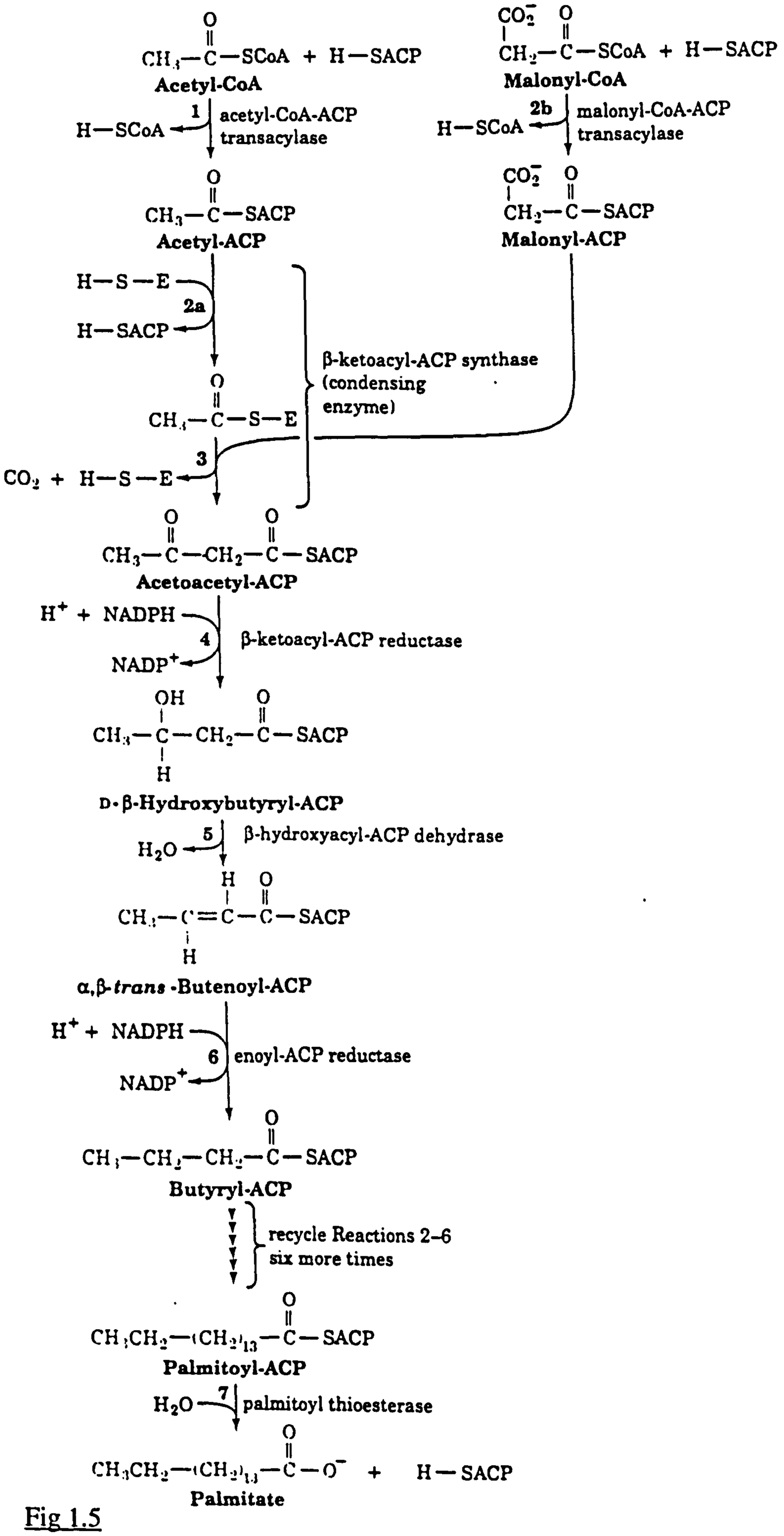



Fig 1.5

The seven step of fatty acid synthesis. Taken from Voet & Voet, (1993).

Actinorhodin produced by *S. coelicolor* is the best studied example of polyketide production by a type II PKS. All the genes for actinorhodin biosynthesis lie on a 22kb stretch of the *S. coelicolor* chromosome (Malpartida & Hopwood, 1984; Malpartida & Hopwood, 1986). Many *act* mutations had been mapped to this stretch of DNA which complemented *act* mutants of class I and III. These mutants were deduced to be blocked in backbone building as they did not secrete any actinorhodin intermediates but would produce actinorhodin from intermediates secreted by later blocked mutants (Rudd & Hopwood, 1979). DNA from this region was used to identify PKS genes from other antibiotic producing *Streptomyces* spp. including the *gra* cluster for granaticin production from *S. violaceoruber* (Sherman *et al* 1989), the *tcm* genes from the tetracenamycin producer *S. glaucesens* (Bibb *et al*, 1989) and of special interest to this work, the *otc* genes from *S. rimosus* which encode the oxytetracycline PKS. (Butler *et al*, 1989). This type of work indicates that genes encoding PKS enzymes occur in distinct clusters in the Streptomycetes. This is true for all polyketide antibiotic genes so far identified and also for other types of antibiotic synthesis enzymes. The regulatory and resistance genes are also present within these clusters.

Polyketide gene clusters share a common aspect of their organisation, a group of three open readings frames that show marked similarity (Fig 1.6). The first and second ORFs encode β ketoacylsynthase-like peptides, although their activity as such is not proven. It is thought that they are translationally coupled to produce a heterodimeric protein. The third ORF encodes the acyl carrier protein (ACP). This is important as it shows that formation of aromatic polyketides by PKS's is by a pathway using ACP and not similar to the plant PKS's which do not require an ACP. A fourth ORF is also present, often directly downstream of these three well conserved regions. ORF 4 encodes cyclase, the type of which varies between the different clusters. The FAS systems which the aromatic PKS's resemble also contain acyl transferase, required for transfer of the malonyl CoA to the ACP and a thioesterase for release of the completed chain. Neither of these components were identified in a study of 12 aromatic PKS systems by Katz & Donadio (1993). A motif typical of acyl transferase function has been identified in the C terminal portion of ORF 1 (Fernandez-Moreno *et al* 1992a) but whether any AT activity is associated with ORF 1 remains to be seen.

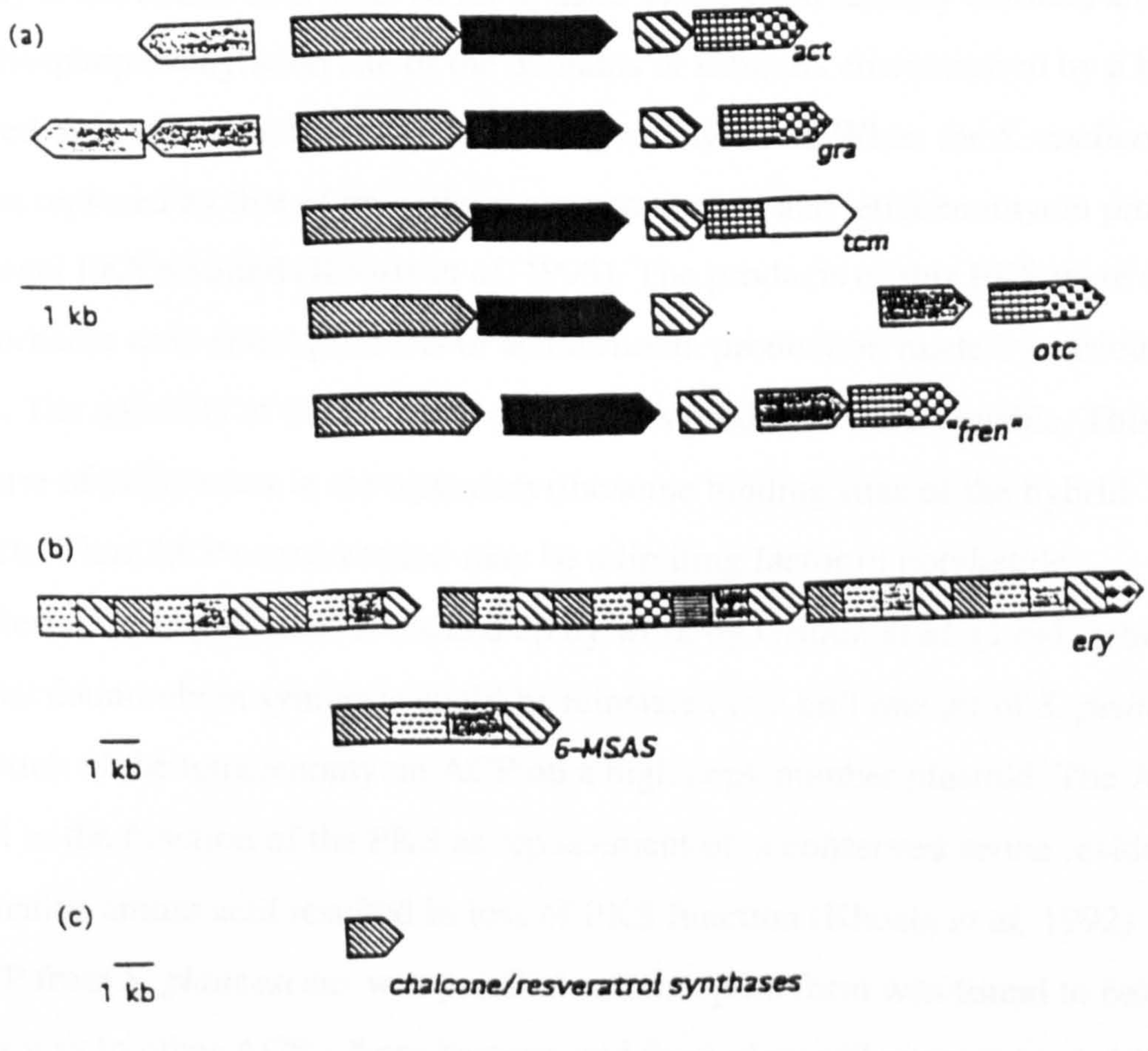



Fig 1.6

Linear alignment of polyketide synthase genes (Hopwood & Khosla, 1992). from (a) Type II PKS (*act*, actinorhodin; *gra*, granaticin; *tcm*, tetracenomycin; *otc*, oxytetracycline and *fren*, frenolicin); (b) Type I PKS (*ery*, erythromycin and *6-MAS*, 6 methylsalicylic acid); (c) Type III PKS

### 1.3.2.2 The Acyl Carrier Protein (ACP)

All the ACP's from fatty acid and polyketide synthesis have shown substantial similarity at the amino acid level so far to date. The highest identity is found around the 4' phosphopantenylation site of the domains or subunits characterised by a highly conserved serine residue, the putative site of pantenylation. When the *S. coelicolor* ACP was replaced by that of granaticin, oxytetracycline and tetracenomycin producers, a functional PKS resulted (Khosla *et al*, 1993). The products of this PKS were similar to actinorhodin or to shunt products of actinorhodin production made by various *act* mutants. The amounts of the different compounds produced were variable. This may be because of differences in the upstream ribosome binding sites of the hybrid constructs. Thus ACP concentration may be a limiting factor in polyketide biosynthesis. This hypothesis is backed up by work by Grimm *et al*, (1994) which stated that daunorubcin synthesis could be reinstated in a null mutant of *S. peuticus* by introduction of the tetracenomycin ACP on a high copy number plasmid. The ACP is essential to the function of the PKS as replacement of a conserved serine residue with an alternative amino acid resulted in loss of PKS function (Khosla *et al*, 1992).

The ACP from *S. glaucescens* was purified and this pure form was found to behave in a similar way to other ACP's from bacteria and from plants (Summers *et al*, 1995). The organisation of *S. glaucescens fabD, H, C* and *B* genes was shown to be very similar to that of the FAS genes of E. coli (*fabH, D, F, G,* and *acpP*). This suggests that the PKS genes of *S. glaucescens* may be similar to those involved in fatty acid synthesis.

## 1.4 The OTC cluster

The major steps in biosynthesis of tetracycline antibiotics were elucidated in the late 1960's from studies on the chlortetracycline producer, *Streptomyces aureofaciens.* (McCormick, 1966, 1968). At that time the genetics of industrial *Streptomyces* spp. were not well established, preventing studies on the biosynthesis genes. During the 1970's genetic investigations of *Streptomyces* spp. began to develop, including genetic analysis of the oxytetracycline producer *Streptomyces rimosus* (Friend & Hopwood, 1971). In 1979, Pigac & Alacevic showed that the genes for OTC biosynthesis were located on the chromosome of *S. rimosus.* By the early 1980's, genetic manipulation

of *Streptomyces* was well under way. In 1981, Rhodes and his coworkers at Pfizer, UK isolated a series of OTC production negative mutants and classified them into nine separate groups according to the results of cosynthesis tests. The nine groups could then be further subdivided into two classes, class I mutants, which were the convertors and class II mutants which were the secretors. Experiments using $^{14}C$ labelled acetate enabled the labelling of the pool of OTC precursors within the cell. This mycelium was then washed and transferred to media containing growing unlabeled mycelia of the other non-producing strain. It was envisaged that if cosynthesis were occurring by transfer of an intermediate between strains, labelled OTC would be produced when the excretor was labelled but the product would not be labelled if the convertor was the labelled strain.

Previous work by Miller *et al* (1960) had shown that certain *S. aureofaciens* mutants did not produce chlortetracycline because they lacked a flavin-like cofactor, termed Cosynthetic factor 1, (CSF1). This is required in catalytic amounts for the final step in the biosynthetic pathway. As an identical step also occurs in OTC biosynthesis, Rhodes and co-workers investigated the importance of CSF 1 in *S. rimosus*. Culture filtrate from a CSF1 producing strain of *S. aureofaciens* was supplied to each OTC mutant. Results showed that there was a requirement by *S. rimosus* for catalytic amounts of the active substance (Rhodes *et al*, 1981). Further studies identified the Class I mutants as being defective for CSF1 whilst class II mutants produce their own CSF 1. Attention was then switched to the Class II OTC mutants.

In order to determine where in the OTC pathway (Fig 1.7) the various mutants were blocked, known OTC intermediates were added to liquid cultures of all the class II mutants. When the supplied compound was converted to OTC, the mutant was presumed to be blocked at the previous step to the added intermediate. All but one of the class II mutants (*otc*-25) could convert 7-chloro-dehydrotetracycline (7- chloro-DHTC) to 7- chloro-oxytetracycline (7- chloro-OTC) via 7- chloro-dehydro-oxytetracycline, (7- chloro-DHOTC). This mutant was assumed to be unable to perform the DHOTC to OTC conversion (step 9) rather than being blocked in the previous step. As DHTC naturally converts to tetracycline (TET), this product was not identified. This mutation was assigned to a locus called *otcA*. Another class II mutant (*otc*-75) could carry out the conversion of 7-chloro-DHTC to 7-chloro-OTC but not

ATC to OTC, thus it was blocked in the C-6-hydroxylation of ATC (*otcC*). A further mutant, (*otc*-4) was able to convert anhydrotetracycline (ATC) to oxytetracycline but not 4-amino anhydrotetracycline to OTC demonstrating a formation of 4-amino-ATC this mutant was assigned to locus *otcD*. None of the mutants that appeared to be blocked before 4-amino ATC (*otc-20, -90* and *-151*) could convert the earlier precursors 6MPT or 4-hydroxy-6MPT to OTC. As thse mutants could cosynthesize with each other they were deduced to be blocked at different steps and were assigned to loci *otcX, Y* and *Z*. The deduced product of *otcZ* shows similarity to hydroxyindole o-methyl transferases and to the carboxyl domain of a putative o-methytransferase. Since the *otcZ* mutant was blocked somewhere before 4-amino-ATC (Rhodes *et al*, 1981) it was suggested that its gene product was involved in 6-methylation of the pretetramid structure (polyketide backbone). However as the mutant could convert neither 6MPT or 4-amino-6MPT to OTC the implication is that it contains mutations at more than one locus. Transcriptional analysis (McDowall *et al*, 1991) has revealed the existence of two divergent promoters, one which directs transcription of a polycistronic message containing *otcC-otcZ-otrA* and the other direction transcription of *otcX* genes. The product of *otcX*-ORF1 shows similarity to actVA-ORF2 sequence of actinorhodin biosynthesis (Caballero et al, 1991). These two genes are likey to be involved in ring-hydroxylation reactions.

From work done by Aranthip Thamchaipenet in this laboratory, it is now known that *otcD* contains two ORF's, *otc*D ORF1 which encodes a putative cyclase/deydrase enzyme predicted to catalyse one or more of the aldol cyclisations of the polyketide chain before the pretetramid is formed, and *otc*D ORF2 which lies directly downstream from *otc*D ORF1 and encodes a product similar to several bacterial hydroxylases including the TcmG hydroxylase of the tetracenomycin cluster (Decker & Hutchinson, 1993). This protein is predicted to act as a hydroxylase either before the formation of 4 amino ATC or in the later hydroxylation step at the fifth carbon position of OTC. These findings are in dispute with the original assignment of *otcD* as a methylase. The current model of the OTC cluster, ( Fig 1.8) includes two methylase genes, *otcZ* as discussed above, and *otc*Y1 ORF5, which is thought to fulfill the second methylation requirement of OTC biosynthesis in converting 4 amino ATC to ATC.

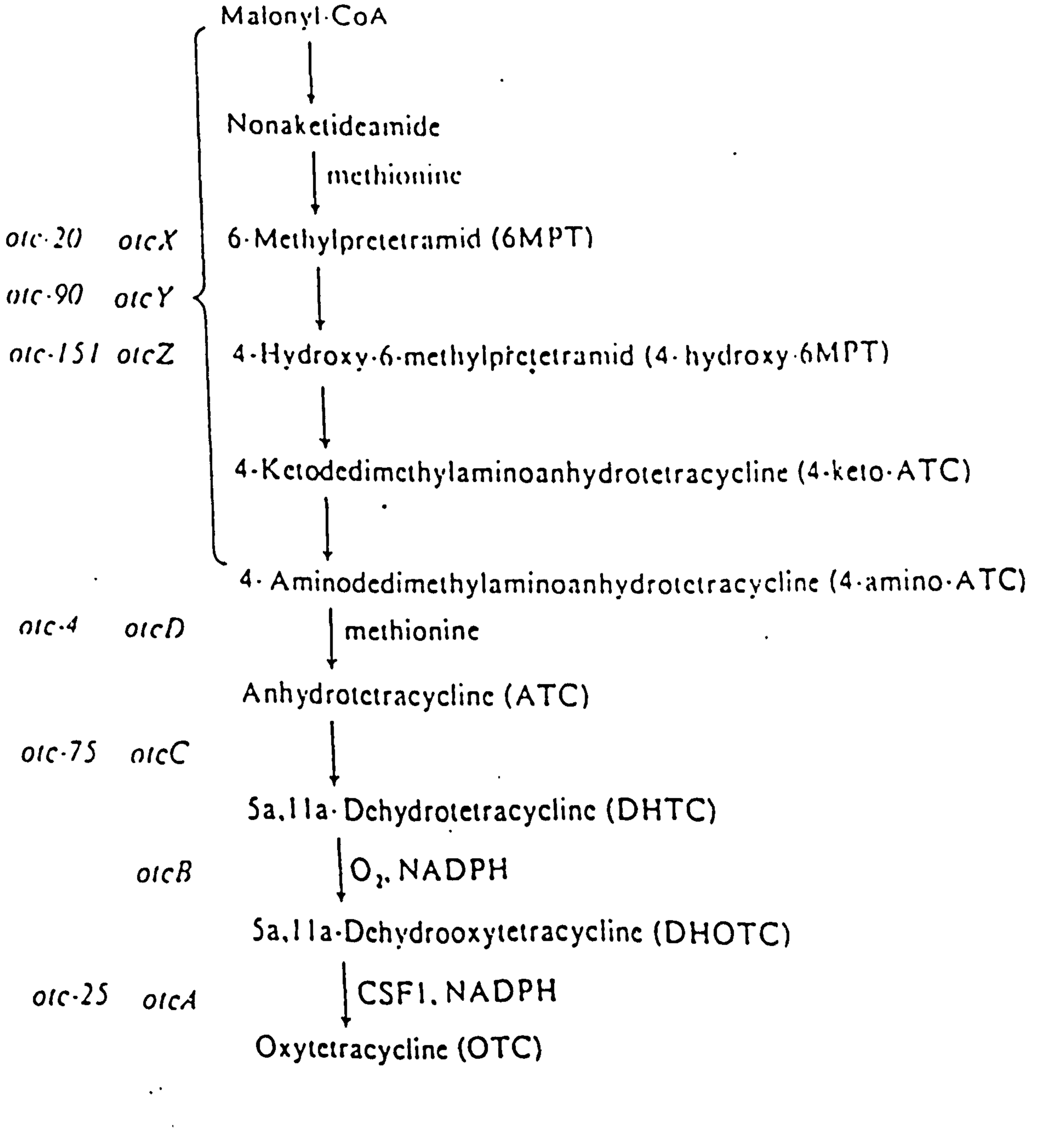



Fig 1.7

Oxytetracycline pathway and assignment of blocked mutants (taken from Rhodes *et al*, 1981).

Further work by Thamchaipenet (1994) has also gone some way to revealing the functions of the *otcY* locus. Initial work probing *S. rimosus* genomic DNA with ketosynthase genes from the *S. coelicolor act* cluster showed *otcY* to be present as two clusters, *otcY1* and *otcY2* separated by 8Kb of chromosome, contrasting with the organisation in *S. coelicolor* where the genes lie adjacent to one another. The *otcY2* region was sequenced (Thamchaipenet, 1994) revealing the presence of four open reading frames, discussed below.

**_otc_Y2 orf1:**

Encodes a polypeptide with considerable homology to the *actIII* gene product of *S. coelicolor,* the putative ketoreductase enzyme and also to a similar peptide from the *gra* cluster. It is suggested that *otcY2* orf1 encodes a protein that catalyses the ketoreduction at the ninth carbon from the end of the polyketide chain.

**_otc_Y2 orf2:**

The deduced peptide sequence of this gene showed some similarities to the β subunit of benzene and toluene dioxygenase from *P. putida* and also to the *act*VI and *fren orfX* gene products, the functions of which are not known.

**_otc_Y2 orf3:**

The putative product of this gene displayed similarities to the adenylate forming enzyme super-family, members of which originate from a variety of sources. This gene is thought to encode a malonamate-CoA ligase, probably involved in malonamyl CoA (starter unit) formation.

**_otc_Y2 orf4:**

Lying downstream of *otcY2 orf3*, *otcY2 orf4* encodes a small protein, highly similar to one of the *S. aureofaciens* tetracycline dehydrogenase subunits. This subunit is involved in the final step of antibiotic synthesis in that pathway. A similar function can be postulated for the *otcY2 orf4* gene product in the OTC pathway.

The position of this "late" gene amongst genes which encode rather more "early" functions in OTC synthesis is interesting.

Subsequent to this, the region between *otcY1* and *otcY2* was also sequenced, (Thamchaipenet, this laboratory) and was shown to contain two *otcY* ORFs, that did not hybridise to the *act* ketosynthase probe, *otcY2-4* and *otcY2-5* are thought to encode dehydrase and methylase respectively. Also present in this region are two further *otcY1* ORFs, *otcY1-4*, thought to encode asparaginase and *otcY1-5*, thought to code for hydroxylase. Once all the sequencing data was collated it became apparent that *otcY1* and *otcY2* lie next to each other on the chromosome as in other polyketide biosynthesis cluster found in *Streptomyces*

## Architecture of *otc* cluster

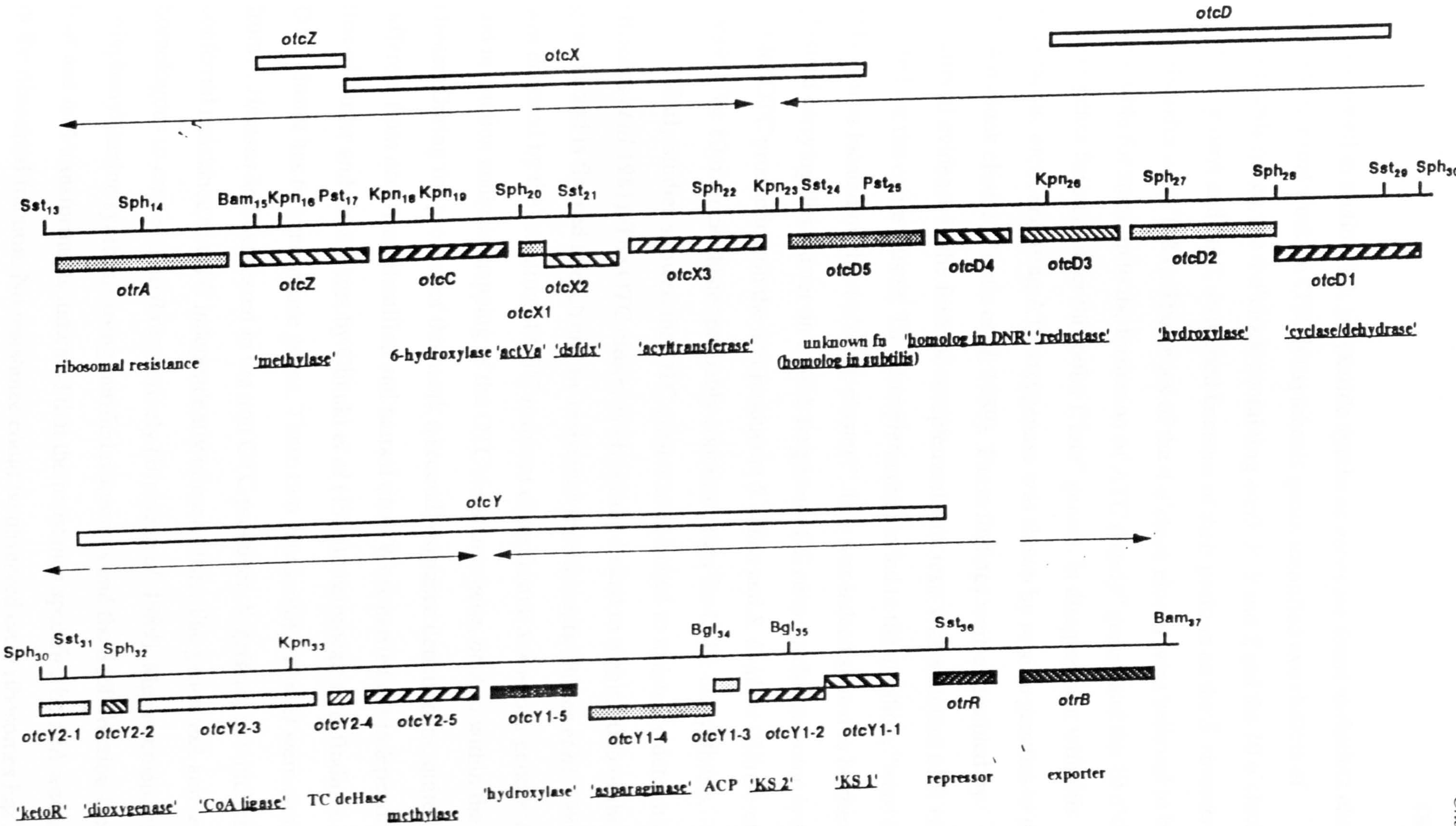



Fig 1.8 Architecture of the *Streptomyces rimosus* OTC biosynthesis cluster (Hunter & Hill, 1997).

As discussed in section 1.3, polyketide synthesis genes are found as distinct clusters. However, initial work on OTC biosynthesis genes identified two clusters of biosynthesis genes, the 4 o'clock containing *otcD, X, Y* and *Z* and the 10 o'clock containing *otcA* and *otcC*, so named because of their position on the *S. rimosus* genetic map (Rhodes *et al*, 1984). The genes of the 4 o'clock cluster were believed to be responsible for steps up to the formation of ATC ("early" genes) and the 10 o'clock cluster genes for steps after this point ("late" genes). In disagreement with this hypothesis, *otcC*, encoding ATC oxygenase was shown by reverse genetics to map to the 4 o'clock cluster (Binnie *et al*, 1989). These findings were augmented by functional evidence in the form of complementation tests with a mutant of *S. rimosus* deleted for the entire cluster. This complementation led to the question, "were all the OTC genes located on the 4 o'clock cluster?" This was indeed shown to be the case. A plasmid carrying this cluster on a 34kb fragment of *S. rimosus* chromosome is able to confer OTC production on the non producers *S. albus* and *S. lividans* (Binnie *et al*, 1989). The 10 o'clock cluster probably contains genes for CSF1 biosynthesis.

Early studies showed the OTC cluster to be linked to resistance determinants (Rhodes *et al* 1981). The OTC cluster is physically linked to within 30kb of the genome and is flanked at each end by resistance determinants (Butler *et al*, 1989). This was deduced by the isolation of OTC resistant clone from a *S. rimosus* genetic library, and restriction analysis mapping of the OTC resistance gene, *otrA* to within the OTC cluster. During the process of this work a second resistance determinant, quite different from *otrA* was identified and named *otrB*. Work carried out independently of that of Butler and co-workers by Ohnuki *et al* (1985), reciprocates the findings that the OTC cluster has two resistance genes. These two genes, *tetA* and *tetB* were cloned from *S. rimosus* and expressed in the non OTC producer, *S. griseus*, in which they conferred resistance to OTC independently of each other. The genes *tetA* and *tetB* are homologous to *otrA* and *otrB* respectively (Binnie *et al*, 1989). Measurement of polyphenylalanine synthesis using purified ribosomes and the S150 fraction from both host and *tetA* transformants indicated that the resistance specified by *tetA* was mainly in the ribosomal fraction. No resistance could be observed once ribosomes has been washed in ammonium chloride. Resistance encoded by *otrA* is mediated by a cytoplasmic factor but not a constituent of the ribosome (Ohnuki *et al*, 1985).

The sequence of *otrA* has been determined (Doyle, 1987). The predicted amino acid sequence shows high homology with TET resistance genes from other bacteria that mediate their resistance through non-covalent modification of ribosomes, e.g. *tetM* of *Streptococcus* and *tetO* from *Campylobacter* ( Doyle *et al*, 1991). The N-terminal domain of OtrA shows marked similarity with the GTP binding sites of elongation factors required for protein synthesis, suggesting that GTP hydrolysis may be important in the function of OtrA. Genes homologous to *tetM* and *tetO* are present in a number of bacterial spp. (Salyers *et al*, 1990). It is clear that they are all related to *otrA* . Indeed it has been speculated that resistances of this type found in clinical isolates may be due to horizontal transfer at some point in the past, followed by sequence divergence (Benveniste & Davies, 1973). Promoter analysis of *otrA* has lead to the finding that during vegetative growth, *otrA* is transcribed from its own promoter *otrA p1* (Doyle *et al*, 1988), with the transcription start 127 nucleotides upstream of the translational start site. *otrA p1* is similar to the consensus sequence of many *E. coli* promoters. During antibiotic production this promoter is no longer used and *otrA* is transcribed as part of a polycistronic transcript which also includes *otcC* and *otcZ*. This change in promoter activity may reflect the use of multiple sigma factors as seen in morphological differentiation of other *Streptomyces* spp. In contrast to these results, the protein synthesis of *S. griseus* carrying the *tetB* gene was sensitive to OTC. Measurement of uptake of labelled OTC revealed that the *tetB* cells accumulated less intracellular OTC than their *tetA* counterparts. This reduced accumulation was also observed for *S. rimosus*. It was then suggested that *otrB (tetB)* encoded a protein responsible for the reduced accumulation of OTC inside the cell.

Efflux as a tetracycline resistance mechanism has been widely studied in the enteric bacteria (McMurray *et al*, 1980; Kaneko *et al* 1985; Chopra,1986). *otrB* was cloned into *E. coli* and sequenced as *tet347* (Reynes *et al*, 1988). The deduced 347 amino acid peptide sequence shows a limited but significant homology to other tetracycline resistance proteins characterised at the time. Later work by MacGregor-Pryde (1995) in this laboratory showed that this sequence was truncated and contained some mistakes. This work also presented preliminary analysis of the deduced *otrB* product, thought to be a transmembrane protein containing 14 membrane spanning alpha helices (MacGregor-Pryde, 1995). Further work to clarify sequence

details is presented in Chapter 3 of this thesis, resulting in a modified *otrB* sequence. (Fig 3.10). The *otrB* gene and gene product are discussed more thoroughly in the experimental chapters of this thesis.

## 1.5. Antibiotic resistance

Organisms that produce antibiotics must somehow render themselves resistant to their products if they are to survive. Early ideas on antibiotic production were that such molecules were only produced once the growth phase comes to an end. However, in some organisms antibiotic production overlaps with growth phase, even though the maximum production of antibiotic still occurs as the organisms reach stationary phase. In such organisms, expression of genes that encode resistance functions should be constitutive. Although many producer organisms comply with the early hypothesis, constitutive expression of resistance genes does occur in some organisms (Cundliffe, 1989). Antibiotic producers have a range of defensive options open to them. Mechanisms of resistance employed by producer organisms include, (1) Modification of intracellular drug molecules causing inactivation, (2) Modification of the target site, (3) Synthesis of novel cellular components in place of the drug target and (4) Efflux of the drug from the cell.

The mechanisms of resistance discussed below are all taken from *Streptomyces* spp. as befits the context of this thesis.

### 1.5.1 Modification of drug molecules.

Antibiotic modification in producer organisms usually occurs due to one of two main mechanisms, N-acetylation of amino groups from an acetyl CoA donor or O-phosphorylation of hydroxyls at the expense of ATP, these methods of resistance are also found in some clinical isolates. Although these are the main modification steps other methods are also used; O-adenylation is common amongst clinical strains as a resistance mechanism (Hollingshead & Vapnek, 1985) but has not been found in antibiotic producers (Cundliffe, 1989). Another modification to attain drug resistance is the O-acetylation of chloramphenicol. Although this is widespread, both in Gram negative and positive bacteria including *Streptomyces* it has not been found in organisms that produce the drug. (Nakano *et al*, 1977). Examples of drugs that are

inactivated by modification are the aminoglycosides neomycin and kanamycin. Streptomycin is another example. Neomycin producers such as *S. fradiae* have aminoglycoside phosphotransferase (APH) and acetyltransferase (AAC) activities. The ribosomes of these organisms are sensitive to the drug throughout the life cycle. It is thought that the two modifying enzymes may act synergistically to bring about full resistance as strains of *S. lividans* expressing only one of the genes from *S. fradiae* were less resistant than strains transformed with both (Thompson *et al*, 1982).

The producer of kanamycin, *S. kanamyceticus*, displays AAC(6') activity although its function does not appear to be resistance. Activity of the enzyme is at a maximum during the growth phase and declines just prior to drug production (Satoh *et al*, 1975). Extra copies of the AAC gene resulted in higher yields of kanamycin and when transformed into *S. fradiae*, in higher neomycin yields, even though this activity is not normally present in that organism (Crameri & Davies, 1986). It seems then that AAC activity may be involved in acetylation of intermediates in kanamycin production, its role in neomycin production is unclear.

In *S. griseus* and *S. bikiniensis*, streptomycin phosphorylation occurs. The principal enzyme in both species is SPH(6), originally identified as the product of the *aphD* gene of *S. griseus*. SPH (3'') and (3'α) phosphotransferase activities have also been found in *S. griseus* and *S. bikiniensis* respectively (Walker & Skorvaga, 1973a, 1973b). Other examples of antibiotic inactivation by modifcation are the 7''-O-phosphorylation of hygromycin B in *S. hygroscopicus* (Leboul & Davies, 1982) and puromycin-N-acetyltransferase activity in *S. alboniger* (Sugiyama *et al*, 1985).

### 1.5.2 Modification of target sites.

Higher levels of resistance can be achieved by this mechanism than are reached by antibiotic inactivation (Cundliffe, 1989). One of the most well documented modifications of cellular components in producer organisms is that of the ribosome. Ribosomal modification is common in those Streptomycetes that produce antibiotics whose mode of action is inhibition of protein synthesis by binding to ribosomes. Not all strains that produce this type of antibiotic protect themselves by ribosomal modification, e.g. streptomycin in *S. griseus* as discussed above. Analysis by Calcutt & Cundliffe (1989) has shown that ribosomal resistance is most often due to some

property of the rRNA. Resistance in several *Streptomyces* strains is mediated by methylation of the rRNA. In most of the organisms identified, ribosomal resistance can be attributed to the product of a single methylase gene acting upon a single specific site in the rRNA resulting in resistance to a certain antibiotic or group of antibiotics (Cundliffe, 1989). There are however exceptions: *S. tenebrarius* has two RNA methylase genes and displays resistance to a much larger array of drugs.

Another method of ribosomal protection in Streptomycete antibiotic producers is by interaction with soluble cytoplasmic proteins, as exemplified by the *otrA* gene product in *S. rimosus*. The exact way in which resistance is bestowed on the ribosome by these interactions is not yet known. The 72 kDa OtrA product shows similarity to the TetM and TetO resistance proteins of Gram-positive bacteria. These three peptide sequences all also show some sequence identity with the GTP binding sites of known elongation factors EF-Tu and EF-G (Doyle *et al*, 1991; Salyers *et al*, 1990). However although it hydrolyses GTP, TetM does not substitute for EF-Tu in *Bacillus subtilis* or *E. coli* (Burdett, 1991).

### 1.5.3 Synthesis of novel resistant cellular components

This type of antibiotic resistance is typified by the novobiocin producer, *Streptomyces sphaeroides*. Novobiocin acts by inhibiting the β subunit of the DNA gyrase enzyme. This enzyme is a heterodimer comprising $\alpha_2 \beta_2$ subunits, it acts as a topoisomerase, introducing negative supercoils into covalently closed circular DNA. The target of novobiocin is the β subunit, an ATPase which energises the nuclease reaction carried out by subunit 1. In *S. sphaeroides* there is a novel DNA gyrase β subunit, *gyrB'* whose function is not affected by novobiocin and is induced in the presence of the drug. Peculiarly, *S. sphaeroides* also possesses a novobiocin sensitive β subunit. It is unclear why this has not been superceded in the cell by the resistant version but may concern the status of the DNA superhelicity. Regulation of the resistant *gyrB'* subunit is controlled by a promoter that is sensitive to the supercoiling of the DNA at that locus, *gyrB'* is induced by a reduction of superhelicity, a phenomenon that can result from the action of novobiocin.

### 1.5.4 Export of the drug

Each organism that produces an antibiotic possesses a mechanism for secretion of the final product. This mechanism can also be looked upon as a resistance mechanism exporting any molecules which have re-entered the cell. Export as a resistance mechanism is prevalent among clinical strains and very well documented in the literature, see section 1.6.2 for a discussion of tetracycline efflux.

Drug efflux in the *Streptomyces* was first identified in *S. coelicolor*, which produces a number of antibiotics including methylenomycin. The methylenomycin biosynthetic genes are entirely plasmid-borne, as is the resistance determinant *mmr*. The amino acid sequence of the Mmr protein was shown to have many similarities to the export genes of Gram negative bacteria (Neal & Chater, 1987). It is now known that Mmr and several other *Streptomyces* drug exporters, including OtrB, the focus of this study, belong to a wide ranging and rapidly expanding family known as the Major Facilitator Superfamily ( Paulsen *et al*, 1996). The MFS and its members are discussed thoroughly in section 1.7.

## 1.6 Tetracycline & Tetracycline resistance.

### 1.6.1 Tetracyclines

Tetracyclines are a group of polyketide antibiotics based on a four ring structure (tetra = 4), all of which are produced by Streptomycetes with *S .aureofaciens* and *S. rimosus* being the most commercially important producers. *S. rimosus* is the producer of oxytetracycline *S. aureofaciens* strains produce tetracycline and chlortetracycline, the ratio of which can be varied with the concentration of chloride ions present in the growth media. Tetracyline are important broad spectrum antibiotics that are still in wide use today despite being amongst the first antibiotic drugs discovered. As well as the molecules mentioned above, the tetracycline group also includes demethyltetracyclines (McCormick *et al,* 1957) and decarboxymide tetracycline derivatives, (Hochstein *et al*, 1960). Modern semisynthetic tetracycline analogues include doxycycline, minocycline and the group of drugs known as the glycylglycines (Fig 1.9). As well as being active against Gram-positive and negative bacteria, tetracyclines are also used to combat infections by spirochaetes, mycoplasmas and rickettsiae. Although they are all closely related in structure they differ in

antimicrobial activity and therapeutic effect. CTC tends to be used on Gram-positive cocci whereas TC is said to be more active against Gram-negative bacilli (Durckheimer, 1975).

The mode of action of tetracyclines is the inhibition of bacterial protein synthesis, with the target site of the drug localised on the ribosome. The drug works by interfering with binding of the ternary complex aminoacyl-tRNA-EFTu-GTP on the acceptor site of the 30s subunit of the ribosome. Atypical tetracyclines such as chelocardin and anhydrotetracycline exhibit bactericidal activity but are poor inhibitors of protein synthesis (Chopra, 1994). It is suggested that these drugs target the cytoplasmic membrane rather than the ribosome (Oliva *et al*, 1992), as they cause alterations in cell morphology and trigger the release of β-galactosidase from the cytoplasm. As would be expected the well-documented tetracycline resistance determinants that operate by efflux or ribosomal protection do not impart resistance to these compounds (Oliva & Chopra, 1992). Such molecules are of no therapuetic use due to the severe side effects experienced by users, probably because the action of such drugs on the membrane is not prokaryote-specific (Rogalski, 1985).

Several properties of the drugs make for near-perfect therapuetic agents,

1) They are active against a broad range of pathogens

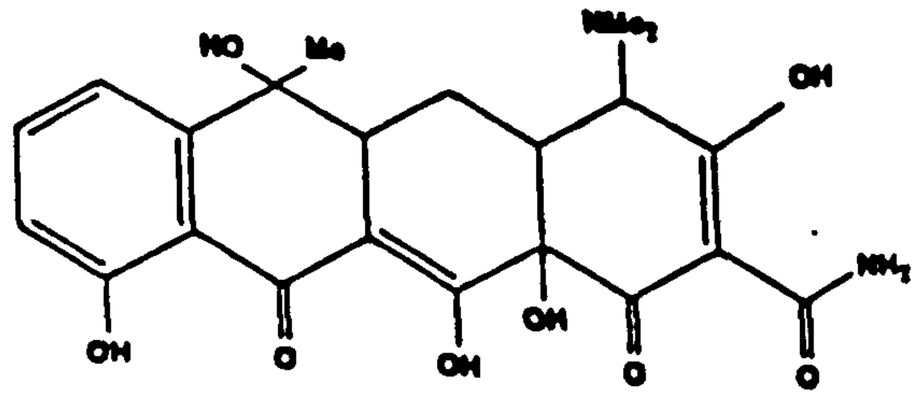
2) They are well absorbed when administered orally

3) They are relatively non-toxic

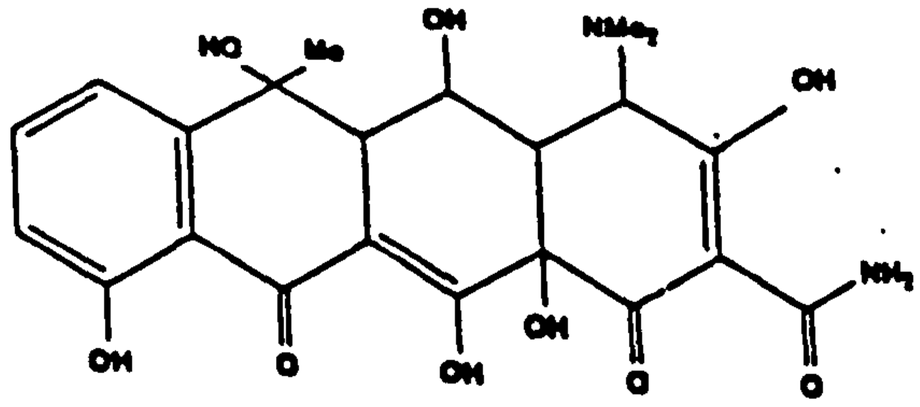
4) They cause few allergic reactions

5) They are cheap to produce

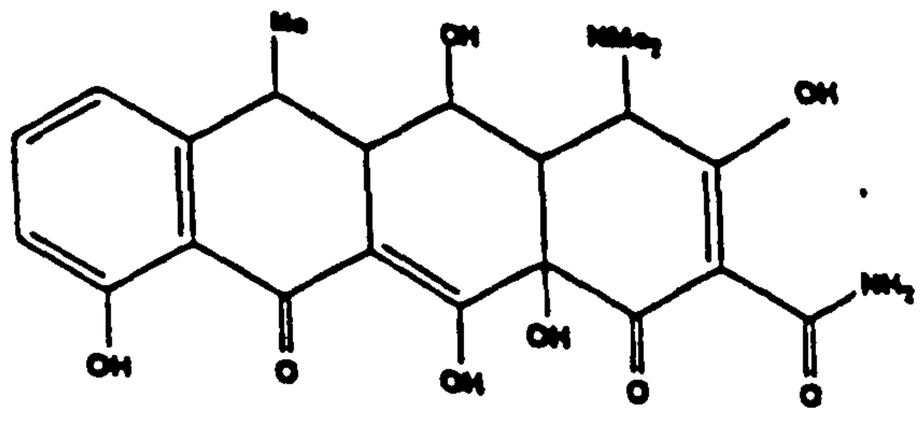
These properties have led to massive use of the drugs for over 40 years, not only in medicine but also as food supplements in animal husbandry and fish farming. As is to be expected, due to the selective pressure caused by such extensive use a growing number of bacterial species are acquiring resistance to tetracyclines.



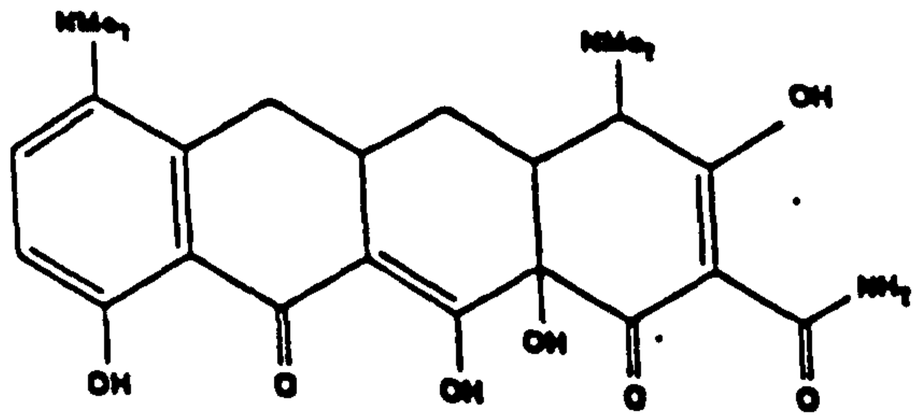
Tetracycline



Oxytetracycline



Doxycycline



Minocycline

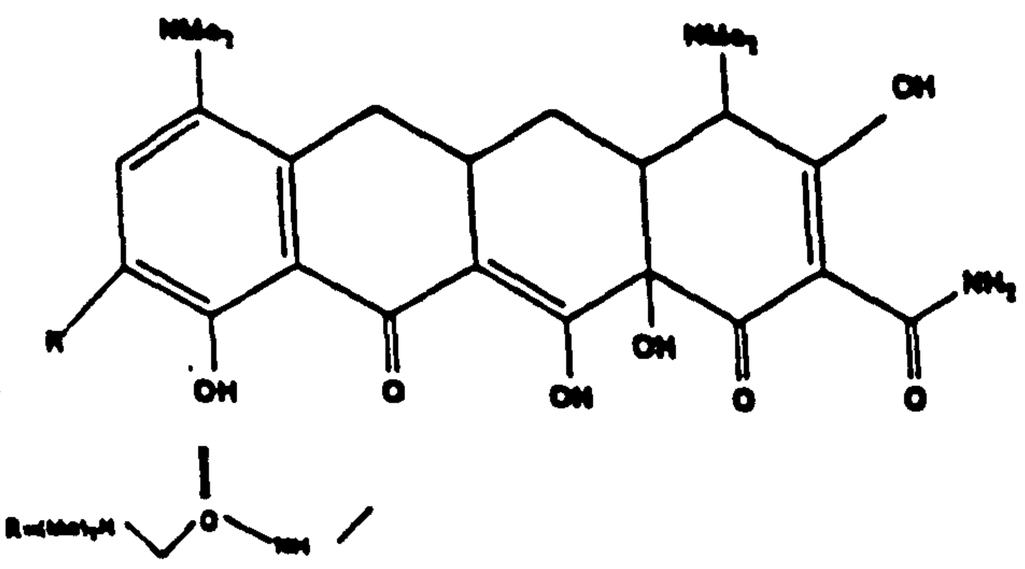


Fig 1.9 DMG-Mino 9-(dimethylglycylamido)minocycline

Tetracycline and tetracycline analogues; Tetracycline, Oxytetracycline, Doxycycline, Minocycline and DMG-Mino (a member of the glycylglycine group of drugs).

### 1.6.2 Resistance to tetracyclines

Up until the start of the 1990's, reduced accumulation was the only mechanism of tetracycline resistance found. Resistance due to tetracycline efflux in this way is the most dominant resistance mechanism amongst the Enterobacteriacae but is not the main method of resistance in other groups of bacteria.

There are three mechanisms of tetracycline resistance observed so far in bacteria, (1) reduced cellular accumulation (efflux), (2) ribosomal protection, (3) antibiotic inactivation. Classification of tetracycline resistance proteins is based on the mechanism of resistance with each *tet* class being designated a capital letter. Proteins of the same class are grouped together in the alphabet (Levy *et al*, 1989). Although resistance determinants of producer organisms are usually physically linked in the genome to biosynthesis genes this is obviously not the case in non antibiotic producers. In these organisms, resistance genes are often located on transposons, conjugative plasmids or other such mobile genetic elements (Roberts, 1994).

#### 1.6.2.1 Tetracyline efflux

There are presently 11 classes of tetracycline-specific efflux pumps, TetA(B-E,) (G), (H), K, L, P and OtrB. (Table 3.1) (TetA(B-E) and TetA(G) and (H) occur in Gram-negative bacteria, the amino acid identities between these proteins range from 45-78%. TetK, L, P and OtrB are all found in Gram-positives but show only slight similarities in their amino acid sequences (Sheridan & Chopra, 1991). These proteins do not constitute a family of homologous proteins, although members of the TetK and TetL groups do show significant sequence identity with each other (64-80%; Ives & Bott, 1990). TetP and OtrB on the other hand show no more similarity to TetK/L than they do to the Gram-negative *tet* proteins (Sloan *et al*, 1994; this work).

Efflux determinants from Gram-negative bacteria share a common genetic organisation in that *tetR*, a repressor gene is located next to *tetA* in a divergent orientation. The TetR protein binds to two tet operator sequences located in the intergenic region in the absence of tetracycline and by doing so, represses transcription of *tetA* and *tetR*. When tetracycline is present, the TetR-[TET-$M^{2+}$] complex, consisting of the repressor protein, tetracycline and a metal dication, is formed. This allows coinduction of *tetA* and *tetR* expression.

It was thought that no such transcriptional repressor protein existed in the Gram-positive tetracycline resistant bacteria (Schnappinger & Hillen, 1996). However, since 1991, it has been recognised that antibiotic efflux resistance systems from *Streptomyces* spp. probably also possess resistance/repressor systems utilising resistance and repressor genes transcribed divergently from overlapping promoters (Neal & Chater, 1991, Cabellero *et al*, 199a, Guilfoile & Hutchinson, 1992).

The upstream region of the *otrB* resistance determinant from *S. rimosus* was found to encode a repressor-like protein with high similarity to the TcmR protein from *S. glaucesens* (MacGregor-Pryde, 1995). TcmR itself is similar to several TetR repressor proteins, including that from Tn*10*. Also revealed by these studies is the similarity between OtrR and the *marO* gene product of *E.coli*. MarO is a tetracycline inducible multiple resistance protein. All of these proteins contain a helix-turn-helix (HTH) motif similar to that found in the bacteriophage 434 repressors. Analysis of the *otrR/otrB* intergenic region by MacGregor-Pryde (1995) revealed similarity to resistance/repressor regions from other antibiotic-producing *Streptomyces* a. Several of these regions contained palindromic sequences, thought to be operator sequences. The presence of these putative operator sequences coupled with the similarity of OtrR to other HTH-containing repressor proteins led to the assignment of a *tet* repressor function to *otrR*.

If other Gram-positive Tet proteins are not repressor regulated, what is the mode of regulation? TetK and TetL mRNA's show sequence similarity to mRNA's regulated by translational attenuation, which has been suggested as the most probable regulatory mechanism for expression of these genes (Speer *et al*, 1992). The method of regulation of TetA(P) however, seems to be different again. TET resistance from this locus is inducible but mRNA from TetA(P) shows no structures typical for proteins regulated by attenuation. TetA(P) is present on a conjugative plasmid. As plasmids that contain only the resistance genes mediate constitutive resistance, regulation must be by some other method. Sequence analysis by Sloan *et al* (1994) revealed no putative regulatory gene in the vicinity of *tetA(P)*.

### 1.6.2.2 Gram-negative TetA proteins, structure/function studies

Most of the structure-function studies to determine the molecular basis for tetracycline resistance by such resistance "pumps" have been carried out on TetA(B) from the *E.coli* transposon, Tn*10*. There is now a wealth of information concerning the topology, three dimensional structure and putative substrate binding site of the protein and also information concerning energisation of the efflux process, much of which has been elucidated by the group of Yamaguchi and co-workers in Japan. Much of the knowledge regarding TetA(B) can also be applied to the other Gram-negative *tet* proteins.

To summarise, these proteins are a group of membrane proteins comprised of 12 membrane spanning regions, most probably α helices, connected by cytoplasmic and periplasmic loops (Eckert & Beck, 1989). The TetA proteins are metal-tetracycline/proton antiporters that pump their substrate into the periplasm (Thanassi *et al*, 1995). The energy required for this process is provided by the pH gradient across the cytoplasmic membrane but the electrical transmembrane potential is not required. (Kaneko *et al*, 1985). These observations lead to the reasoning that tetracycline efflux by the TetA group of resistance proteins is an electrically neutral process.

Residues essential to tetracycline binding and translocation are situated in various locations in the protein but only one such essential amino acid, D66, is found in the loop regions (Yamaguchi *et al* 1990). D66 is to be found in the cytoplasmic loop that link the second and third transmembrane helices (TMH2 and TMH3). This loop contains a conserved motif, GXXXXRXGR running from G62 -R71, that is common not only to the *tet* proteins but to a much wider series of transport proteins, the Multi-Facilitator Superfamily (see section 1.7 and Chapter 3 for more information). Mutational analysis of this loop has shown that only R70 is essential for function of the protein (Yamaguchi *et al*, 1992a). As mutation of G62 or G69 to amino acids that have a low propensity to form β turns leads to reduced activity of the protein, it is thought that the ability of loop 2-3 to form a β turn is important for protein function (Yamaguchi *et al*, 1992a). The Gram-negative Tet proteins are hypothesised to have evolved from a gene duplication event as they are comprised of

two structurally symmetrical halves (Rubin *et al*, 1990). Loop 8-9 corresponds to loop 2-3 according to this hypothesis (Yamaguchi *et al*, 1992b). However none of the residues in this region is required for protein function (Yamaguchi *et al*, 1993; 1994) It has been suggested that loop 2-3 and loop 8-9 might form the substrate entrance gate (Yamaguchi *et al*, 1990, 1993). Helix 3 is thought to be important for transport of the TET-$M^{2+}$ complex as substitution of S77 with a C residue shows no transport activity. (Yamaguchi *et al*, 1992c). It is thought that side chains of the helix 3 residues protrude into the [TET-$M^{2+}$ ] transport channel. Helix 7 is thought to contain residues important for substrate recognition, TetA(B) variants that confer resistance to DMG-mino, a glycylglycine type drug have been identified (Guay *et al*, 1994). These variants contain mutations in transmembrane helices 7 and 10.

In summary, residues extending over the whole length of the TetA primary sequence contribute to its function as a tetracycline efflux pump. In contrast to the proton antiporter activity of TetA(B), TetA(C) from pSC101 shows potassium uptake activity. TetA(C) mutants which lack the ability to supply tetracycline resistance but still transport $K^+$ have been used to indicate some residues involved specifically with tetracycline transport (McNicholas *et al,* 1992). Some of the findings are different from that regarding TetA(B), e.g. mutations that abolish tetracycline efflux have been located to four residues of connecting loop 10-11 (McNicholas *et al*, 1995). There is evidence that TET resistance proteins of Gram-negative bacteria function as multimers (Curiale *et al*, 1984). A multimerisation domain has been mapped to the N terminal half of TetA(B) (McMurray & Levy, 1995), suggesting a role for oligomerisation in TetA function. A model of the two dimensional topology of tetracycline specific efflux pumps is shown in Fig 1.10

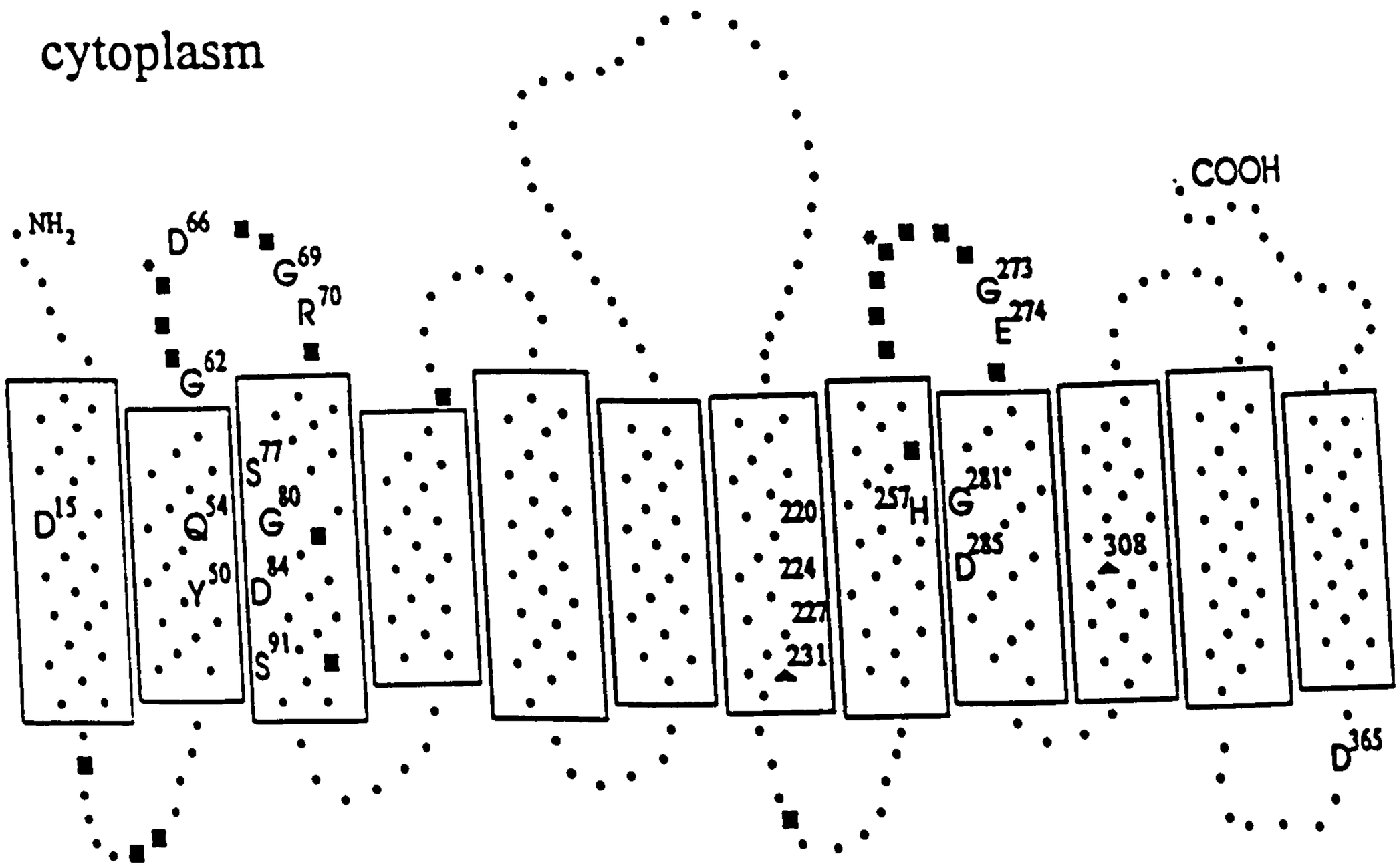



Fig 1.10

Two-dimensional model of the topology of tetracycline specific efflux pumps. Amino acids, given by one letter abbreviations are given for positions at which mutations result in at least 80% loss of activity in TetA(B) (modified from Schnappinger & Hillen, 1996).

### 1.6.2.3 Gram-positive Tet proteins

Much less is known about the molecular basis of TET resistance in the Gram-positive derived exporters. Studies of the TetA(L) protein from *Bacillus subtilis* have identified marked differences to the Enterobacteriacae TET exporters. TetA(L) is chromosomally encoded and will only confer significant levels of tetracycline resistance on the cell when present in multiple copies after an amplification event (Ives & Bott, 1990). Studies by Guffanti & Krulwich (1995) show that TetA(L) is not a specific tetracycline/ $H^+$ antiporter but also antiports $K^+/Na^+$ and $K^+/H^+$ .These antiport activities are electrogenic in contrast to the electroneutral antiport of the Gram-negative *tet* proteins. These results firmly support the conclusion that TetA(L) is a multifunctional antiporter. It is suggested that the protein may have two distinct binding domains for its metal and complex organic substrates.

Studies on OtrB have shown that it is no more similar to the Gram-positive *tet* proteins than the Gram-negative tet resistance pumps, showing most similarity to other *Streptomyces* antibiotic efflux pumps (MacGregor-Pryde, 1995; this work). OtrB also shows high similarity to some multidrug exporters, raising the question that it too may be a multifunctional antiporter along the lines of Tet(L).

Tet(K) is also capable of transporting $K^+$ ions in a similar way to the TetA(C) protein (Guay *et al*, 1993b). It was postulated from hydropathy profiles that the Tet(K) and Tet(L) proteins were comprised of 14 transmembrane helices in contrast to the 12 TMH arrangement of the Gram-negative *tet* proteins (Guay *et al*, 1993a, 1993b,) This has now been further demonstrated by the construction and analysis of phoA Tet(K) fusions which firmly demonstrate the presence of 14 transmembrane regions (Ginn *et al*, 1997). Previously little work detailing mutagenesis experiments to identify functionally important residues of any Gram-positive *tet* proteins had been published. However, recent work in the Yamaguchi laboratory has demonstrated that four transmembrane glutamic acid residues of Tet(K), located in helices 1, 5 and 11 are functionally important and can be compared to the conserved transmembrane aspartyl residues of TetA(B) (Fujihara *et al*, 1996). It is important to remember that information gained from studies of Tet(K) may not be applicable to other Gram-positive *tet* proteins as they show so little similarity to each other.

### 1.6.2.4 Tetracycline resistance by ribosomal protection

Ribosomal protection as a method of tetracycline resistance was first observed in the Streptococci (Burdett, 1986). Six classes of proteins that confer resistance in this way are now known: TetA(M), (O)-(Q), (S) and OtrA. (Martin *et al*, 1986; Taylor, 1986; LeBlanc *et al*, 1988; Doyle *et al*, 1991; Dittrich & Schrempf, 1992; Nikolich *et al*, 1992; Charpenteir *et al*, 1993; Sloan *et al*, 1994). Initial work on this group of proteins was done using TetM. Since the other classes in the group have sequence similarity of over 40% (Sloan *et al*, 1994), it is reasonable to expect the protection method to be the same for all *tet* proteins that act by ribosomal protection.

The mechanism of ribosomal protection is discussed in section 1.5.2. In summary, it appears that these TET resistance proteins mediate resistance by inhibiting the binding of the aminoacyl-tRNA:EF Tu: GTP complex to the A site of the ribosome. The exact mechanism of inhibition remains to be determined at a molecular level.

### 1.6.2.5 Tetracycline modification.

A third mechanism of tetracycline resistance is to chemically modify the antibiotic, rendering it non-toxic. This phenomenon was first observed by Park & Levy (1988) who described that transposon Tn*4400* from *Bacteroides fragalis* mediated TET resistance in *E. coli* by degradation of the drug. Independently of these findings it was shown that a closely related *Bacteroides* transposon Tn*4351* also carries a gene for a tetracycline-inactivating enzyme. Before the finding of tetracycline modifying activity from Tn*4400*, Park *et al* had previously shown that Tn*4400* mediated TET resistance by an efflux mechanism (Park *et al*, 1987) and had designated the Tn*4400* encoded TET resistance gene *TetF*, in line with the nomenclature guideline proposed by Levy *et al* (1989). It is now known that Tn*4400* and Tn*4351* encode tetracycline modifying products, which require NADP and oxygen to carry out the modification reaction, the gene was re-named *tet(A)X* to infer a new class of resistance determinant. TetX is presumed to be an NADP requiring oxidoreductase. (Speer *et al*, 1991). The Tn*4400 tetA(X)* provides only a fifth of the resistance levels of that encoded by the Tn*4351 tetA(X)*. This is thought to be due to a 4 base difference in their upstream regions which provides Tn*4351* with a better -35 site, presumably

leading to higher levels of transcription (Speer *et al*, 1991). The efflux properties of Tn*4440* observed by Park & Levy (1988) are encoded on a separate ORF within the Tn*4400* sequence. Although the product of this gene has tetracycline efflux activity it does not confer TET resistance on *E.coli* when present without *tetA(X)* (Speer & Salyers, 1990). The two genes do not act synergistically as the resistance levels attained by *tetA(X)* alone are indistinguishable from those observed when both genes are present (Speer & Salyers, 1990). The efflux activity was able to be increased to that conferred by pBR328 (Manniatis *et al*, 1982) by cloning the region into a high copy plasmid. Surprisingly, despite the increased efflux, cells were still sensitive to tetracycline. The entire efflux determinant present on Tn*4400* is thought to be included in an insertion element. Tn*4400* and Tn*4351* are 90% homologous, the difference being a 0.5kb insertion element in the region of the efflux gene of Tn*4400*. Alternatively, Tn*4351* could be the result of a deletion event in *Tn4400* in which only part of the efflux gene was removed (Speer *et al*, 1991).

## 1.7 The Mutifacilitator Superfamily (MFS)

The tetracycline efflux proteins discussed above are part of a growing and very wide -ranging family of transport proteins. Located in the cell membrane, members of the multifacilitator superfamily transport a variety of substrates, from sugars to Krebs cycle intermediates. Despite the differences in substrates and origin, MFS members have been observed in bacteria, fungi, plants and mammals.

The concept of a superfamily of transport proteins was suggested in 1993 by Marger & Saier, following a symposium on solute transport (Saier, 1990). This described a number of classes of proteins including the homologous glucose carriers of prokaryotes and eukaryotes. Homology between these and other sugar transporters is well documented in the literature, most work having been done by Henderson and co-workers (Henderson & MacPherson, 1986; Maiden, 1987; Maiden *et al*, 1988, Muiry, 1989; Henderson, 1990). Initially, extensive similarities were identified between passive transporters from yeast and mammalian cells and some bacterial proton/sugar symporters (Baldwin & Henderson, 1989). A model was proposed for the topology of the glucose transporter which by analogy was to apply to the AraE and XylE proteins (Baldwin & Henderson 1989). This group of related (Proton motive force) dependent

proteins was extended to include other transport proteins from *E.coli* including the citrate importer and tetracycline antiporter (Henderson & Maiden, 1990). Some conserved features of the sequences were observed, a particularly good example is the motif RXGRR, where R can be substituted by K and X is often a large hydrophobic residue. This motif, has been expanded to GXXXXRXGRR and is present in the cytoplasmic loop connecting TMH2 and 3 in all MFS proteins (Paulsen *et al*, 1996). These early comparative analyses provided a firm background on which to establish the existence of the larger multifacilitator superfamily.

Marger & Saier, (1993) identified five families of transport proteins which they showed by extensive statistical analyses to be members of one large superfamily (Fig 1.11).

Class 1: drug resistance

Class 2: sugar uptake

Class 3: uptake of Krebs cycle intermediates

Class 4: phosphate-ester/phosphate antiport

Class 5 oligosaccharide uptake.

Analysis of Class 1 proteins gave relatively low intraclustal comparison scores, 15-20SD's (standard deviations) compared with scores of 20-110 SD's for families 2 to 5. Thus, members of family 1 show the lowest degree of sequence similarity. To further determine the phylogenetic relationships within cluster 1, analysis was performed on the seven most divergent members of the group: NorA (quinolone resistance in *S. aureus*), TetA (tetracycline resistance in *E.coli*), QacA (antiseptic resistance in *S. aureus*), MmrA (methylenomycin A resistance in *Streptomyces coelicolor*),TetE (tetracycline resistance in *Streptococcus pneumoniae*), EmrA (multidrug resistance in *E.coli*) and Atr-1(aminotriazole resistance in *Saccharomyces cerevisiae*). The outcome of this revealed that the drug resistance proteins of family 1, split into four branches (Fig 1.12), with NorA and TetA the first branch, MmrA and EmrA the second and TetE and Atr-1 the third and fourth respectively. The greatest similarity between the seven proteins was found in their N-terminal regions between residues 100 and 160- now known to be the location of motif C, a motif present in the 12 and 14 TMH families but not the other MFS proteins (Paulsen *et al*, 1996). Further sequence analyses (Griffith *et al*, 1992; Paulsen & Skurray, 1993) led

to a revision in the phylogeny of the MFS and to the proposal by Paulsen *et al* (1996) that the drug exporter family should be further sub-divided into two groups, those proteins with 12 TMH's being one group and those comprised of 14 TMH's the other. This has led to a revised phylogeny, with the MFS consisting of six separate families. (Figs 1.13 and 1.14) each of which contains many separate transport proteins.

With the publication of increasing numbers of entire genome sequences, the number of MFS members is quickly expanding, although many of these proteins are hypothetical transporters of unknown substrate specificity. The number of proteins in the 12 and 14 TMH families alone was reported at over 100 in 1996 and has presumably increased considerably since then.

Family 1 is now the 14 TMH family. Its members include known or probable multidrug efflux proteins, tetracycline resistance pumps from Gram-positive bacteria and vesicular amine transporters from higher eukaryotes (Lineal, 1993), which are involved in neurotransmission but can also mediate drug resistance (Schuldiner, 1994; Schuldiner *et al*, 1995) plus other hypothetical or as yet uncharacterised 12 TMH proteins. Family 2, the 12 TMH family has members which include a number of known and putative multidrug exporters from bacteria and fungi, other drug resistance proteins, many from Gram-positive bacteria including antibiotic producer organisms and as in Family 1, a number of uncharacterised or hypothetical proteins identified by "sequence gazing".

Both families 1 and 2 are comprised of several phylogenetically distinct groups. Family 1 is divided into six clusters on this basis;

a) a cluster of fungal proteins, b) a cluster containing various bacterial proteins including the multidrug resistance proteins Bcr and EmrD plus two hypothetical yeast proteins, c) contains monoamine and acetylcholine transporters from higher eukaryotes, d) a cluster containing two bacterial chloramphenicol resistance proteins plus other bacterial transporters, e) the Gram-negative *tet* cluster which as well as containing tetA(B)-(E) also includes three multidrug proteins from Gram-positive bacteria, and f) this cluster contains some more distantly related members of family 1 such as the multidrug resistance protein LmrP from *Lactococcus lactis* (Bolhuis *et al*,

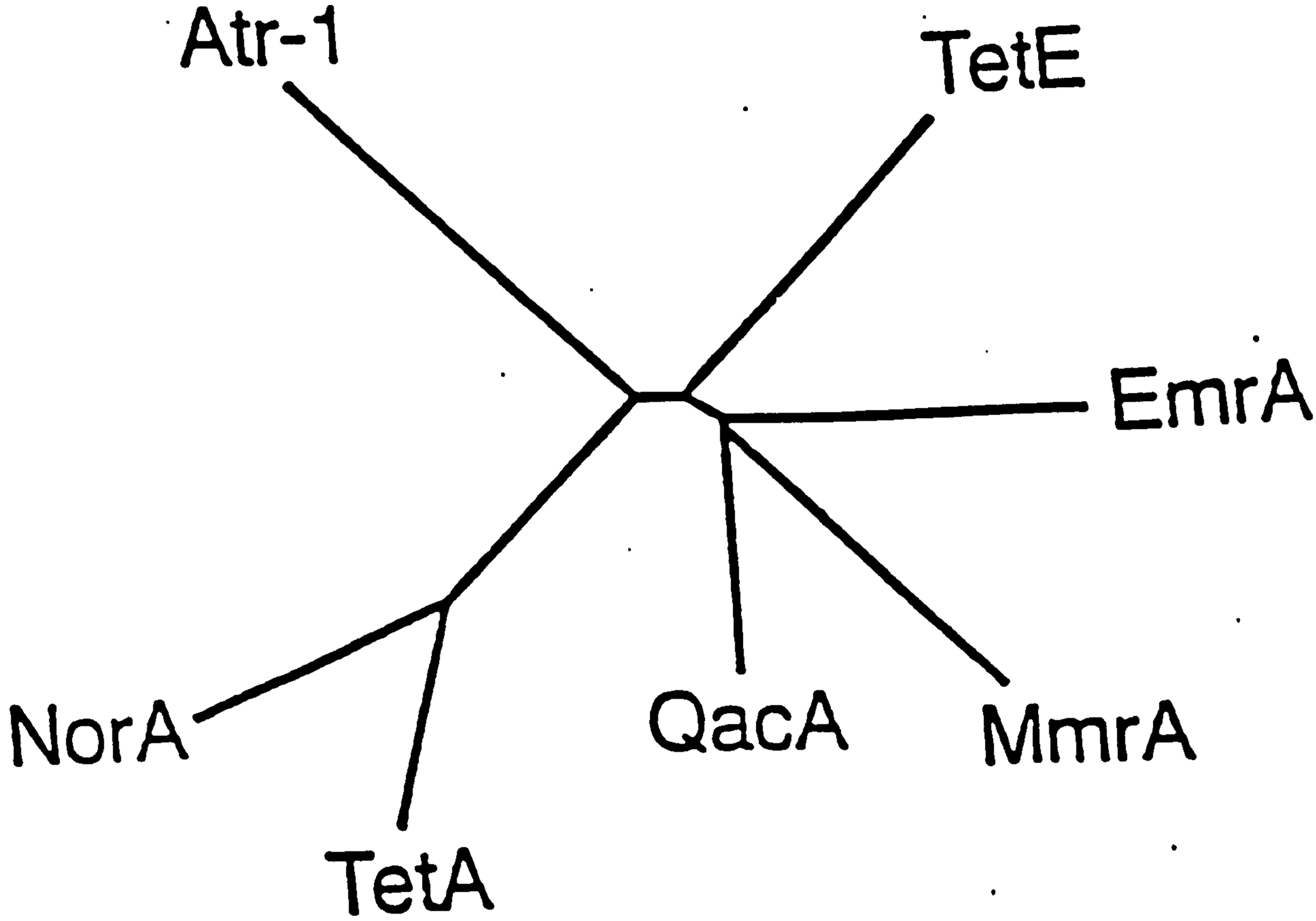



Fig 1.11

Phylogenetic tree of seven transport proteins, (taken from Marger & Saier, 1993).

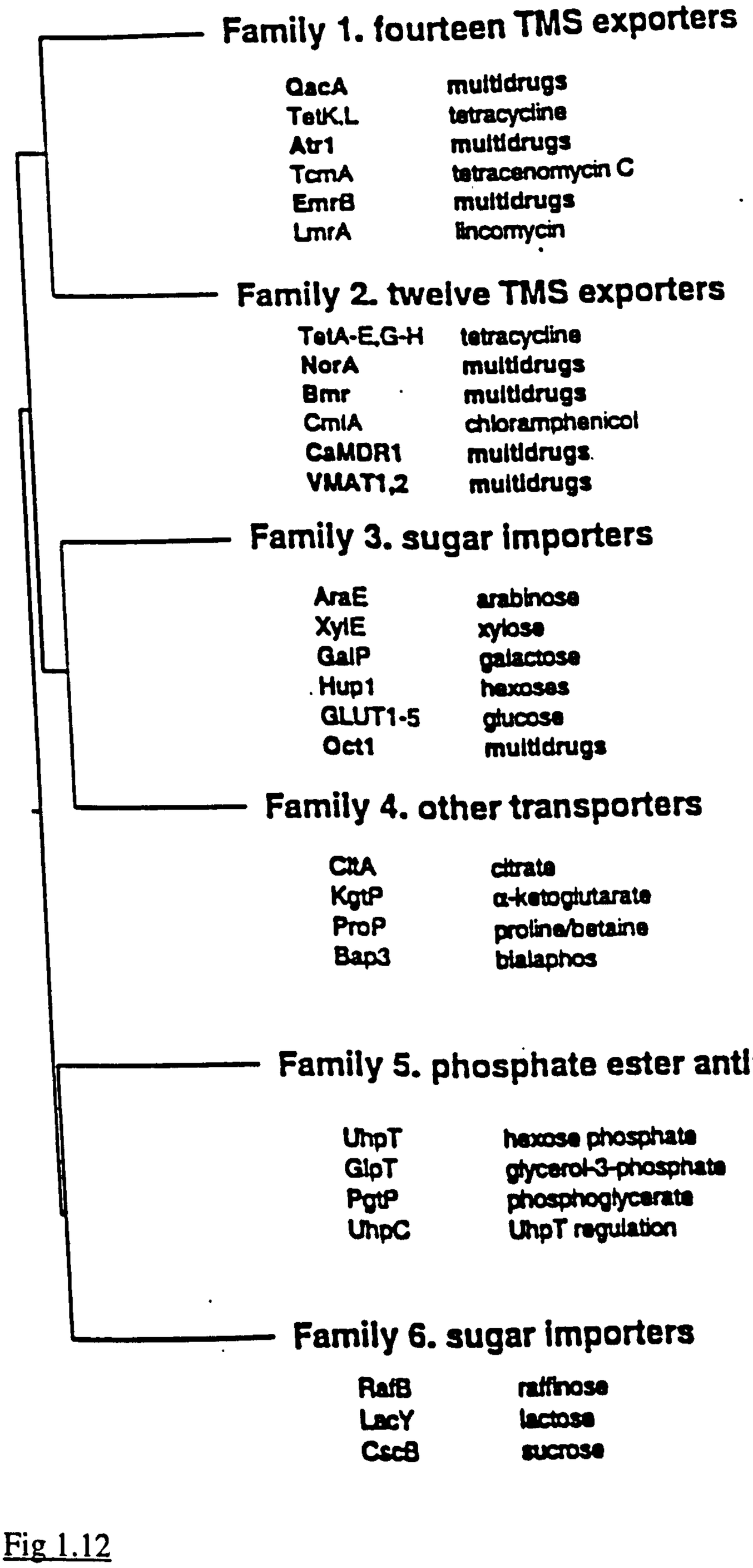



Fig 1.12

Revised phylogenetic relationship between six families of the MFS, taken from Paulsen *et al*, 1996).

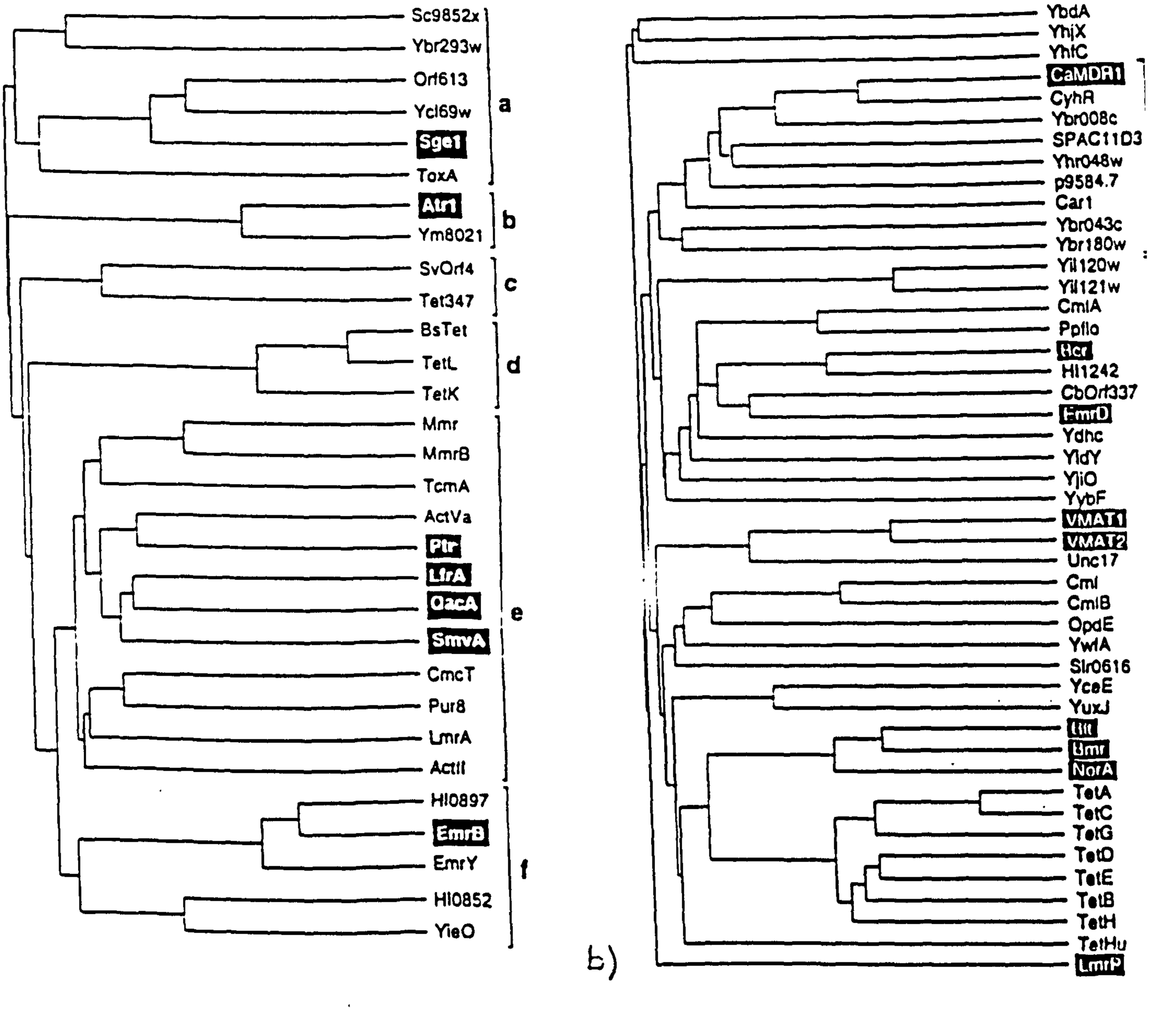



Fig 1.13

Phylogenetic trees showing relationships among (a) 14 TMS and (b)12 TMS families of the MFS. Multidrug transporters are shown in reverse type. The clusters a to f indicated in (a) are as described in the text. Taken from Paulsen *et al*, (1996). Tet347 in figure a) should now be known as OtrB.

1994). Family 2 is clustered in a similar way, it has five clusters; a) contains several yeast proteins including the multidrug efflux protein Sge1 (Amakasu *et al*, 1993). b) is a small cluster containing Atr-1, a yeast multidrug efflux protein (Kanazawa *et al*, 1988) c) contains two *Streptomyces* proteins, the oxytetracycline resistance protein, Tet347 (OtrB) from *Streptomyces rimosus* and the hypothetical product of the *Streptomyces violaceoruber SV*-ORF4 gene which probably encodes granaticin resistance (Hunter pers. comm), d) is the Gram-positive tet cluster, containing TetK, (Noguchi *et al*, 1986) TetL, (Hoshino *et al*, 1985) and a *Bacillus subtilis* TET resistance protein, e) is a large cluster of bacterial drug resistance proteins, mostly from Gram-positive organisms and includes several drug resistance/secretion proteins from antibiotic producing Actinomycetes as well as some multidrug exporters, also from Gram-positive bacteria. Lastly, cluster f) contains some Gram-negative derived 14 TMH proteins.

From the information presented above it can be seen that clusters usually appear to be divided on the basis of their function/substrate specificity, although both families do contain a numbers of proteins as yet uncharacterised. This method of division however does not account of the fact that multidrug resistance proteins/efflux pumps are presented in many of the various lineages. Lewis (1994) suggested that the multidrug proteins do not form a related family. The phylogenetic analyses of Paulsen *et al*, 1996 and Paulsen & Skurray, 1993, echo this sentiment. It is proposed that multidrug resistance is a phenomenon that has arisen separately on a number of occasions. Multidrug resistance could possibly arise from gradual loss of substrate specificity from a substrate specific transporter (Lewis, 1994).

# CHAPTER 2

## Materials and Methods

## 2.1 Introduction

This chapter contains the general procedures used in the experiments which were the basis of this thesis. The chapter is presented as five sections, bacterial strains, plasmids and chemicals; microbiological techniques and standard media; DNA methods ; Protein methods; and computer methods.

## 2.2 Bacterial strains, vectors and chemicals.

Table 2.1, Bacterial strains

| **Strain** | **Genotype** | **Source / Reference** |
|---|---|---|
| *E coli* Strain | | |
| BL21(DE3) | (E. coli B) $F^-$ompT$r^-_B$ $m^-_B$ (lambda DE3 lysogen) (lambda DE3 lambda 21; in the *int* gene of lambda is cloned a fragment bearing *lacI* and the T7 polymerase gene regulated by the lacUV5 promoter. | Novagen |
| DH5α | *endA1 hsdR17* ( $r^-_K$ $m^+_K$) *supE44 thi1 recA1 gyrA96 relA1* Δ*( LacZYA- ArgF)* U169 (ø80 d*lacZ* ΔM15) *deoR* | Hanahan (1983) |
| CB51 | *dam-3, ara-14, thi-1,* Δ (*lac-pro* ) | C. Boyd, University of Edinburgh |
| DS941 | *recF*143, *proA7, str*31, *thr1, leu*6, *tsx*33, *mt*12, *his*4, *arg*E3, *lac*Y+,*lac*ZΔM15, *lac*$I^q$, *gal*K2, *ara*14, *sup*E44, *xyl*5 | D. Sherrat, University of Glasgow |
| ET12567 | F *dam*-13::Tn9, *dcm*-6, *hsd*M, *hsd*R, *rec*F143, Tn10, *gal*K2, *gal*T22, *ara*14, *lac*Y1, *xyl*-5, *leu*B6, *thi*-1, *ton*A31, *rps*L136, *his*G4, *tsx*-78, *mtl*-1, *gln*44, *sup*E, *hsd*Δ5, *thi* Δ(*lac pro*A B) | MacNeil (1988) |
| ***Streptomyces rimosus*** strains | | |
| M15883 | $OTC^R$ | Binnie *et al.*, (1989) |

| | | |
|---|---|---|
| M15883$^S$ | OTC$^S$ | Butler *et al.*, (1989) |
| ***Streptomyces albus*** | OTC$^S$ | |
| *Streptomyces lividans* TK24 | *str6* | Hopwood *et al.*, (1985) |

Table 2.2, plasmids

| | | |
|---|---|---|
| pGLW12 | 2.3kb *Pst* I -*Sac* I *otrB* fragment cloned into pIJ487 (Ward et al, 1986) | L.Woodward, this laboratory |
| pGLW81 | 2kb *Sac*I-*Bam*HI *otrB* fragment cloned into pWHM3 | K. Linton, this laboratory |
| pIJ702 | *Streptomyces* cloning vector | Katz *et al* (1983) |
| pUC18 | -- | Yanisch-Perron *et al* (1985) |
| pIBI24 | derived from pEMBL | Dente *et al*, (1983) |
| pMTL23 | Amp$^R$ col E1 origin cloning vector used to clone PCR product | Chambers *et al* (1988) |
| pT7-7 | Amp$^R$ col E1 cloning and expression vector allowing positioning of the cloned gene in frame with T7 promoter | Tabor & Richardson (1985) |
| pET21b | Amp$^R$ *lacI* pMB1 origin. Cloning vector used to create C-terminal polyhistidine fusions regulated by both T7 promoter and down stream *lac* operator | Novagen |
| pET15b | Amp R lacI pMB1 origin. Cloning vector used to create N-terminal polyhistidine fusions | Novagen |
| pIJ6017 | Kan$^R$, *Streptomyces* expression vector regulated by Tsr inducible *tip*A promoter. | Takano *et al.* (1995) |
| pJJIBI | 460pb I fragment containing start of otrB ORF and small portion of upstream region cloned into pIBI24 | Chapter 3 |
| pJJ 1 | 2.1kb *otrB* fragment cloned *Eco*RI - *Hind* III into pUC18 | Chapter 4 |

| | | |
|---|---|---|
| pJJ 2 | 1.8kb *Nde* I - *Hind* III fragment cloned into pMTL23 | Chapter 4 |
| pJJ 3 | 1.6kb *BsaB* I - *Hind* III fragment from pJJ1 cloned into pJJ2 | Chapter 4 |
| pJJ 4 | 1.8 kb *Nde* I - *Hind* III fragment from pJJ3 cloned into pT7-7 | Chapter 4 |
| pJJ 5 | 1.8kb *Nde* I - *Hind* III fragment from pJJ4 cloned into pET21b | Chapter 4 |
| pJJ7 | 1.7kb *Nde*I-*Eco*RI PCR product cloned into pMTL23 | Chapter 5 |
| pJJ8 | 1.66Kb *Nde*I-*Sac*I fragment cloned into pET21b | Chapter 5 |
| pJJ9 | 1.7kb *Nde*I - *Eco*RI fragment cloned into pET21b | Chapter 5 |
| pJJ10 | 0.5kb PCR product cloned *NdeI- Eco*RI into pMTL23 | Chapter 6 |
| pJJ11 | 0.8 kb PCR product cloned *Eco*RI-*Nde*I into pMTL23 | Chapter 6 |
| pJJ12 | pJJ10 insert into pET21b | Chapter 6 |
| pJJ13 | pJJ11 insert into pET21b | Chapter 6 |
| pJJ14 | 0.8kb *Cla*I-*Eco*RI fragment from pJJ11 cloned into pJJ12 | Chapter 6 |
| pJJ17 | pJJ14 with *Sac*I deletion to make HisTag fusion at C terminus | Chapter 6 |
| pJJ18 | 1.8kb NdeI-*Bam*HI fragment from pJJ6 cloned into pET15b | Chapter 5 |

### 2.2.1 Chemicals and biochemicals

Ampicillin and ethidium bromide were supplied by Sigma Chemical Co. (Poole, UK).

Bactotryptone, yeast extract and Bactotryptone (agar) were supplied by Difco (Detroit, USA).

TEMED (N, N, N', N'- tetramethylethylene diamine) and Tris buffer were supplied by Boehringer Mannheim (Lewes, UK).

DMSO was supplied by BDH Chemicals (Poole, UK).

Agarose, IPTG (isopropyl-β-thiogalactoside) were supplied by BRL (Gibco Ltd., Paisley, UK).

Acrylamide, bisacrylamide and SDS were supplied by FSA laboratory supplies (Loughborough, UK).

Oligonucleotides were synthesised on either an Applied Biosystems Model 280A DNA synthesiser at the Institute of Biomedical and Life Sciences, University of Glasgow, using reagents from Cruachem (Science Park, Glasgow UK) or obtained from BRL (Gibco Ltd., Paisley, UK).

All other chemicals were of analytical reagent grade and were supplied from one of the following suppliers; BDH Ltd., Formachem Ltd., Sigma Chemical Co. Poole, UK).

### 2.2.2 Enzymes and kits

All restriction enzymes, T4 DNA ligase, alkaline phosphatase *Pfu* and *Taq* polymerases were supplied by BRL or Promega Corporation.

Purification of DNA from agarose gels was performed using Qiaex gel extraction kit from Qiagen Ltd. (Germany).

Plasmid DNA was prepared for sequencing using plasmid PURE mini-prep kit from Sigma -Aldrich. (Dorset, UK)

## 2.3 Standard media and microbiological techniques

### 2.3.1 Media for the growth of *E. coli*

Chemicals of good quality were used in the preparation of the growth media and solutions (AnalR grade when available). The source of many of the chemicals varied during the course of this work. The most common suppliers were BDH Chemicals Ltd.; Difco Laboratories (Detroit, Michigan, USA) and Sigma Chemical Co. Ltd.

a) L-broth

1% (w/v) tryptone, 0·5% (w/v) yeast extract, 0·5% (w/v) NaCl, 0·1% (w/v) glucose, 0·002% (w/v) thymine, made up in distilled water and adjusted to pH 7·0 with NaOH.

b) L-agar

As L-broth, but without glucose and with the addition of 1·5% (w/v) bacto-agar.

c) 2xYT medium

1·6% (w/v) bacto-tryptone, 1% (w/v) bacto-yeast extract, 0·5% (w/v) NaCl, made up in distilled water and adjusted to pH 7·0 with NaOH.

d) Terrific Broth.

1.2% (w/v) bacto tryptone, 2.4% (w/v) bacto yeast extract, 0.4% (v/v) glycerol, dissolved in 900ml. After autoclaving add 100 ml of a solution of 0.17mM $KH_2PO_4$ and 0.72mM $K_2HPO_4$.

e) Minimal agar

Agar was made to 1·75% with water. To 75 ml of this, 25 ml of D+M salts were added and glucose and thiamine were added to final concentrations of 2 mg/ml and 20 µg/ml, respectively.

f) SOC

2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl per litre.(pH 7 ) After autoclaving, add $MgCl_2$ to 10 mM and glucose to 20 mM

g) Davis and Mingoli (D+M) Salts (x4)

2·8% (w/v) $K_2HPO_4$, 0·8% (w/v) $KH_2PO_4$, 0·4% (w/v) $(NH_4)_2SO_4$, 0·1% (w/v) trisodium citrate, $MgSO_4.7H_2O$, made up in distilled water.

### 2.3.2 Media used for growth of *Streptomyces*

a) Soya Mannitol Agar (SM)

This was used as a general plating medium for *Streptomyces*, particularly for production of spores. It consists of 2% (w/v) mannitol, 2% (w/v) soya bean flour, and 1·6% (w/v) agar, made up in tap water.

b) Yeast extract-Malt extract (YEME)

Used for the growth of *Streptomyces lividans* It consists of 0·3% (w/v) Difco yeast extract, 0·5% (w/v) Difco bacto peptone, 0·3% (w/v) Oxoid malt extract, 1% (w/v) glucose, 34% (w/v) sucrose, made up in distilled water. To prevent pelleting, and produce a well-dispersed growth, $MgCl_2$ was added to a final concentration of 5mM, and glycine to a final concentration of 0·5% (w/v).

c) Tryptone Soya Broth. (TSB)

30g/l in distilled water

Used for the growth of S. *albus and* S. *rimosus.*

For *S. albus* add glycine, (10%) 5 ml/l and $MgCl_2$, (100 mM) 5 ml/l

d) TSB agar.

1.5% agar in TSB broth

### 2.3.3.Sterilisation of media

All growth media and buffer solutions were sterilised by heating to 120°C for 15 minutes in an autoclave. Antibiotics solutions and amino acids were sterilised by filtration through Nalgene 0.22µm pore membranes.

### 2.3.4 Antibiotics and indicators

The antibiotics used throughout for liquid and plate selection were as follows:

| Drug | Final concentration in µg/ml | Stock concentration in mg/ml | Storage temp. in °C |
|---|---|---|---|
| Ampicillin | 100 | 200 | -20 |
| Chloramphenicol | 34 | 34 | -20 |
| Thiostrepton | 50 | 10 | -20 |

Kanamycin was used to select for *Streptomyces lividans* transformed with pIJ6017 based plasmids, 100 µg/ml was used for selection on solid media and 5 µg/ml used in liquid. When these plasmids were used in *S. albus*, a concentration of 200 µg/ml was used in solid media and 100 µg/ml in liquid.

### 2.3.5 Growth of *Streptomyces*

#### 2.3.5.1 Growth in liquid media

Cultures were typically grown in 500 ml conical flasks containing 100 ml of medium at 30°C on an orbital shaker at 200 rpm. Expression experiments employed TSB or YEME medium which produced reproducible and rapid growth. Spores from frozen suspensions or suspensions that were freshly-prepared from a frozen slope were used to inoculate the medium. Cells grown on YEME could be harvested after 48-72 hours and stored indefinitely as a cell pellet at -20 °C.

#### 2.3.5.2 Harvesting of mycelia

After growth, mycelia were diluted in an equal volume of distilled $dH_2O$, recovered from the media by centrifugation (10 minutes, 10000 g) and resuspended in $dH_20$.

### 2.3.6 Growth of *E.coli*

Liquid cultures of *E. coli* strains from which plasmids were to be isolated were grown in L- broth with the appropriate antibiotic selection (typically ampicillin at 100 µg/ml and/or chloramphenicol at 34 µg/ml). The volume of broth inoculated depended on the quantity of plasmid required. Routinely 1.5 ml and 3 ml culture

were used for Doly and Sigma plasmid preparations respectively. (sections 2.4.2.1 and 2.4.2.2).

For the preparation of membrane fractions and the preparation of competent cells, *E coli* were grown in Terrific-broth. To maximise aeration of the culture, the volume of the Erlenmeyer flask used was at least five times that of the broth. All cultures were incubated at 37°C in an orbital shaker at ca. 250 rpm.

#### 2.3.6.1 Expression of cloned genes in *E. coli* using IPTG.

Liquid cultures of *E. coli* strains in which cloned genes were to be expressed from an expression vector containing the T7 promoter were inoculated from overnight cultures and grown in 800 ml Terrific-broth with the appropriate antibiotic selection. Cultures were monitored for growth by spectrophotometry of 1 ml samples at 600 nm. On reaching $OD_{600}$ = 0.6, cultures were induced to express protein by the addition of 0.4 mM IPTG and grown for a further 2.5 hours. Cultures were then harvested by centrifugation in Beckman J2-21 M/E centrifuge using a JA14 rotor at 5000 x g for 5 minutes.

### 2.3.7 Preparation of *Streptomyces* spores

#### 2.3.7.1 Preparation and storage of suspensions

Concentrated spore suspensions were required for inoculating liquid cultures of *S. rimosus, S. albus* or *S. lividans.* The protocol described by Hopwood *et al.* (1985) was followed with minor modifications:

A boiling tube containing a slant of SM agar (produced by pouring ca. 20 ml of molten agar into the tube and allowing it to solidify with the tube held at +5° from the horizontal) was inoculated with 150 µl of a spore or mycelial fragment suspension and incubated at 30°C. After 5-10 days the surface of the culture was covered in a dark grey mass of spores. The slant was then sealed using parafilm and frozen at -20°C. Spores could be harvested immediately or stored indefinitely at -20°C. The spores were harvested by adding 5 ml of $dH_2O$ to the frozen slant and rubbing the surface of the slant with a 10 ml glass syringe. Contaminating agar or mycelial fragments were removed by a single passage through a cotton wool filter, as

described in Hopwood *et al.* (1985). The filtered spore suspension was then either used fresh to inoculate YEME or TSB media or frozen at -20°C.

### 2.3.7.2 Spore counts

Colony forming units were determined by plating suitably diluted spore samples on Soya plates. Counts of the number of colonies were made after incubation at 30 °C for 5 days.

## 2.3.8 Preservation of *E.coli* strains

*E.coli* strains were stored in glycerol. An 800 µl aliquot of an overnight culture was mixed with an equal volume of 40% (v/v) glycerol, 2% peptone (w/v) and frozen at -70$^0$C The strains were revived by scraping the surface of the frozen suspension with a toothpick and either inoculating liquid broth or streaking onto agar to isolate a single colony.

## 2.3.9 Introduction of plasmid DNA into *E. coli*

### 2.3.9.1 Preparation of competent cells

a) $CaCl_2$ method.

Cells prepared by this method have an efficiency of ca. $1x10^5$ cfu per µg of plasmid DNA so were used for moving plasmids from strain to strain.

An overnight culture of the recipient strain was inoculated 1: 100 into fresh L-broth with antibiotic if appropriate and incubated with shaking at 37°C for 90 - 120 minutes to a density of approximately $10^8$ $ml^{-1}$ cells ($OD_{600}$ 0.4 - 0.6) The cells were harvested by pelleting in a bench top centrifuge @ 10,000 x g, for 5 minutes and resuspended in half the original culture volume of ice cold 50 mM $CaCl_2$. The cells were pelleted again and resuspended in 1 ml ice-cold 50 mM $CaCl_2$ for every 30 ml of original culture volume and stored on ice for at least 20 minutes before use. The competent cells prepared by this method were used the same day.

b) Hanahan's method

Cells prepared by this method have an efficiency of ca. $1x10^{6-}$ $1x10^7$ cfu per µg of plasmid DNA so were used for transformation of ligations.

Cells were prepared by a modified method of Hanahan, (1983).

An overnight culture of the recipient strain was inoculated 1:100 into 2xYT with antibiotic if appropriate. The culture was incubated at 37°C with shaking (as above) until $OD_{600}$ reached 0.4 - 0.6. Cells were pelleted in a bench top centrifuge and resuspended in 2.5 ml ice cold TFB buffer for every 30 ml of culture. The cells were left on ice for 15 minutes. 100 µl of DMSO was added after 5 minutes, after a further 10 minutes 100 µl of DTT/KAc** was added. After another 5 minutes a further 100µl of DMSO was added. After a further 5 minutes, the cells were divided into 200 µl aliquots and were ready to transform. Competent cells prepared by this method were used the same day.

TFB buffer : 10 mM MES* pH6.3, 45 mM $MnCl_2.4H_2O$, 10 mM $CaCl_2.2H_2O$, 100 mM KCl, 3 mM cobalt chloride.

*A 1 M stock of MES was prepared by dissolving 19.52 g of MES in 80 ml of distilled $H_2O$, pH was adjusted to 6.3 with 5 M KOH, add $dH_20$ to 100 ml and filter sterilised.

** DTT/KAc: 2.25 M DTT, 40 mM Potassium acetate, pH6.

<u>b) Electrocompetence method</u>

Cells prepared by this method have an efficiency of ca. $1x10^{-9}$ cfu per µg of plasmid DNA so were used for transformation of ligations in lower efficiency strains

An overnight culture of the recipient strain was inoculated 1:100 into 500 ml fresh L broth and incubated for 3 - 3.5 hours until OD 600 reached 0.4 - 0.6. The cells were then harvested by pelleting in a Beckman centrifuge @ 15000 x g , 4°C for 15 mins and resuspended in 250 ml ice cold, sterile $dH_2O$, the cells were pelleted again and resuspended in 50 ml of ice cold, sterile, 10% glycerol and pelleted again. The cells were finally resusupended in 0.8 ml ice-cold sterile 10% glycerol, and 40µl aliquots dispensed into 0.5 ml microcentrifuge tubes on dry ice before storage at -80°C. These cells can be stored for months once prepared.

## 2.3.9.2 Transformation procedures

<u>a) CaCl2 transformation</u>

Transformations were carried out in sterile 1.5 ml microcentrifuge tubes. An aliquot (maximum volume 10 µl) of ligation mix or plasmid DNA was added to 100 µl

aliquot of competent cells, mixed gently and incubated on ice for at least 20 minutes. The DNA/ cell mixture was heat shocked at 42 °C for 90s. 1 ml of L broth was added and the transformation mix incubated at 37 °C with shaking, for 45 minutes to allow any antibiotic resistance to express before being plated onto L- agar containing the appropriate antibiotic / chromogenic substances and incubated at 37 °C overnight.

b) Hanahan's method

As above but using 0.8 ml SOC in place of L-broth before incubation.

c) Electrotransformation

An aliquot of competent cells was removed from -80°C and thawed on ice, to this was added no more than 2.0 µl ligation mix or plasmid DNA the mixture was incubated on ice for 1 minute. The mixture was transferred to a pre chilled electrotransformation cuvette, the cuvette was dried and placed in dried, prechilled cuvette holder and subjected to a pulse of 250kV (resistance 200Ω) using a Biorad pulse controller. 1 ml of ice cold SOC liquid media was added to the cuvette and mixed using a Pasteur pipette. The mix was incubated at 37 °C for 45 minutes to allow expression of antibiotic resistance before plating onto L-agar as the previous method.

### 2.3.9.3 Selection of pUC derived recombinant clones

a) Ampicillin

Stock solutions (200 mg/ml made up in water) were added to molten agar (cooled to 55°C) to a final concentration of 50 µg/ml.

b) Xgal ( 5-bromo-4-chloro-3-indolyl-β-D-galactosidase)

This was used in conjunction with IPTG (indolyl-β-D galactoside) to identify the *E. coli* clones containing pUC-based vectors with inserts in their polylinker cloning sites. Recombinants containing inserts are generally white whilst those with no insert are blue. X-gal was stored at a concentration of 20 mg/ml in dimethylformamide (DMF) at-20°C. IPTG stock was at 24 mg/ml in $dH_2O$ and stored at -20°C. Final concentrations in agar were 20 µg/ml and 50 µg/ml respectively.

### 2.3.10 Introduction of plasmid DNA into *Streptomyces*

Plasmids were introduced into protoplasts of *Streptomyces* species using the polyethylene glycol method described by Hunter (1985).

### 2.3.10.1 Protoplast method of Transformation

**Reagents.**

Medium P: 5·73 g N-tris[hydroxymethyl]methyl-2-aminoethanesulfonic acid (TES), 103 g sucrose, 2·93 g $MgCl_2.7H_2O$, 0·5 g $K_2SO_4$, 3·68 g $CaCl_2.2H_2O$, 2 ml trace element solution; adjusted to pH 7·4 with NaOH and made up to 1 litre in distilled water.

Lysozyme solution: 10% (w/v) sucrose, 25 mM TES buffer (pH 7·2), 2·5 mM $K_2SO_4$, 2 ml trace elements (Hopwood *et al.*, 1985), 2·5 mM $MgCl_2$, 2·5 mM $CaCl_2$. $KH_2PO_4$ (0·005% [w/v]) and lysozyme (0·3 mg/ml) were added immediately prior to use.

PEG solution: 1 g of polyethylene glycol 1540 (supplied by BDH) was melted in a microwave (600 W) for 10 seconds on the reheat setting and then mixed with 3 ml of medium P.

### 2.3.10.2 Preparation of protoplasts

30 ml cultures were grown in the appropriate medium at 30°C. The optimal time to harvest the mycelium in order to recover the most competent protoplasts was species-dependent (i.e. *S. lividans* mycelia were harvested after 65 hours and *S. rimosus* and *S. albus* after 48 hours, respectively). The mycelium was pelleted at 12000 g for 10 minutes and washed twice in 10·3% (w/v) sucrose. The pellet was then resuspended in 4 ml of lysozyme solution and incubated at 37°C for 15-30 min. The formation of protoplasts was monitored using a microscope, and the reaction terminated by adding 5 ml of P medium.

The protoplasts were then filtered through cotton wool (Hopwood *et al.*, 1985), pelleted using a centrifuge (12000 g for 10 minutes) and washed twice in P medium.

Finally, they were resuspended in 2 ml of medium P, dispensed into 100 µl aliquots and frozen at -70°C.

### 2.3.10.3. Transformation of protoplasts.

The protoplasts were thawed on ice. DNA was added in a volume of less than 10 µl and the mixture incubated on ice for 30 seconds. 400 µl PEG solution was added, the solution incubated on ice for a further 1 minute and, finally, 800 µl medium P was added. Dilutions of the transformation mix were then made in medium P and plated onto regeneration medium.

### 2.3.10.4 Regeneration of transformed protoplasts

Selection of transformants, due to plasmid-borne resistance markers, was performed by overlaying the regeneration plates with 1 ml of a 10·3% (w/v) sucrose solution containing a suitable antibiotic after 16-20 hours of non-selective growth at 30°C. The antibiotics used in this work for plasmid selection in *Streptomyces* were thiostrepton (obtained from E. R. Squibb, New Jersey, USA) and kanamycin. Thiostrepton was dissolved in DMSO to make a 1% (w/v) stock solution. Kanamycin was dissolved in distilled $H_2O$ at 25 mg/ml Transformed protoplasts were selected by overlaying with 220µg/ml thiostrepton or 2 mg/ml kanamycin for *S. lividans*, 5 mg/ml for *S. albus*

## 2.4 General DNA Methods

### 2.4.1 Commonly used buffers

a) TE buffer (10x)

100 mM Tris-HCl pH8, 10 mM EDTA, Sterilised at 120 °C and 20 psi using an autoclave and stored at room temperature, used as a 1x solution

b)TAE buffer (10x)

4.48% (w/v) Tris, 1.64% (w/v) sodium acetate, 0·36% (w/v) $Na_2EDTA.2H_2O$, made up in distilled water, pH adjusted to 8·2 with glacial acetic acid, used as a 1x solution.

c) λ/*Hin*d III DNA markers

λ *c*I857 *S*7 DNA was obtained from BRL. This DNA was cleaved with the restriction enzyme *Hin*d III and resulting DNA fragments diluted to a final concentration of 27 ng/µl in TE (final concentration 1x) with loading buffer added to 1x concentration. Typically, 10 µl was used on agarose gels as markers for comparing the size and concentration of bands in samples.

d) Agarose gel loading buffer

0.5 % (w/v) bromophenol blue, 0.05 % (w/v) xylene cyanol, 50 % Ficoll, 1 % (w/v) SDS, 100 mM EDTA.

## 2.4.2 Preparation of plasmid DNA

### 2.4.2.1 Reagents for isolation of plasmid DNA

Protocols based on the alkaline lysis method (Birnboim and Doly, 1979) were used for the isolation of plasmid from small (5 ml) cultures of *E. coli*. An alternative alkaline lysis developed by Kieser, was used for isolation of plasmid DNA from *Streptomyces*.

a) Birnboim Doly I (BDI)

50 mM glucose, 25 mM Tris-HCl pH 8·0, 10 mM EDTA. If streptomycete plasmids were being isolated, lysozyme was added immediately before use to a final concentration of 5 mg/ml.

b) Birnboim Doly II (BDII)

0·2 M NaOH, 1% (w/v) SDS which was stored in a plastic container.

c) Birnboim Doly III (BDIII)

Prepared by mixing 60 ml 5 M $CH_3COOK$ with 11.5 ml glacial acetic acid and 28.5 ml $dH_2O$. The resultant solution is 3 M with respect to potassium and 5 M with respect to acetate.

d) DNase-free RNase

Pancreatic RNase (RNase A) was dissolved at a concentration of 10 mg/ml in $dH_2O$, heated to 100°C for 15 mins and allowed to cool slowly to room temperature. The RNase was then aliquoted and stored at -20°C.

e) Acid phenol/chloroform

50 g phenol was dissolved in 50 ml chloroform and 10 ml $dH_2O$. Hydroxyquinoline was added as an anti-oxidant

f) Alkaline SDS

M NaOH, 2% (w/v) SDS.

### 2.4.2.2 Small-scale plasmid preparation from *E.coli*

Routinely, plasmids were isolated from 1·5 ml of *E. coli* cultures. The cells were pelleted by centrifugation in a 1·5 ml microfuge tube (12000 g for 30 secs) and resuspended in 100 µl of BDI, containing lysozyme at a concentration of 1 mg/ml, using a vortex mixer. This was followed by the addition of 200 µl of BDII and repeated inversion of the microfuge tube to thoroughly mix the suspension. Immediately afterwards, 150 µl of pre-chilled BDIII was added to the viscous bacterial lysate, mixed gently on the vortex mixer and placed on ice for 5-10 mins. The cell debris and most of the chromosomal material were harvested by centrifugation (12000 g, 4°C for 10 min) in a microfuge. The supernatant was transferred to a fresh tube and extracted with half volumes of phenol/chloroform and chloroform. The nucleic acid was then precipitated by the addition of 2 volumes of ethanol and allowed to stand at room temperature for 5 mins. The precipitate was harvested by centrifugation in a microfuge (12000 g, 4°C for at least 15 mins). The resulting pellet was rinsed twice with 70% (v/v) ethanol before it was allowed to dry by leaving the tube open on the bench. The nucleic acid was then resuspended in 50 µl $dH_2O$ containing DNase-free RNase (20 µg/ml).

If further purification of the DNA was required (for example for double-stranded sequencing reactions), the pellet was resuspended in 16 µl of 1x TE, to which 4 µl of 4 M NaCl and 20 µl of 13% PEG were added, followed by incubation on ice for 20 mins. Centrifugation for 15 mins at 4 $^0$C produced a pellet, which was washed in 70% ethanol, and resuspended in 1x TE.

### 2.4.2.3. Preparation of plasmid DNA by the "Kieser" method

(Hopwood *et al.*, 1985)

Up to 1 ml of *Streptomyces* culture was diluted 1:1 with $dH_2O$ and pelleted by centrifugation. The pellet was washed in TES buffer, resuspended in 500 µl TES buffer (containing 2 mg/ml lysozyme), and incubated at 37°C for up to 30 mins. 250 µl of alkaline SDS was added, the solution homogenised immediately, and incubated at 70°C for 15 mins. After cooling to room temperature, the solution was mixed with 80 µl of acid phenol/chloroform, centrifuged, and the supernatant isolated. This step was repeated using 80 µl of chloroform, and the supernatant added to 60 µl 3 M sodium acetate and 600 µl isopropanol. The solution was mixed and left at room temperature for 10 mins, followed by centrifugation. The pellet was washed in 70% ethanol, and resuspended in 1x TE.

## 2.4.3 Precipitation of DNA

DNA solutions were precipitated by the addition of 1/10 volume of 5 M NaCl and 2 volumes of cold ethanol or an equal volume of isopropanol. After mixing, the DNA was pelleted by centrifugation (27000 g, 4 °C, 30 mins for volumes of 7·5-20 ml, or 12000 g, 4 °C, 15 mins for small volumes in microfuge tubes). The pellet was washed in 70% (v/v) ethanol and dried briefly in a vacuum desiccator, or in an open tube on the bench.

## 2.4.4 Spectrophotometric measurement of nucleic acid

Nucleic acid concentrations were determined spectrophotometrically at 260 nm. In a 1 cm path length an absorbance value of 1·0 corresponds to 50 µg/ml for double-stranded DNA, 33 µg/ml for single-stranded DNA and 20 µg/ml for oligonucleotides.

## 2.4.5 Digestion of DNA with restriction enzymes

Restriction digests were carried out using the BRL restriction enzymes and REact buffers which were provided with each batch of enzyme. There are ten different REact buffers with a range of salt concentrations, each one suitable for a range of enzymes. Alternatively, digests were performed using New England Biolabs restriction enzymes and the accompanying restriction buffers. Analytical digests were

carried out in a volume of 20 µl at 37°C. Preparative digests were carried out in larger volumes. When DNA was digested with two restriction enzymes, the endonuclease requiring the lower salt buffer was used first. After the recommended duration of digestion, the salt concentration was adjusted and the second enzyme added.

### 2.4.6 Ligation of DNA fragments

The ligation of DNA fragments was carried out usually at a DNA concentration of 6 µg/ml. The molar ratio of insert fragment to vector was 3:1, when the vector could not ligate to itself (for example when using a vector that has been dephosphorylated or has been cut with two non-complementary enzymes). A molar ratio of 10:1 was used when the ends of the vector could ligate to each other. Ligations were performed usually in 20 µl of 1x ligation buffer provided by BRL, containing 1 U of T4 ligase per µg of DNA. The reactions were incubated for 4 hours at room temperature or overnight at 4°C

### 2.4.7 Removal of the 5' phosphate from linearised DNA

<u>10x CIP Buffer:</u> 200 mM Tris-HCl pH 8·0, 10 mM $MgCl_2$, 10 mM $ZnCl_2$ and 0·5 mg/ml Bovine Serum Albumin.

<u>Procedure:</u> Calf Intestinal Alkaline phosphatase (CIP) was used to remove the 5' phosphate from DNA. Around 5 pmoles of 5'-terminal phosphorylated DNA with 5' protruding ends (approximately 7 µg of a 5 kb molecule) were incubated in 1x CIP buffer, containing 0·1 U of CIP at 37 °C for 60 mins. The reaction was terminated by heating to 65 °C in 1x gel loading buffer for 10 mins. The 5'-terminal dephosphorylated DNA was recovered from an agarose gel after electrophoresis. (2.4.10)

### 2.4.8 Filling in of 5' overhanging bases resulting from restriction digests.

After digestion, 8µl of dNTP's at a 1.25 mM concentration and 2-3U of T4 DNA polymerase was added to the 20µl reaction and incubated for a further 1 hour at 37 °C. The dNTP's were removed by electrophoresis and gel purification before further manipulation.

### 2.4.9 Addition of linkers to blunt ended fragments

The molar ratio of linker to blunt ended vector must be 200:1 4 µl of 5x ligation buffer and 200 ng of vector DNA (0.6 pmol) were added to 22.5 pmol of kinased linker and 1U T4 DNA ligase in a 20 µl reaction volume and incubated at 4 °C overnight.

### 2.4.10 Removal of excess linker after ligation.

The reaction was heated to 65 °C for 10 minutes to stop the reaction. The DNA was precipitated (2.4.4) and reconstituted in 8 µl 1x TE to which was added 1µl of suitable ReAct buffer and 1U of restriction enzyme. The restriction digest was carried out as normal (2.4.6.) The excess DNA was removed by electrophoresis and gel purification, whereupon the fragment was religated to itself and transformed before further manipulation.

### 2.4.11 Agarose gel electrophoresis

#### 2.4.11.1 Preparation of agarose gels

DNA was visualised on horizontal neutral agarose gels. Although 0·9 % (w/v) gels were most commonly used, 1-2 % (w/v) gels were occasionally used to separate fragments of <1·5 kb. Gels were routinely prepared and run in TAE buffer. λ-*Hind III* markers were used on all gels as size markers and for quantification of the amount of DNA by comparing the intensity of bands to those of the samples (2.4.1c).

a) Mini gels

BRL model H6 gel kits were used for the rapid analysis of DNA after digestion with restriction enzymes or precipitation steps. 0·18 g agarose was added to 20 ml of 1x TAE, boiled, and then cooled to 60°C. The molten agarose was poured into a 7·6 cm x 5·1 cm gel caster with an 8 well slot former (4·1 x 0·8 mm wells). Alternatively, smaller gels were made by pipetting 8-12 ml of molten agarose onto a 7.6 cm x 5.1 cm glass slide, wells were formed with a modified 8 well slot former. After the gel had set, the slot former was removed and the gel placed in the tank with 500 ml of 1x

TAE. Depending on the time available and the level of resolution required, the DNA was separated by electrophoresis for 30-60 mins with an applied voltage of 2-10 V/cm. 1/10 volume loading buffer (2.4.1d) was added to the DNA samples before they were loaded onto the gel. The separated DNA molecules were visualised on a 302 nm UV transilluminator

b) Large gels

200 ml gels were also used to ensure good separation of DNA fragments for accurate sizing .They were made by pouring 200 ml of molten agar into a 16·5 x 23 cm gel former with a 14 space slot former. The gels were run overnight at 20 V in 1x TAE buffer. DNA samples were mixed with 1/10 volume of 10x loading buffer (2.4.1d) before loading onto the gel.

### 2.4.11.2 Staining of gels with Ethidium Bromide

Gels were stained by placing in a solution of EtBr dissolved in 1x TAE or distilled $H_2O$ at 0.5 µg/ml for 10-20 minutes. Background staining caused by unbound EtBr was removed by soaking in 1 x TAE.

### 2.4.11.3 Photography of agarose gels

Gels stained with ethidium bromide were viewed on a 302 nm UV transilluminator and photographed using Polaroid type 67 land film Alternatively, pictures were obtained using a Mitsubishi video copy processor attached to a UVP video camera.

## 2.4.12 Recovery of DNA from agarose gels

Using Qiaex-suspension

A Qiaex gel extraction kit was obtained from Qiagen, Germany. The gel slice containing the excised DNA fragment was mixed with 3 vols. QX1 solubilisation buffer and Qiaex-suspension (10 µl per 5 µg DNA). The gel slice was incubated at 50°C for 10 mins with frequent mixing to solubilise the gel and allow the DNA to absorb to the Qiaex-suspension. The suspension was pelleted, washed twice with 500 µl QX1 wash buffer, and twice with 500 µl PE wash buffer. The suspension was pelleted and air-dried for 10-15 mins, followed by resuspension in 1x TE buffer or water, 10 mins incubation at 50°C with frequent mixing and pelleted again. The

DNA was contained in the supernatant, in a condition ready for any subsequent manipulations.

### 2.4.13 DNA sequencing

Plasmid DNA was sequenced automatically by Mr. R. Tate at the Molecular Biology Laboratory, Department of Physiology and Pharmacology, University of Strathclyde using an, ABI (Applied Biosystems) 373A automated sequencer run using dye terminators and Amplitaq FS DNA polymerase (Cycle sequencing).

#### 2.4.13.1 Preparation of plasmid template for sequencing

Plasmid DNA was prepared for sequencing using a plasmidPURE kit (Sigma).

#### 2.4.13.2 Preparation of sequencing primers.

Oligonucleotide primers for sequencing were obtained from BRL (Paisley,UK) as a dry pellet, primer was reconstituted with 200 µl 1 x TE this was used as a stock and stored at -20°C. The DNA concentration of the stock was determined spectrophotometrically, (2.4.4). Stock was diluted with sterile distilled $H_2O$ to a suitable concentration for the sequencing reaction.

### 2.4.14 Polymerase Chain Reaction.

The Polymerase Chain Reaction (PCR) was employed several times in this work to introduce mutations into DNA. Conditions must be optimised for each primer/template combination but a general protocol is given below.

<u>a) PCR mix ingredients</u>.

5-10 ng plasmid template

100 pmole primer I

100 pmole primer II

5 µl DMSO

2.5 mM $MgCl_2$

1.25 mM dNTP's

5 µl 10x buffer

total volume to 50 µl with $dH_2O$

b) Reaction conditions.

step 1. Denature DNA at 94°C for 4 minutes add 0.5 Units *Taq* or *Pfu* polymerase (1µl) plus 1 drop of mineral oil to prevent evaporation from top of tube.

step 2. 94 °C for 30s

step 3. $T_m$* - 5 °C for 60s

step 4. 72 °C for 120s

Repeat steps 2- 4 for 30 cycles.

*$T_m$ is calculated by the equation:

$T_m = 59.9 + 41$ [% G+C] - 657/ $L_p$

(where $L_p$ is length of primer in bases)

### 2.4.14.1 PCR Primers

Several primers were employed during work presented in this thesis.

5'CCCAGTCACGACGTTGTAAAACG3'

Universal (forward) M13/pUC Primer

5'AGCGGATAACAATTTCACACAGG3'

Universal (reverse) M13/pUC Primer

5'GATCCTGACGCATATGTCGGGACTGCT3'

Custom primer 1, includes *NdeI* site, underlined (4.4.2)

5'CGGAATTCCTCGAGTCAGAGCTCCCAGGCGTCCGACGCGGG3'

Custom primer 2, changes otrB stop codon to TGG, encoding tryptophan (5.2.2)

5'CCCGGAATTCTAGGTATCGATGCCGACCATGCGGTGCTCGCGCCG3'

Custom primer 3, creates N-terminal half of truncated *otrB* when used with universal forward primer (5.8)

5'CATATGGGCATCGATACCCTGCGGCGCGGCGAC3'

Custom primer 4, creates C-terminal half of truncated *otrB* when used with reverse primer (5.8)

## 2.5 General protein methods

### 2.5.1 Preparation of crude protein extracts

Cell pellets were resuspended in ice cold extraction buffer, (100 mM potassium phosphate buffer pH7.0 [6.15 ml 1 M $K_2PO_4$, 3.85 ml 1 M $KH_2PO_4$ in 100 ml], 5 mM EDTA, 1 mM benzamidine, 0.4 mM DTT; benzamidine and DTT added immediately before lysis) and lysed by passage through an automatic French Pressure cell under 750 psi pressure or by use of a constant systems model 40/AA cell disrupter. The suspension was centrifuged at 10,000 x g for 15 minutes at 4°C and the supernatant isolated.

### 2.5.2. Preparation of membrane fractions

#### 2.5.2.1 Disruption of cells

Cell pellets were resuspended in not more than 30 ml 20 mM Tris, pH8, 0.5 mM EDTA and French pressed as above. After French pressing the cells were centrifuged to bring down whole cells and cell debris using a Beckman JA20 rotor at 13750 rpm, (1500 x g) for 30 minutes. The supernatant was decanted off and ultra-centrifuged in a Beckman Ti60 rotor at 45,000 rpm, (131,000 x g) for 90 minutes to pellet the mixed inner and outer membranes. The membrane pellets were each resuspended in 4 ml 50 mM sodium phosphate buffer, pH8 and quick frozen on dry ice and ethanol. These membrane samples can be stored at -20°C.

#### 2.5.2.2 Solubilisation of membranes

Membrane pellets were thawed on ice and resuspended in a total of 8 ml solubilisation buffer, 50 mM sodium phosphate, pH8, 10%(v/v) glycerol, 20 mM imidazole pH8, 1% (w/v) dodecylmaltoside (DDM) using a blue tip at first then sequentially finer needles. The suspension was mixed on a rocker for 2 hours at 4°C.

After mixing, the membrane samples were centrifuged in a Beckman SW 50 rotor at 290,000 x g for 90 minutes.

### 2.5.3 Isolation of His-tagged protein

#### 2.5.3.1 Preparation of HisResin

4 ml of HisResin (Novagen) equivalent to 2 ml settled bed volume, was washed twice with distilled $H_2O$ then charged with 50 ml 400 mM $NiSO_4$ by rocking for 8 hours at 4°C.

After this time the tubes were left upright for 30 minutes for the agarose resin to settle, the supernatant was then removed. The resin was then washed twice in denaturing equilibration buffer, 50 mM sodium phosphate, pH8, 10%(v/v) glycerol, 20 mM imidazole, pH8, 0.1% DDM, 6 M urea.

#### 2.5.3.2 Binding of HisTagged proteins to resin

Solubilised membrane fractions were added to the settled resin and the volume made up to 50 ml with denaturing equilibration buffer in falcon tubes. The tubes were rocked overnight at 4 °C in order that the HisTagged protein would bind to the Ni-resin.

#### 2.5.3.3 Column purification of HisTagged proteins.

The resin from the overnight binding reaction was left to settle in the tubes and then loaded into a column made from 2.5 ml syringe with a glass wool frit. The resin was left to settle in the column which was washed through with the remaining volume from the 50 ml of equilibration buffer.

The column was then washed with 10 volumes of elution buffer I, 50 mM sodium phosphate buffer, pH8, 20 mM imidazole, pH8, 10% (v/v) glycerol, 0.1% (w/v)DDM, 6 M urea. 1.5 ml fractions were collected in eppendorf tubes.

The column was then washed through with 10 volumes of elution buffer II, as above but with 200 mM imidazole in place of 20 mM. 1.5 ml fractions were again collected.

The final wash was done using citrate buffer, 100 mM sodium citrate, pH4, 10% (v/v) glycerol, 0.1% (w/v) DDM, 6 M urea. 1.5 ml fractions were collected in

eppendorfs as before. The column purification should be carried out at a low temperature but 4 °C may cause the urea to precipitate.

### 2.5.4 Protein Assays

The amount of protein in membrane samples was measured prior to gel electrophoresis using the bicinchoninic acid assay for soluble protein.

#### 2.5.4.1 BCA assay

REAGENTS

Reagent A: 1% BCA-$Na_2$, 2% $Na_2CO_3.H_20$, 0.16% $Na_2$ Tartrate, 0.4% NaOH and 0.95% $NaHCO_3$. (Adjust to pH 11.25 by the addition of solid $NaHCO_3$).

Reagent B: 4% $CuSO_4.5H_2O$ in sterile distilled $H_2O$.

Working Reagent: 100 vol. Reagent A + 2 vol. Reagent B.

BSA Protocol

A series of BSA (Bovine Serum Albumin) standards ranging from 0 µg/ml to 1000 µg/ml was prepared. To 100 µl of standard or suitably diluted sample, 2 ml of working reagent was added. The tubes were then incubated at 37°C for 30 minutes after which the absorbance was measured at 562 nm against a 0 µg/ml BSA blank.

#### 2.5.4.2 DC- protein assay of column eluates

As column eluates contained 6 M urea and 0.1% DDM, the Detergent Compatible assay system from BioRad was used to assay for protein in these samples. This is compatible with 4 M urea. BSA standards were prepared in 4M urea in concentrations ranging from 0 to 250 µg/ml. To 200 µl of standard or suitably diluted sample (4 M urea), 100 µl solution A' and 800 µl of solution B from the kit were added. The tubes were incubated at room temperature for 20 minutes after which the absorbance was measured at 655 nm against a blank containing 4 M urea.

### 2.5.5 Denaturing polyacrylamide gel electrophoresis of proteins (SDS PAGE)

#### 2.5.5.1 Reagents for SDS PAGE

a) 4 x load buffer 62.5 mM Tris. HCl (pH8.6), 5% glycerol, 1.6% (w/v) SDS, 0.0025% (w/v) bromophenol blue. immediatley prior to use add 3 µl β-mercaptoethanol per 100 µl

b) Anode buffer. 0.302% (w/v) Tris base, 1.44% (w/v) glycine, 0.1% (w/v) SDS.

c) Cathode buffer. 0.302% (w/v) Tris.HCl (pH8.4).

d) Molecular weight standards. Supplied as a prestained mixture of six proteins by Biorad, (Hemel Hempstead, UK). Contains; Rabbit-muscle Phosphorylase B, 97,400 Da, Bovine Serum Albumin, 66,200 Da, Hen Egg White Ovalbumin, 45,000 Da, Bovine Carbonic Anhydrase, 31,000 Da, Soybean Trysin-inhibitor, 21,000 Da, Hen Egg White Lysozyme, 14,400 Da.

### 2.5.5.2 Apparatus for SDS PAGE

5 x 17 cm gels were cast between glass plates with 1.5mm thick spacers and a 20 tooth 1.5 mm thick teflon comb. Electrophoresis was performed using a model 15-17 vertical gel electrophoresis apparatus from BRL, Paisley, UK or using the Atto mini gel system from Genetic Research Instrumentation Ltd. Dunmow, Essex, UK. This system used 90 mm x 80 mm x 1mm glass plates with a 12 well 0.8 mm thick teflon comb.

### 2.5.5.3 Protocol for SDS PAGE

Electrophoresis of the supernatant from the cell lysate, inner membrane fraction or eluate from the purification column was performed in the presence of 0·1% (w/v) SDS in a 4% (w/v) stacking gel and a 15% (w/v) running gel. A ratio of 8.6 : 1 acrylamide : bis-acrylamide was used in the gel, and polymerisation was induced by the addition of 0·1% (w/v) fresh ammonium persulphate and 0·1% (v/v) TEMED. When using the minigel apparatus, the gel was made up as a 5 % (w/v) stacking gel and 10 % (w/v) resolving gel at a ratio of 3.75 : 1 acrylamide: bisacrylamide.

1 volume of 4 x protein loading buffer was added to three volumes of sample, and the mixture denatured at 37°C for 20 minutes. 10-30 µl of the sample was loaded onto the gel, and the gel electrophoresed at a constant current of 200 mA. A prestained molecular weight ladder was also run on the gel to allow estimation of subunit Molecular weight and later to determine extent of transfer of proteins to membrane.

### 2.5.6 Staining of polyacrylamide gels

The gel was stained after electrophoresis by either the Coomassie method or the silver-stain method.

a) Coomassie method.

The gel was placed in Coomassie blue stain, (0·25% (w/v) Coomassie brilliant blue R250 in 105 (v/v) acetic acid, 45% (v/v) methanol for 30 mins, and destained in 10% (v/v) acetic acid, 45% (v/v) methanol.

b) Silver Stain method (Heukshoven & Dernick, 1988).

1)The gel was fixed in 30% (v/v) ethanol, 10% (v/v) acetic acid for 3 hours or overnight.

2) The gel was washed in10% (v/v) ethanol for 5 minutes

3) The gel was washed three times in deionised $H_2O$ for 10 minutes each time.

4) The gel was placed in Farmer's reagent for 1 minute.

Farmer's reducing reagent: 0.3% (w/v) $Na_2S_2O_3.5H_2O$, 0.15% (w/v) $KFe(CN)_6$, 0.005% (w/v) $Na_2CO_3$.

5) The gel was washed three times in deionised $H_2O$ for 10 minutes each time.

6) The gel was placed in 0.1% (w/v) $AgNO_3$ for 30 minutes .

7) The gel was quickly washed in deionised $H_2O$, (10-20 seconds).

8) The gel was placed in developer made from 2.5% (w/v) $Na_2CO_3$, 0.02% (w/v) Formaldehyde until colour developed, (ca. 5 mins).

9) The reaction was stopped by placing the gel in 30% (v/v) methanol, 10% (v/v) acetic acid for 30 minutes.

### 2.5.7. Transfer of proteins to nitro-cellulose membrane (Western blotting).

#### 2.5.7.1. Transfer Buffer

a) Transfer Buffer was prepared by the method of Towbin *et al* (1979)

25 mM Tris, 192 mM glycine. The buffer was measured for pH which should be 8.5 depending on the quality of reagents. The pH was not adjusted with acid or base as this affects the conductivity.

### 2.5.7.2 Apparatus for transfer

The transfer was performed using a model 2005 Transfer tank from LKB Bromma for large gels or a Biorad mini trans-blot tank for small gels.

### 2.5.7.3 Protocol for transfer

Following electrophoresis, the gels were rinsed for 15 minutes in transfer buffer to remove electrophoresis buffer salts and detergents, during this time the buffer was changed 3 times.

A piece of nitro-cellulose membrane (Hybond-C, Amersham, UK) cut to the dimensions of the gel was slid slowly at an angle of 45° into transfer buffer, abrupt wetting was avoided as this can lead to formation of bubbles within the matrix.

Two pieces of filter paper, (Whatman 3 mm) per gel were cut to the dimensions of the gel and saturated with transfer buffer, the sponge pads from the apparatus were also completely saturated with transfer buffer.

The apparatus tank was half filled with pre chilled transfer buffer and the cooling system, set to 12 °C and pre-run for 30 mins before the start of blotting.

The gel sandwich was built up on one plate by placing a pre-wetted sponge pad, followed by a pre-wetted filter paper, followed by the equilibrated polyacrylamide gel. The wetted membrane was lowered on top of the gel using forceps, making sure all air bubbles were excluded, and a second piece of filter paper placed on top, this was followed by the second pre-wetted sponge pad, finally the second plate was added to the sandwich and clipped firmly into place. Good contact was maintained throughout.

The whole gel holder was then placed inside the blotting tank in the correct orientation, (membrane towards the positive electrode )and the tank topped up with pre-chilled transfer buffer. The transfer was run at 60 V ( about 500 mA ) for 5 hours or at 30 V overnight (16-18 hours) for the large gels or for 1.5 hours at constant current of 200 mA when using the Bio-Rad mini trans-blot equipment. Following transfer, if not blocked straight away, membranes were removed from the sandwich using forceps , washed briefly (1-2 minutes) in transfer buffer and air dried. Once dried, the membranes were stored between two sheets of filter paper in sealed plastic bags at 4°C until required.

## 2.5.8 Immunodetection

### 2.5.8.1 Reagents for immunodetection

a) TBS buffer: 50 mM Tris-HCl, pH7.5, 150 mM NaCl.

b) TBS-Tween® buffer: 50 mM Tris-HCl pH7.5, 150 mM NaCl, 0.1% Tween® 20

c) Primary antibody: Anti-PentaHistidine, Mouse Ascites fluid, Mouse IgG (Qiagen, UK).

d) Chemiluminescence Western blotting kit, ( Boehringer Mannheim , Lewes, UK) contains luminol substrate with enhancer, starting solution, Blocking reagent, Secondary antibody: Anti-mouse IgG-POD, Fab fragments.

e) Primary blocking buffer, 10% marvel in TBS-Tween-20®

### 2.5.8.2 Protocol for immunodetection

The following washing and incubation steps were carried out at room temperature on a shaking platform.

1) Blocking

The membrane was placed in a trough with the side that had been in contact with the gel uppermost. 50ml of primary blocking buffer was added to the trough to cover the membranes and the blocking reaction was carried out overnight at room temperature.

2) Addition of primary antibody

The blot was overlaid with 50 ml of primary antibody at a dilution of 1:2000 in blocking buffer and incubated overnight at 4 °C.

3) Washing

The membrane was washed twice for 10 minutes each in at least 50 ml TBS/Tween and then twice for 10 minutes each in a similar volume of TBS.

4) Addition of secondary antibody

The membrane was incubated for 30 mins with 20 ml (125µl of solution per $cm^2$ of membrane) of secondary antibody at a dilution of 40 mU/ml in 0.5% blocking buffer from the kit (or in 10% Marvel solution as step 2)

5) Washing

As step 3.

6) Detection.

Transparent plastic roll was cut and formed into a bag slightly bigger than the size of the membrane, the membrane was placed inside the bag. The bag was sealed on three sides using a vacuum sealer. The premixed detection reagent (100 volumes of solution A to 1 volume of solution B as directed by the kit) was added to the bag at 20 µl/cm$^2$ (between 1.5 to 3.0 ml depending on size of membrane) The bag was sealed after eliminating air bubbles and incubated for 60 seconds. The membrane was removed from the bag and was placed protein side uppermost between two acetate sheets into a film cassette.

In a dark room, a photographic film was placed on top of the bag and exposed for 60 seconds. The film was then developed. A second film was placed in the cassette and exposed for a suitable time period estimated by intensity of initial signal.

## 2.6 Computer methods

The DNA and protein sequences in this thesis were analysed using the Sequence Analysis Software package, Version 8 from the Genetics Computer Group at the University of Wisconsin (GCG) (Devereux *et al*, 1984).

**BIBBPROGAC**

analysis; uses the algorithm of Bibb *et al* (1984) to find potential coding regions in the nucleotide sequences of Streptomycetes.

**GAP**: uses the algorithm of Needleman & Wunsch, (1970) to find the alignment of two complete sequences that maximises the number of matches and minimises the number of gaps.

**PEPPLOT**:

plots measures of protein secondary structure and hydrophobicity in parallel panels of the same plot. Uses a choice of 10 panels , the analysis done in this thesis used the algorithm of Goldman, Engelman and Steitz, (Engelman *et al*, 1986) and the hydropathy measure of Kyte and Doolittle (Kyte & Doolittle, 1982) to identify nonpolar transbilayer helices over a window of 9 residues.

**PILEUP**:

creates a multiple sequence alignment from a group of related sequences using progressive, pairwise alignments. It can also plot a tree showing the clustering relationship used to create the alignment.

# CHAPTER 3

## *otrB* gene and gene product

# Chapter 3.

# The *otrB* gene and gene product.

## 3.1 Introduction.

The 30 kb *otc* cluster lies 550-670 kb from one of the chromosomal ends and is known to be flanked by resistance determinants, *otrA* and *otrB*, with the *otrB* gene being closer to the chromosome end (Garven, 1994; Peric & Hranueli unpublished results). Initially *otrA* was shown to specify resistance to OTC mainly in the ribosomal fraction. OtrA shows considerable homology to the TetM and TetO gene products of *Streptococcus* and *Campylobacter*. It is suggested that the product of *otrA* confers resistance by means of a noncovalent modification of the ribosome. At the onset of this work the *otrB* gene had been cloned on 3 separate occasions (Rhodes *et al*, 1984; Ohnuki *et al*, 1985 and Reynes *et al* 1988). The function of the gene product had been demonstrated on two separate occasions (Ohnuki *et al*, 1985; Butler *et al* 1989). Ohnuki showed that the protein synthesis machinery of *S. griseus* carrying the *tetB ( otrB )* resistance determinant on a plasmid was sensitive to tetracycline and by the use of inverted membrane vesicles demonstrated that this strain exhibited reduced uptake of tritium labelled tetracycline. As reduced uptake is equivalent to increased export when using inverted vesicles it was proposed that the function of the *otrB* gene product is to implement transport of the produced antibiotic from the cell.

Two sequences for *otrB* were available, (Reynes *et al*, 1988; and MacGregor-Pryde, 1995). Reynes *et al* named the gene *tet* 347, as sequencing had revealed that the gene product was 347 amino acids in length. Studies showed that the protein was membrane associated and supported the hypothesis that it was involved in tetracycline transport. The amino acid sequence of Tet347 showed similarity to the exporter families but was smaller than the usual size for a membrane transporter and did not appear to have an upstream repressor gene. For this reason, *tet 347* was given its own branch of the evolutionary tree (Sheridan & Chopra, 1991). This hypothesis was contested by the MacGregor-Pryde (1995) sequence which clearly identifies an upstream ORF, *otrR* reading divergently from *otrB*. Analysis of this sequence revealed that the gene product of this ORF is a putative repressor belonging to the

family of *mcr* resistance/ repressor genes and that a putative helix-turn-helix DNA binding motif is present in the upstream sequence. The inverted and direct repeats of the intergenic region show similarity to the operator sequence of bacteriophage 434 as do all of the resistance repressor gene systems considered by MacGregor-Pryde (1995).

Analysis of the *otrB* sequence by MacGregor-Pryde identified one large open reading frame of 1611 bp, encoding a 537 amino acid product which shows high sequence similarity to the *tcmA* gene product (Guilfoile & Hutchinson, 1992). This implies that the OtrB protein acts as an export pump powered by the transmembrane electrochemical gradient. Studies of the deduced *otrB* gene product using hydropathy profiling and comparison analysis indicate that the protein is a member of the major facilitator superfamily (MFS), showing most similarity with members of the class III family. These and other data from the thesis of MacGregor-Pryde suggest that the Reynes sequence was erroneous and that the *otrB* gene product does conform to known Gram-positive families of tetracycline export proteins and also shows similarity to other Streptomycete antibiotic export genes.

## 3.2 Aims of this chapter

During the course of this work it was discovered that the MacGregor-Pryde sequence also contained some errors. For this reason, a region at the 5' end of the gene was cloned into pIBI24 and resequenced. The data resulting from this allowed a modified amino acid sequence for OtrB to be deduced. The aims of this chapter were:

(1) To elucidate the definitive sequence of the *otrB* gene.

(2) Analysis of the amino acid sequence with respect to other tetracycline resistance determinants and other members of the Major Facilitator Superfamily.

## 3.3 Re-sequencing of parts of *otrB*

### 3.3.1 Sequencing strategy

In order to express the OtrB protein in *E. coli* it was necessary to engineer an *Nde* I (CATATG) site into the ATG start codon of the gene. This was done using PCR mutagenesis with a mutagenic oligonucleotide primer (4.4.3). As the PCR reaction was performed using *Taq* polymerase it was important to sequence the PCR product

and compare the sequence to the known *otrB* sequence to eliminate mistakes made by the enzyme, which has no proof reading activity. The PCR product was cloned using restriction enzymes *Nde* I - *Hind* III, into pMTL23 which has annealing sites for universal and reverse sequencing primers just outside of the polylinker region and sequenced from the 5' end using an ABI automated sequencer to give c 300 bp of sequence at the start of the gene. When this sequence was compared to the MacGregor-Pryde sequence, a number of differences were immediately apparent.

The number of differences was thought to be too high to be due to the infidelity of *Taq* polymerase activity. The lack of proof reading activity of the enzyme results in some misincorporation of bases. The fidelity of *Taq* polymerase has been reported to be $2 \times 10^{-4}$, error/bp i.e. a mistake every 5 Kb, ( Saiki *et al,* 1988), although later reports are that the error rate is actually slightly lower at $8.9 \times 10^{-6}$ error/bp (Cline *et al,* 1996). In order to rule out PCR mistakes, it was decided that the PCR template itself should be sequenced . However, direct sequencing of pJJ1 was impractical because the 1.6 kb *otrB* gene was present within a larger 2.34 kb *Sac*I-*Bam*HI fragment on pJJ1 with 500 bp of DNA upstream of the ATG start site (Fig 3.1a ).

The 460 bp *Apa* I fragment containing the start codon and 401 bp of *otrB* was cloned into pIBI24 to create pJJIBI and sequenced. This sequence also showed some differences to the previous work.

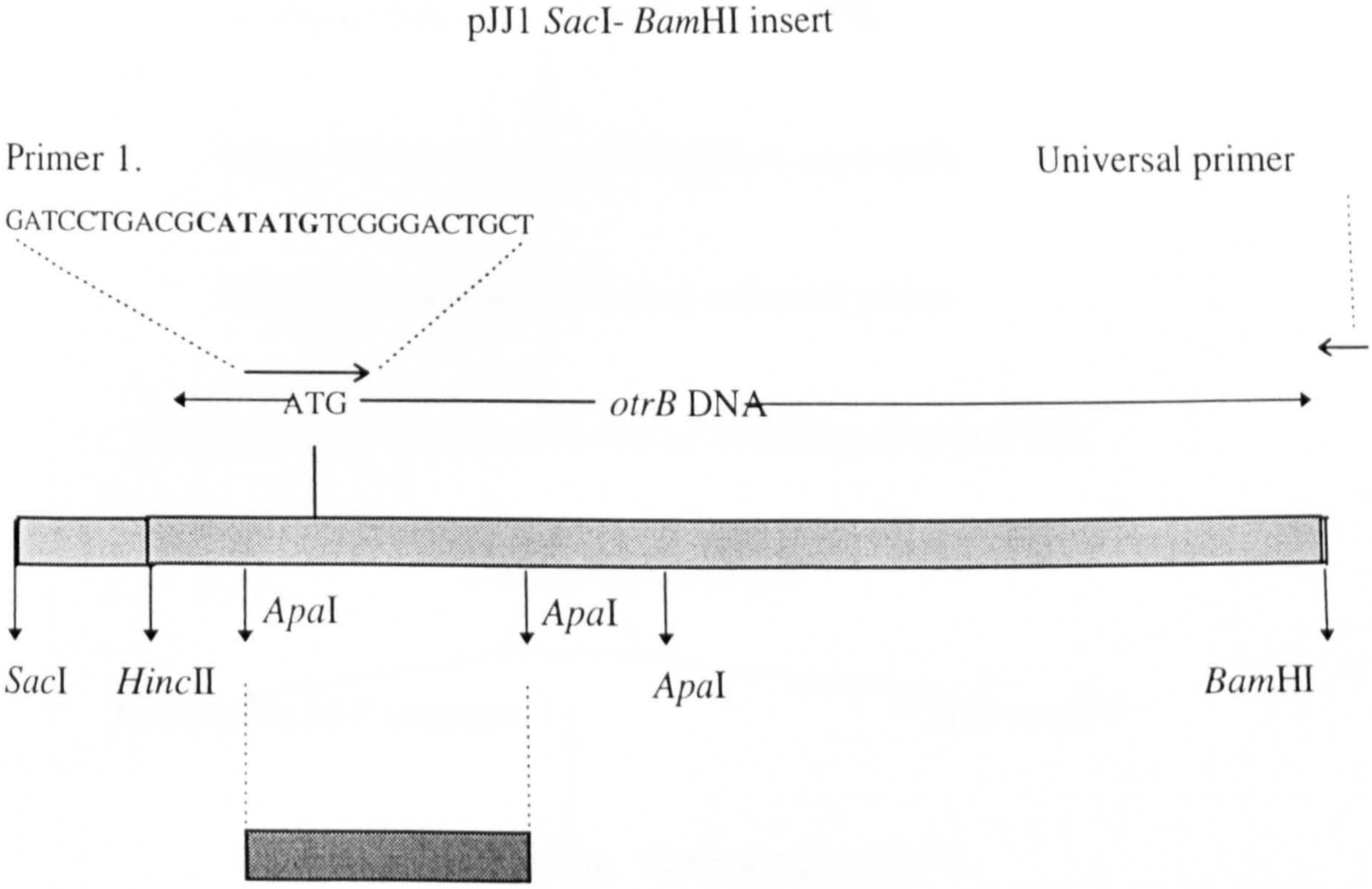



Fig. 3.1a

PCR and cloning strategy for *otrB*

Introduce *Nde*I site at ATG codon by PCR.

↓

Clone PCR product into pMTL23 to create pJJ2.

↓

Sequence from *Nde*I end using universal primer.

↓

Compare sequence data with that of MacGregor-Pryde (1995).

Differences observed

Mistakes in M-P sequence ?

PCR errors ?

↓

Sequence PCR template, subcloning required.

↓

Subclone genomic *Apa*I fragment into pIBI24, creating pJJIBI and sequence from both ends using universal and reverse primers.

↓

Mistakes in M-P sequence.

↓

Compare M-P sequence and Reynes sequence with new sequence data.

↓

PCR mistake found in pJJ3.

↓

Fresh pJJ3 clone sequenced, thought to have no PCR mistakes, this is named pJJ6.

↓

Design sequencing primers to determine sequence in positions where M-P, Reynes and new data do not agree. Analyse these data and produce definitive *otrB* sequence

Fig 3.1b

Plan of work for Chapter 3.

### 3.3.2 Analysis of sequence data

DNA sequence data were analysed using the "Sequence analysis software package, version 9" from the University of Wisconsin, ( GCG 9; Devereux *et al* 1984). This package contains a series of programs written specifically for molecular geneticists requiring to compile and analyse sequence data quickly and with accuracy. The programmes used in this analysis are listed below. Further information can be found in the previous chapter (2.6). The GCG programs used to analyse sequence data were; BESTFIT, TRANSLATE, PILEUP and COMPARE.

## 3.4 Comparison of sequence data for otrB.

The sets of sequence data were compared using the GCG program BESTFIT. 335 bases of pJJ3 were sequenced, the pJJ3 sequence differing from the MacGregor-Pryde sequence at 24 positions of the 335, more than be attributed to polymerase infidelity. Comparing the pJJIBI sequence to that of pJJ3 , the two sets of data were in agreement apart from one deletion in pJJ3, this was presumed to be a PCR mistake. A second pMTL based clone, pJJ6, was chosen at random and sequenced in the same way. The pJJ6 data was compared to the pJJIBI sequence and no differences were found in the first 266 bases of the *otrB* open reading frame (Fig 3.2), supporting the theory that the deletion in pJJ3 was due to infidelity of the *taq* polymerase. The new *otrB* sequence was then compared to the *tet347* sequence (Reynes *et al,* 1988) and to *otrB* (MacGregor-Pryde, 1995). (Figs 3.3 and 3.4.).

```
1 ATGTCGGGACTGCTGCTGGCCGTGTTCCTCGCGGCCCTGGACCAgACGGT 50
  ||||||||||||||||||||||||||||||||||||||||||||||||||||
  1 ATGTCGGGACTGCTGCTGGCCGTGTTCCTCGCGGCCCTGGACCAGACGGT 50

 51 CATCGCCACCGCGATGCGCACCATCGCGGACGACCTCCACGGCCAgACCG 100
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
 51 CATCGCCACCGCGATGCGCACCATCGCGGACGACCTCCACGGCCAgACCG 100

101 AGCAGGCATGGGCGACGACGGGCTACCTCATCGCCTCCGTCCTGGCGATG 150
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
101 AGCAGGCATGGGCGACGACGGGCTACCTCATCGCCTCCGTCCTGGCGATG 150

151 CCGTTCTACGGCAAGCTGTCCGACATCTACGGGCGCAAGCCCATGTACCT 200
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
151 CCGTTCTACGGCAAGCTGTCCGACATCTACGGGCGCAAGCCCATGTACCT 200

201 GATCTCCATCGTGGTGTTCATCGGCGGCTCGGTGCTGTGCGGCACGGCCG 250
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
201 GATCTCCATCGTGGTGTTCATCGGCGGCTCGGTGCTGTGCGGCACGGCCG 250

251 GCTCGATGTGGGAGCT 266
    ||||||||||||||||
251 GCTCgATGTGGGAGCT 266
```

Fig 3.2

BESTFIT of the wild type *otrB* sequence from pJJIBI (top) against sequence from the PCR product in pJJ6 (bottom).

```
  1 ATGTCGGGACTGCTGCTGGCCGTGTTCCTCGCGGCCCTGGACCAGACGGT 50
    ||||||||||||||||||||||||||||||||||||||||||||||||||
  1 ATGTCGGGACTGCTGCTGGCCGTGTTCCTCGCGGCCCTGGACCAGACGGT 50
          .         .         .         .         .
 51 CATCGCCACCGCGATGCGCACCATCGCGGACGACCTCCACGGCCAGACCG 100
    ||||||||||||||||||||||||||||||||||||||||||||||||||
 51 CATCGCCACCGCGATGCGCACCATCGCGGACGACCTCCACGGCCAGACCG 100
          .         .         .         .         .
101 AGCAGGCATGGGCGACGACGGGCTACCTCATCGCCTCCGTCCTGGCGATG 150
    ||||||||||||||||||||||||||||||||||||||||||||||||||
101 AGCAGGCATGGGCGACGACGGGCTACCTCATCGCCTCCGTCCTGGCGATG 150
          .         .         .         .         .
151 CCGTTCTACGGCAAGCTGTCCGACATCTACGGGCGCAAGCCCATGTACCT 200
    ||||||||||||||||||||||||||||||||||||||||||||||||||
151 CCGTTCTACGGCAAGCTGTCCGACATCTACGGGCGCAAGCCCATGTACCT 200
          .         .         .         .         .
201 GATCTCCATCGTGGTGTTCATCGGCGGCTCGGTGCTGTGCGGCACGGCCG 250
    ||||||||||||||||||||||||||||||||||||||||||||||||||
201 GATCTCCATCGTGGTGTTCATCGGCGGCTCGGTGCTGTGCGGCACGGCCG 250
          .         .         .         .         .
251 GCTCgATGTGGGAGCTGGCCCTCTTCCGGGCCGTCCAAGGGACTGGGCGG 300
    |||||||||||||||| |||||||||||||||||| ||||||||||||||
251 GCTCGATGTGGGAGCT.GCCCTCTTCCGGGCCGTCC.AGGGACTGGGCGG 298
          .         .         .         .         .
301 CGGCGGGCTGATGTCCCTGCCCACCGCGGTGGTCGCCGACCTCGCCCCGG 350
    ||||| ||||||||||||||||||||||||||||||||||||||||||||
299 CGGC.GGCTGATGTCCCTGCCCACCGCGGTGGTCGCCGACCTCGCCCCGG 347
               .         .
351 TGCG.CGAGCGCGGCCGCTACTTCGCC 376
    |||| ||||||||||||||||||||||
348 TGCGCCGAGCGCGGCCGCTACTTCGCC 374
```

Fig 3.3

Sequence comparison of the pJJ6 *otrB* PCR product (top) with *tet347* sequence (Reynes *et al*, 1988) (bottom ).

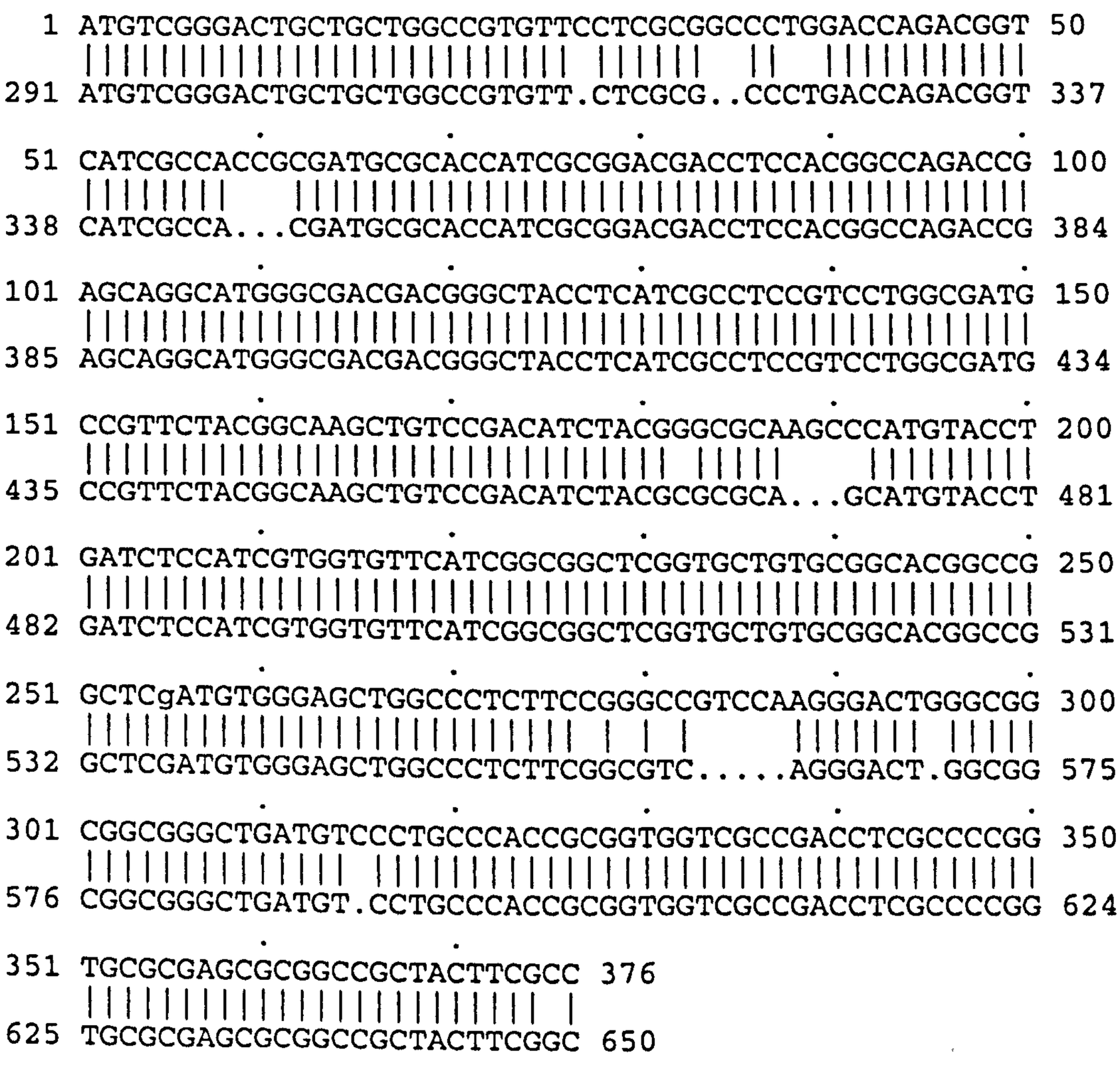



Fig 3.4

Sequence comparison of the pJJ6 *otrB* PCR product (top) against the original *otrB* sequence (bottom) (MacGregor-Pryde, 1995). It should be noted that the original MacGregor-Pryde sequence contains sequence upstream of *otrB* and that the *otrB* ORF begins at position 291.

### 3.4.1 Outcome of initial sequence analysis

As *tet347*, *otrB* (MacGregor-Pryde, 1995) and the sequence of the PCR product in pJJ3 showed a G at position 255, it appeared that the A from the automated sequence of the pJJ6 insert was a mistake, mis-naming of bases was resolved by careful study of the chromatogram (bases which have been re-called or were un-named (N's) in the original chromatogram are denoted in lower case in Figs 3.2 - 3.4.)

The mistakes were often due to peaks being spaced unevenly and overlapping each other or peaks being so small that they were easily confused with the background noise.

From the BESTFITs of *tet347* against the pJJ6 sequence, there appear to be three frameshifts in the Reynes sequence. These occur at positions 288, 305 and 355. *tet347* has an A missing at 288, a C missing at 305 and a C inserted at 355 (Fig 3.3). The C at 355 is not present in the pJJ6, pJJIBI or MacGregor-Pryde sequences. The presence of the bases at 288 and 305 in the other three sequences plus analysis of codon usage using BIBBPROGAC, (2.6) was evidence that the two bases at 288 and 305 were actually present in the gene and not sequencing artefacts and that the inserted C at 355 is also a mistake

BESTFIT analysis of the later parts of *otrB* (MacGregor-Pryde, 1995) and *tet347* shows that the two sequences are in agreement apart from insertion of a C residue at 787 and 3 GC inversions at 650, 700 and 1060 in the *otrB* sequence of MacGregor-Pryde. These residues correspond to bases 497, 359, 409 and 765 of the coding sequence from the MacGregor-Pryde data. By using BIBBPROGAC it can be shown that the "extra" C residue is required at base 497 to maintain codon preference in the reading frame.

At this stage it became apparent that further sequencing was required to resolve the three remaining differences (359, 409, 765) between the MacGregor-Pryde and Reynes sequence data.

## 3.5 Second stage sequencing strategy.

As initial sequencing data of the 5' terminal region of from pJJIBI and pJJ6 (see above) was not in agreement with that presented by MacGregor-Pryde (1995) or Reynes and co-workers (1988) further sequencing was carried out to resolve these differences and the three differences found further downstream between the MacGregor-Pryde and Reynes data (Fig 3.5).

### 3.5.1 Primer design

Oligonucleotide primers for DNA sequencing were designed against the MacGregor-Pryde sequence using the programs, Oligo 4.0-s® (distributed for MedProbe for NBI, Europe) and GeneRunner, version 3.5 (Hastings software). These programs analyse the template-primer duplex for annealing temperature, stable bond formation and G-C clamps. Three primers were designed, the first (Primer 1), being a 20-mer which read forward from base 294 the second (Primer 2), an 18-mer, reading forward from base 587, the third (Primer 3) also an 18-mer, reading in reverse from base 600. All are shown in Fig 3.5.

Following the discovery of a GC inversion at position 1469 late on in the project (See Section 5.12), it became necessary to sequence the remainder of the *otrB* gene. This step had not been taken before as the Reynes' and MacGregor-Pryde sequences were in agreement apart from the anomalies mentioned above. Four further sequencing primers were designed, (Primers 4 to 7).These are also shown in Fig 3.5

### 3.5.2 Sequencing results

Data from the sequencing reaction using Primer 1 showed that the sequence at base 359 was CC as in the Reynes' sequence and not GC as written by MacGregor-Pryde.
The data from Primer 3 resolved the differences at base 409, in this case in agreement with the MacGregor -Pryde (GGCGG) and not the Reynes sequence (GCGGC). The sequence from the reaction performed with Primer 2 gave data in agreement with the Reynes sequence at base 765, i.e. GC. not the MacGregor-Pryde data, CG.
The data from primers 4 to 7 gave the *otrB* sequence from base 287 to the end of the gene. BESTFIT analysis was performed against the previously compiled sequence.

Several differences to the Reynes/ MacGregor-Pryde sequence were found in the region not previously sequenced for this thesis; an extra C at position 1157, a deleted G at position 1168, a C-G inversion at 1296 and the previously mentioned GC-CG inversion at 1469. Later in the project sequencing of the far end of the gene was verified using T7 terminator primer, a reverse primer which read into the gene from a cloning vector (section 5.13).

### 3.5.3 BIBBPROGAC analysis

In order to verify that the sequence contained no frame shift mutations, the nucleotide sequence was translated into peptide sequence to check for stop codons, and run through a codon analysis check using BIBBPROGAC. Both analyses showed that no frame shifts were apparent in the new sequence. The BIBBPROGAC results are presented as Fig 3.6. The revised *otrB* sequence was deposited as AF061335 (Fig 3.7).

## 3.6 Results of peptide analysis

As noted above, the nucleotide sequence generated from resequencing was translated into peptide sequence, this was done using the GCG program TRANSLATE (Fig 3.7). When the peptide sequence was available, several types of analysis were possible. A hydropathy profile of a peptide sequence can be formed using the GCG program, PEPPLOT. This program utilises the hydrophobicity measure of Kyte and Doolittle. From the hydropathy profile of OtrB (Fig 3.8) It can be clearly seen that there are 14 strongly hydrophobic segments, deduced, from their length and composition, to form transmembrane alpha helices (TMH's).

```
   ATGTCGGGACTGCTGCTGGCCGTGTTCTCGCGCCCTGACCAGACGGTCATCGCCACGATG
 1 ---------+---------+---------+---------+---------+---------+ 60
   TACAGCCCTGACGACGACCGGCACAAGAGCGCGGGACTGGTCTGCCAGTAGCGGTGCTAC

   CGCACCATCGCGGACGACCTCCACGGCCAGACCGAGCAGGCATGGGCGACGACGGGCTAC
61 ---------+---------+---------+---------+---------+---------+ 120
   GCGTGGTAGCGCCTGCTGGAGGTGCCGGTCTGGCTCGTCCGTACCCGCTGCTGCCCGATG

   CTCATCGCCTCCGTCCTGGCGATGCCGTTCTACGGCAAGCTGTCCGACATCTACGCGCGC
121 ---------+---------+---------+---------+---------+---------+ 180
   GAGTAGCGGAGGCAGGACCGCTACGGCAAGATGCCGTTCGACAGGCTGTAGATGCGCGCG

   AGCATGTACCTGATCTCCATCGTGGTGTTCATCGGCGGCTCGGTGCTGTGCGGCACGGCC
181 ---------+---------+---------+---------+---------+---------+ 240
   TCGTACATGGACTAGAGGTAGCACCACAAGTAGCCGCCGAGCCACGACACGCCGTGCCGG
```

Primer 4

```
   GGCTCGATGTGGGAGCTGGCCCTCTTCGGCGTCAGGGACTGGCGGCGGCGGGCTGATGTC
241 ---------+---------+---------+---------+---------+---------+ 300
   CCGAGCTACACCCTCGACCGGGAGAAGCCGCAGTCCCTGACCGCCGCCGCCCGACTACAG
```

Primer 1

```
   CTGCCCACCGCGGTGGTCGCCGACCTCGCCCCGGTGCGCGAGCGCGGCCGCTACTTCGGC
301 ---------+---------+---------+---------+---------+---------+ 360
   GACGGGTGGCGCCACCAGCGGCTGGAGCGGGGCCACGCGCTCGCGCCGGCGATGAAGCCG
```

Apa I site (See Fig 3.1)

```
   TTCCTCCAGATGGCGTGGGTGGTCGCCAGCGTCGCGGGCCCGCTGGCGGGCGGCTTCTTC
361 ---------+---------+---------+---------+---------+---------+ 420
   AAGGAGGTCTACCGCACCCACCAGCGGTCGCAGCGCCCGGGCGACCGCCCGCCGAAGAAG

   GCGGAGGCGGGCCAGGTCTTCGGCATCGACGGCTGGCGCTGGGTGTTCCTGCTCAACGTA
421 ---------+---------+---------+---------+---------+---------+ 480
   CGCCTCCGCCCGGTCCAGAAGCCGTAGCTGCCGACCGCGACCCACAAGGACGAGTTGCAT

   CCGCTGGGCCTGCTGGCCCTGGTCACCGTGCGCAAGGCCCTGAACCTGCCGCACGAACGG
481 ---------+---------+---------+---------+---------+---------+ 540
   GGCGACCCGGACGACCGGGACCAGTGGCACGCGTTCCGGGACTTGGACGGCGTGCTTGCC
```

Primer 5

P rimer 2

```
   CGCGAGCACCGCATGGACGTACTGGGCGCGGCGGCGCTGGCGCTGTTCCTGGTGCCCCTG
541 ---------+---------+---------+---------+---------+---------+ 600
   GCGCTCGTGGCGTACCTGCATGACCCGCGCCGCCGCGACCGCGACAAGGACCACGGGGAC
```

Primer 3

```
   CTGATCGTCGCCGAACAGGGCCGGACCTGGGGCTGGGGCTCGCCGGCCGCCCTCGCCCTC
601 ---------+---------+---------+---------+---------+---------+ 660
   GACTAGCAGCGGCTTGTCCCGGCCTGGACCCCGACCCCGAGCGGCCGGCGGGAGCGGGAG

   TTCGCGCTCGGCGCGGCCGGGCTGGCGGTCTTCATCCCCGTCGAGCTGCGGCGCGGCGAC
661 ---------+---------+---------+---------+---------+---------+ 720
   AAGCGCGAGCCGCGCCGGCCCGACCGCCAGAAGTAGGGGCAGCTCGACGCCGCGCCGCTG

   GAGGCCATCCTGCCGCTGGGGCTCTTCCGGCGCGGCAGCATCGCGCTGTCGTCCGCGGTC
721 ---------+---------+---------+---------+---------+---------+ 780
   CTCCGGTAGGACGGCGACCCCGAGAAGGCCGCGCCGTCGTAGCGCGACAGCAGGCGCCAG
```

Primer 6

```
   AACTTCACCATCGGCGTCGGCATCTTCGGCACGGTCACCACCCTGCCGCTGTTCCTCCAG
781 ---------+---------+---------+---------+---------+---------+ 840
   TTGAAGTGGTAGCCGCAGCCGTAGAAGCCGTGCCAGTGGTGGGACGGCGACAAGGAGGTC

   ATGGTGCAGGGGCGGACCCCGACCCAGGCCGGACTGGTGGTCATCCCGTTCATGCTGGGC
841 ---------+---------+---------+---------+---------+---------+ 900
   TACCACGTCCCCGCCTGGGGCTGGGTCCGGCCTGACCACCAGTAGGGCAAGTACGACCCG

   ACCATCGCCTCGCAGATGGTCTCCGGCAAGCTCATCGCGTCCTCGGGCCGGTTCAAGAAA
901 ---------+---------+---------+---------+---------+---------+ 960
   TGGTAGCGGAGCGTCTACCAGAGGCCGTTCGAGTAGCGCAGGAGCCCGGCCAAGTTCTTT

   CTGGCGATCGTGGGCCTGGGCTCGATGGCCGGGGCGCTGCTGGCCATGGCCACCACCGGC
961 ---------+---------+---------+---------+---------+---------+ 1020
   GACCGCTAGCACCCGGACCCGAGCTACCGGCCCCGCGACGACCGGTACCGGTGGTGGCCG
```

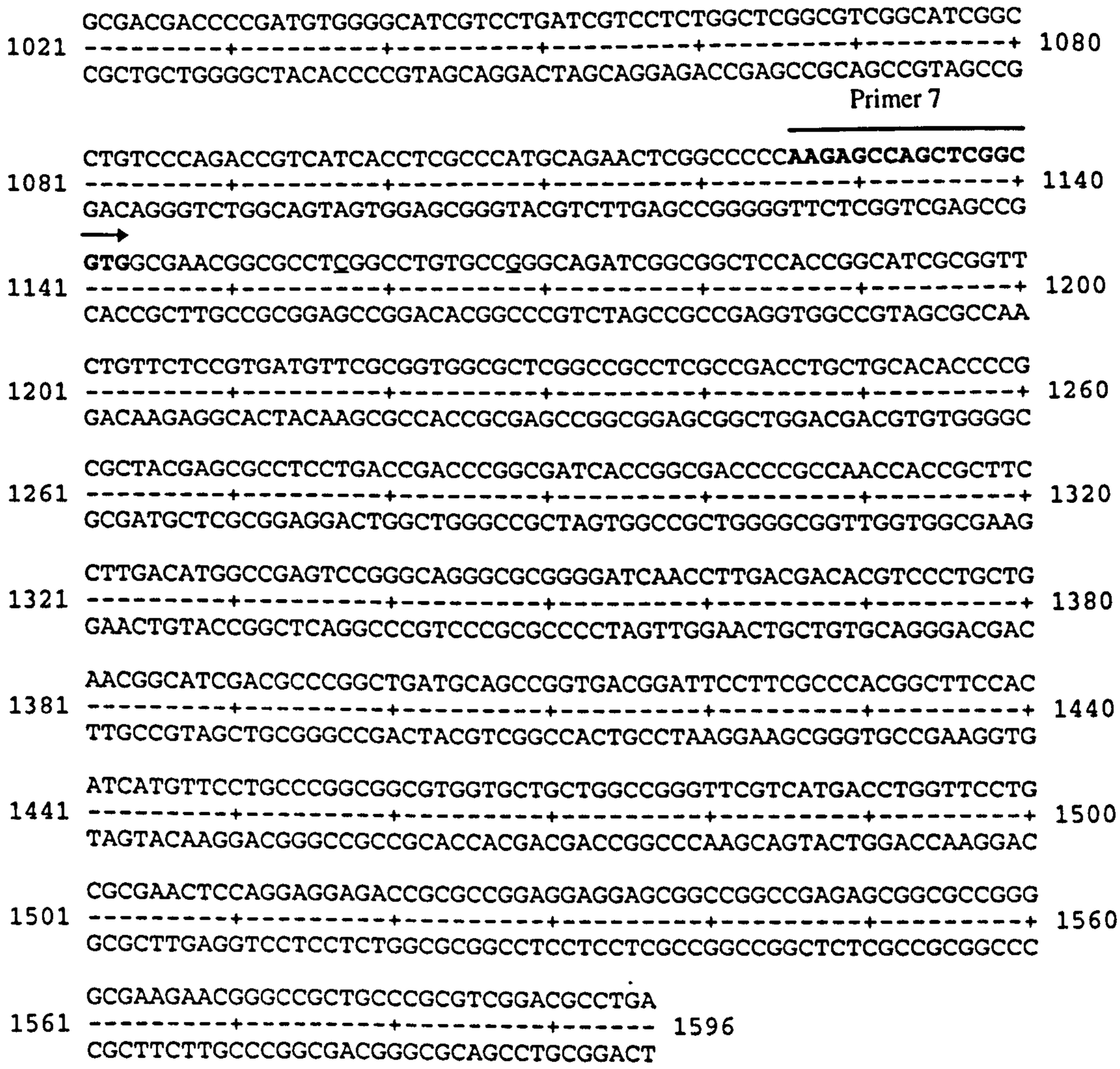



Fig 3.5

MacGregor-Pryde *otrB* sequence and primers designed against it. Numbering refers to the *otrB* ORF, and not the whole otrB sequence. Primers were used to determine definitive sequence over regions with sequence anomalies. The seven sequencing primers (arrows) are shown with their annealing positions. Anomalies between the three sets of sequence data, (this work, MacGregor-Pryde, 1995 and Reynes *et al*, 1988) are underlined. Also shown (boxed) is the second *Apa*I sit (Fig 3.1). DNA sequence upstream of this is discussed in section 3.4. Later in the project sequencing was also performed at the far C-terminal end of the gene reading in from vector sequence using T7 terminator primer (section 5.13).

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 16 | 0.700 | 0.300 | 0.900 | 811 | 0.400 | 0.400 | 1.000 |
| 31 | 0.900 | 0.300 | 1.000 | ...826 | 0.500 | 0.500 | 1.000 |
| 46 | 0.700 | 0.500 | 1.000 | 841 | 0.600 | 0.400 | 1.000 |
| 61 | 0.400 | 0.600 | 1.000 | 856 | 0.800 | 0.300 | 1.000 |
| 76 | 0.700 | 0.400 | 1.000 | 871 | 0.700 | 0.600 | 1.000 |
| 91 | 0.900 | 0.300 | 1.000 | 886 | 0.700 | 0.500 | 0.900 |
| 106 | 0.600 | 0.700 | 0.900 | 901 | 0.700 | 0.300 | 0.900 |
| 121 | 0.500 | 0.700 | 0.900 | 916 | 0.500 | 0.500 | 1.000 |
| 136 | 0.600 | 0.400 | 1.000 | 931 | 0.400 | 0.500 | 1.000 |
| 151 | 0.500 | 0.400 | 1.000 | 946 | 0.400 | 0.500 | 1.000 |
| 166 | 0.400 | 0.300 | 1.000 | 961 | 0.400 | 0.500 | 0.900 |
| 181 | 0.500 | 0.400 | 1.000 | 976 | 0.600 | 0.400 | 0.900 |
| 196 | 0.400 | 0.400 | 1.000 | 991 | 0.700 | 0.500 | 1.000 |
| 211 | 0.400 | 0.200 | 1.000 | 1006 | 0.800 | 0.600 | 1.000 |
| 226 | 0.600 | 0.400 | 1.000 | 1021 | 0.700 | 0.700 | 1.000 |
| 241 | 0.600 | 0.800 | 1.000 | 1036 | 0.400 | 0.800 | 1.000 |
| 256 | 0.600 | 0.700 | 1.000 | 1051 | 0.500 | 0.600 | 1.000 |
| 271 | 0.700 | 0.400 | 1.000 | 1066 | 0.600 | 0.300 | 1.000 |
| 286 | 0.900 | 0.500 | 0.900 | 1081 | 0.700 | 0.400 | 1.000 |
| 301 | 0.800 | 0.600 | 0.900 | 1096 | 0.700 | 0.500 | 1.000 |
| 316 | 0.700 | 0.600 | 1.000 | 1111 | 0.400 | 0.500 | 1.000 |
| 331 | 0.900 | 0.500 | 1.000 | 1126 | 0.400 | 0.600 | 1.000 |
| 346 | 1.000 | 0.500 | 1.000 | 1141 | 0.600 | 0.500 | 1.000 |
| 361 | 0.800 | 0.600 | 1.000 | 1156 | 0.700 | 0.500 | 1.000 |
| 376 | 0.600 | 0.400 | 1.000 | 1171 | 0.700 | 0.700 | 1.000 |
| 391 | 0.600 | 0.400 | 1.000 | 1186 | 0.600 | 0.700 | 1.000 |
| 406 | 0.800 | 0.600 | 1.000 | 1201 | 0.600 | 0.600 | 0.900 |
| 421 | 0.800 | 0.600 | 1.000 | 1216 | 0.500 | 0.400 | 0.900 |
| 436 | 0.800 | 0.600 | 1.000 | 1231 | 0.600 | 0.300 | 1.000 |
| 451 | 0.800 | 0.400 | 1.000 | 1246 | 0.900 | 0.500 | 1.000 |
| 466 | 0.600 | 0.500 | 1.000 | 1261 | 0.900 | 0.500 | 1.000 |
| 481 | 0.600 | 0.400 | 0.900 | 1276 | 0.800 | 0.400 | 1.000 |
| 496 | 0.800 | 0.200 | 0.900 | 1291 | 0.800 | 0.500 | 1.000 |
| 511 | 0.900 | 0.400 | 1.000 | 1306 | 0.700 | 0.600 | 1.000 |
| 526 | 0.700 | 0.400 | 1.000 | 1321 | 0.600 | 0.500 | 1.000 |
| 541 | 0.800 | 0.400 | 0.900 | 1336 | 0.700 | 0.300 | 0.900 |
| 556 | 0.900 | 0.400 | 0.900 | 1351 | 0.800 | 0.500 | 0.900 |
| 571 | 0.900 | 0.400 | 0.900 | 1366 | 0.700 | 0.500 | 0.900 |
| 586 | 1.000 | 0.500 | 0.900 | 1381 | 0.600 | 0.300 | 0.900 |
| 601 | 0.900 | 0.400 | 1.000 | 1396 | 0.600 | 0.400 | 1.000 |
| 616 | 0.800 | 0.200 | 0.900 | 1411 | 0.700 | 0.400 | 1.000 |
| 631 | 0.700 | 0.500 | 0.900 | 1426 | 0.600 | 0.400 | 0.900 |
| 646 | 0.600 | 0.900 | 1.000 | 1441 | 0.600 | 0.400 | 0.900 |
| 661 | 0.800 | 0.800 | 1.000 | 1456 | 0.600 | 0.300 | 1.000 |
| 676 | 0.900 | 0.600 | 1.000 | 1471 | 0.700 | 0.300 | 1.000 |
| 691 | 0.900 | 0.600 | 1.000 | 1486 | 0.900 | 0.400 | 1.000 |
| 706 | 0.800 | 0.400 | 1.000 | 1501 | 0.500 | 0.400 | 1.000 |
| 721 | 0.800 | 0.400 | 1.000 | 1516 | 0.600 | 0.300 | 0.900 |
| 736 | 0.900 | 0.500 | 1.000 | 1531 | 0.900 | 0.400 | 0.900 |
| 751 | 0.800 | 0.400 | 1.000 | 1546 | 0.900 | 0.600 | 1.000 |
| 766 | 0.700 | 0.600 | 1.000 | 1561 | 0.900 | 0.700 | 1.000 |
| 781 | 0.600 | 0.700 | 1.000 | 1576 | 0.700 | 0.700 | 1.000 |
| 796 | 0.400 | 0.500 | 1.000 | 1591 | 0.700 | 0.600 | 1.000 |

Fig 3.6

BIBBPROGAC analysis of revised *otrB* sequence

The analysis shows that the average G+C composition of column 1 is around 70% while for column 2 it is over 50%, column 3 displays a G+C composition of over 90%.This is consistent with an open reading frame starting at base 1. (Percentages are displayed as decimals where 1.000 = 100%). Columns represent codon positions 1,2 and 3 for a 30 base window, step length is 15 bases. (Refer to section 2.6)

```
  ATGTCGGGACTGCTGCTGGCCGTGTTCCTCGCGGCCCTGGACCAGACGGTCATCGCCAC
1 --------+---------+---------+---------+---------+---------+---------+ 60
  TACAGCCCTGACGACGACCGGCACAAGGAGCGCCGGGACCTGGTCTGCCAGTAGCGGTGG

  M  S  G  L  L  L  A  V  F  L  A  A  L  D  Q  T  V  I  A  T   -

  GCGATGCGCACCATCGCGGACGACCTCCACGGCCAGACCGAGCAGGCATGGGCGACGACG
61 --------+---------+---------+---------+---------+---------+---------+ 120
  CGCTACGCGTGGTAGCGCCTGCTGGAGGTGCCGGTCTGGCTCGTCCGTACCCGCTGCTGC

  A  M  R  T  I  A  D  D  L  H  G  Q  T  E  Q  A  W  A  T  T   -

  GGCTACCTCATCGCCTCCGTCCTGGCGATGCCGTTCTACGGCAAGCTGTCCGACATCTAC
121 --------+---------+---------+---------+---------+---------+---------+ 180
  CCGATGGAGTAGCGGAGGCAGGACCGCTACGGCAAGATGCCGTTCGACAGGCTGTAGATG

  G  Y  L  I  A  S  V  L  A  M  P  F  Y  G  K  L  S  D  I  Y   -

  GGGCGCAAGCCCATGTACCTGATCTCCATCGTGGTGTTCATCGGCGGCTCGGTGCTGTGC
181 --------+---------+---------+---------+---------+---------+---------+ 240
  CCCGCGTTCGGGTACATGGACTAGAGGTAGCACCACAAGTAGCCGCCGAGCCACGACACG

  G  R  K  P  M  Y  L  I  S  I  V  V  F  I  G  G  S  V  L  C   -

  GGCACGGCCGGCTCGATGTGGGAGCTGGCCCTCTTCCGGGCCGTCCAGGGACTGGGCGGC
241 --------+---------+---------+---------+---------+---------+---------+ 300
  CCGTGCCGGCCGAGCTACACCCTCGACCGGGAGAAGGCCCGGCAGGTCCCTGACCCGCCG

  G  T  A  G  S  M  W  E  L  A  L  F  R  A  V  Q  G  L  G  G   -

  GGCGGGCTGATGTCCCTGCCCACCGCGGTGGTCGCCGACCTCGCCCCGGTGCGCGAGCGC
301 --------+---------+---------+---------+---------+---------+---------+ 360
  CCGCCCGACTACAGGGACGGGTGGCGCCACCAGCGGCTGGAGCGGGGCCACGCGCTCGCG

  G  G  L  M  S  L  P  T  A  V  V  A  D  L  A  P  V  R  E  R   -

  GGCCGCTACTTCGCCTTCCTCCAGATGGCGTGGGTGGTCGCCAGCGTCGCGGGCCCGCTG
361 --------+---------+---------+---------+---------+---------+---------+ 420
  CCGGCGATGAAGCGGAAGGAGGTCTACCGCACCCACCAGCGGTCGCAGCGCCCGGGCGAC

  G  R  Y  F  A  F  L  Q  M  A  W  V  V  A  S  V  A  G  P  L   -

  GCGGGCGGCTTCTTCGCGGAGGCGGGCCAGGTCTTCGGCATCGACGGCTGGCGCTGGGTG
421 --------+---------+---------+---------+---------+---------+---------+ 480
  CGCCCGCCGAAGAAGCGCCTCCGCCCGGTCCAGAAGCCGTAGCTGCCGACCGCGACCCAC

  A  G  G  F  F  A  E  A  G  Q  V  F  G  I  D  G  W  R  W  V   -

  TTCCTGCTCAACGTACCGCTGGGCCTGCTGGCCCTGGTCACCGTGCGCAAGGCCCTGAAC
481 --------+---------+---------+---------+---------+---------+---------+ 540
  AAGGACGAGTTGCATGGCGACCCGGACGACCGGGACCAGTGGCACGCGTTCCGGGACTTG

  F  L  L  N  V  P  L  G  L  L  A  L  V  T  V  R  K  A  L  N   -

  CTGCCGCACGAACGGCGCGAGCACCGCATGGACGTACTGGGCGCGGCGGCGCTGGCGCTG
541 --------+---------+---------+---------+---------+---------+---------+ 600
  GACGGCGTGCTTGCCGCGCTCGTGGCGTACCTGCATGACCCGCGCCGCCGCGACCGCGAC

  L  P  H  E  R  R  E  H  R  M  D  V  L  G  A  A  A  L  A  L   -

  TTCCTGGTGCCCCTGCTGATCGTCGCCGAACAGGGCCGGACCTGGGGCTGGGGCTCGCCG
601 --------+---------+---------+---------+---------+---------+---------+ 660
  AAGGACCACGGGGACGACTAGCAGCGGCTTGTCCCGGCCTGGACCCCGACCCCGAGCGGC

  F  L  V  P  L  L  I  V  A  E  Q  G  R  T  W  G  W  G  S  P   -

  GCCGCCCTCGCCCTCTTCGCGCTCGGCGCGGCCGGGCTGGCGGTCTTCATCCCCGTCGAG
661 --------+---------+---------+---------+---------+---------+---------+ 720
  CGGCGGGAGCGGGAGAAGCGCGAGCCGCGCCGGCCCGACCGCCAGAAGTAGGGGCAGCTC

  A  A  L  A  L  F  A  L  G  A  A  G  L  A  V  F  I  P  V  E   -
```

```
    CTGCGGCGCGGCGACGAGGCCATCCTGCCGCTGGGGCTCTTCCGGCGCGGCAGCATCGCG
721 ---------+---------+---------+---------+---------+---------+ 780
    GACGCCGCGCCGCTGCTCCGGTAGGACGGCGACCCCGAGAAGGCCGCGCCGTCGTAGCGC

    L  R  R  G  D  E  A  I  L  P  L  G  L  F  R  R  G  S  I  A  -

    CTGTGCTCCGCGGTCAACTTCACCATCGGCGTCGGCATCTTCGGCACGGTCACCACCCTG
781 ---------+---------+---------+---------+---------+---------+ 840
    GACACGAGGCGCCAGTTGAAGTGGTAGCCGCAGCCGTAGAAGCCGTGCCAGTGGTGGGAC

    L  C  S  A  V  N  F  T  I  G  V  G  I  F  G  T  V  T  T  L  -

    CCGCTGTTCCTCCAGATGGTGCAGGGGCGGACCCCGACCCAGGCCGGACTGGTGGTCATC
841 ---------+---------+---------+---------+---------+---------+ 900
    GGCGACAAGGAGGTCTACCACGTCCCCGCCTGGGGCTGGGTCCGGCCTGACCACCAGTAG

    P  L  F  L  Q  M  V  Q  G  R  T  P  T  Q  A  G  L  V  V  I  -

    CCGTTCATGCTGGGCACCATCGCCTCGCAGATGGTCTCCGGCAAGCTCATCGCGTCCTCG
901 ---------+---------+---------+---------+---------+---------+ 960
    GGCAAGTACGACCCGTGGTAGCGGAGCGTCTACCAGAGGCCGTTCGAGTAGCGCAGGAGC

    P  F  M  L  G  T  I  A  S  Q  M  V  S  G  K  L  I  A  S  S  -

    GGCCGGTTCAAGAAACTGGCGATCGTGGGCCTGGGCTCGATGGCCGGGGCGCTGCTGGCC
961 ---------+---------+---------+---------+---------+--------- 1020
    CCGGCCAAGTTCTTTGACCGCTAGCACCCGGACCCGAGCTACCGGCCCCGCGACGACCGG

    G  R  F  K  K  L  A  I  V  G  L  G  S  M  A  G  A  L  L  A  -

    ATGGCCACCACCGGCGCGACGACCCCGATGTGGGGCATCGTCCTGATCGTCCTCTGGCTC
1021 ---------+---------+---------+---------+---------+--------- 1080
    TACCGGTGGTGGCCGCGCTGCTGGGGCTACACCCCGTAGCAGGACTAGCAGGAGACCGAG

    M  A  T  T  G  A  T  T  P  M  W  G  I  V  L  I  V  L  W  L  -

    GGCGTCGGCATCGGCCTGTCCCAGACCGTCATCACCTCGCCCATGCAGAACTCGGCCCCC
1081 ---------+---------+---------+---------+---------+--------- 1140
    CCGCAGCCGTAGCCGGACAGGGTCTGGCAGTAGTGGAGCGGGTACGTCTTGAGCCGGGGG

    G  V  G  I  G  L  S  Q  T  V  I  T  S  P  M  Q  N  S  A  P  -

    AAGAGCCAGCTCGGCGTGGCGAACGGCGCCTCCGGCCTGTGCCGGCAGATCGGCGGCTCC
1141 ---------+---------+---------+---------+---------+--------- 1200
    TTCTCGGTCGAGCCGCACCGCTTGCCGCGGAGGCCGGACACGGCCGTCTAGCCGCCGAGG

    K  S  Q  L  G  V  A  N  G  A  S  G  L  C  R  Q  I  G  G  S

    ACCGGCATCGCGGTTCTGTTCTCCGTGATGTTCGCGGTGGCGCTCGGCCGCCTCGCCGAC
1201 ---------+---------+---------+---------+---------+--------- 1260
    TGGCCGTAGCGCCAAGACAAGAGGCACTACAAGCGCCACCGCGAGCCGGCGGAGCGGCTG

    T  G  I  A  V  L  F  S  V  M  F  A  V  A  L  G  R  L  A  D  -

    CTGCTGCACACCCCGCGCTACGAGCGCCTCCTGACGGACCCGGCGATCACCGGCGACCCC
1261 ---------+---------+---------+---------+---------+--------- 1320
    GACGACGTGTGGGGCGCGATGCTCGCGGAGGACTGCCTGGGCCGCTAGTGGCCGCTGGGG

    L  L  H  T  P  R  Y  E  R  L  L  T  D  P  A  I  T  G  D  P  -

    GCCAACCACCGCTTCCTTGACATGGCCGAGTCCGGGCAGGGCGCGGGGATCAACCTTGAC
1321 ---------+---------+---------+---------+---------+--------- 1380
    CGGTTGGTGGCGAAGGAACTGTACCGGCTCAGGCCCGTCCCGCGCCCCTAGTTGGAACTG

    A  N  H  R  F  L  D  M  A  E  S  G  Q  G  A  G  I  N  L  D  -

    GACACGTCCCTGCTGAACGGCATCGACGCCCGGCTGATGCAGCCGGTGACGGATTCCTTC
1381 ---------+---------+---------+---------+---------+--------- 1440
    CTGTGCAGGGACGACTTGCCGTAGCTGCGGGCCGACTACGTCGGCCACTGCCTAAGGAAG

    D  T  S  L  L  N  G  I  D  A  R  L  M  Q  P  V  T  D  S  F  -
```

```
     GCCCACGGCTTCCACATCATGTTCCTCGCCGGCGGCGTGGTGCTGCTGGCCGGGTTCGTC
1441 --------+---------+---------+---------+---------+---------+--------- 1500
     CGGGTGCCGAAGGTGTAGTACAAGGAGCGGCCGCCGCACCACGACGACCGGCCCAAGCAG

     A  H  G  F  H  I  M  F  L  A  G  G  V  V  L  L  A  G  F  V  -

     ATGACCTGGTTCCTGCGCGAACTCCAGGAGGAGACCGCGCCGGAGGAGGAGCGGCCGGCC
1501 --------+---------+---------+---------+---------+---------+--------- 1560
     TACTGGACCAAGGACGCGCTTGAGGTCCTCCTCTGGCGCGGCCTCCTCCTCGCCGGCCGG

     M  T  W  F  L  R  E  L  Q  E  E  T  A  P  E  E  E  R  P  A  -

     GAGAGCGGCGCCGGGGCGAAGAACGGGCCGCTGCCCGCGTCGGACGCCTGA
1561 --------+---------+---------+---------+---------+---------+- 1611
     CTCTCGCCGCGGCCCCGCTTCTTGCCCGGCGACGGGCGCAGCCTGCGGACT

     E  S  G  A  G  A  K  N  G  P  L  P  A  S  D  A  *  -
```

Fig 3.7

The revised nucleotide sequence for *otrB*, deposited as AF061335 and the translated peptide sequence.

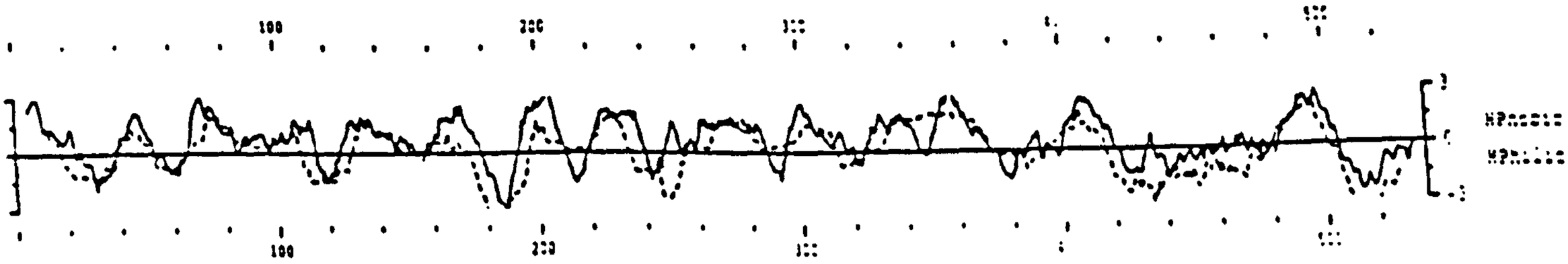


Fig 3.8

Hydropathy plot for OtrB obtained using PEPPLOT.

### 3.6.1 Hydropathy profile of OtrB

The plot (2.6. for details and reference) produced is a curve which is the average of a residue-specific hydropathy index over a window of nine residues. A curve drawn in the upper half of the frame indicates a hydrophobic region and when in the lower a hydrophilic region is indicated. The hydropathy plot for OtrB is shown in Fig 3.8

Von Heijne (1986) made observations that the distribution of certain amino acids is biased with respect to the topology of membrane proteins. Hydrophobic residues (F, I, L, V and M) are twice as likely to appear in membrane spanning regions as in cytoplasmic or periplasmic loops. The "positive inside rule" says that the cytosplasmic and periplasmic loops may be differentiated by the fact that positively charged residues, R and K appear four times as frequently in cytoplasmic loops as they do in the periplasmic variety. The putative positions of the 14 TMH's within the OtrB sequence are shown in Fig 3.9.

### 3.6.2 Major facilitator superfamily

The deduction of 14 putative TMH's of OtrB its function as an antibiotic exporter lead to its assignment within a rapidly expanding group of proteins, the Major Facilitator Superfamily (for a review see Paulsen *et al*, 1996). This family contains membrane transport proteins from all forms of life, from bacteria to higher eukaryotes and has more than 300 members (1.7).

The MFS was originally sub-divided into 5 families on the basis of substrate (Marger & Saier, 1993); transporters for, (i) drug export, (ii) sugar uptake, (iii) uptake of Krebs cycle intermediates, (iv) phosphate-ester/phosphate antiport, and (v) oligosaccharide uptake. Family (i) which contained proteins with both 12 and 14 TMS's is now further divided into the 12 TMH exporters and the 14 TMH exporters, (Paulsen *et al*, 1996). A phylogenetic tree was shown in Fig 1.14. OtrB is presumed to be a member of the 14 TMH exporter sub-family. The best way to assess the relationship of OtrB to other transport proteins was to use the GCG program, PILEUP ( 2.6).

Several PILEUP analyses were performed. Fig 3.10 shows the conserved residues between several Streptomycete 14 TMH drug exporters. Fig 3.11 shows a

MSGLLLAVFL AALDQTVIAT AMRTIADDLH GQTEQAWATT GYLIASVLAM

PFYGKLSDIY GRKPMYLISI VVFIGGSVLC GTAGSMWELA LFRAVQGLGG

GGLMSLPTAV VADLAPVRER GRYFAFLQMA WVVASVAGPL AGGFFAEAGQ

VFGIDGWRWV FLLNVPLGLL ALVTVRKALN LPHERREHRM DVLGAAALAL

FLVPLLIVAE QGRTWGWGSP AALALFALGA AGLAVFIPVE LRRGDEAILP

LGLFRRGSIA LCSAVNFTIG VGIFGTVTTL PLFLQMVQGR TPTQAGLVVI

PFMLGTIASQ MVSGKLIASS GRFKKLAIVG LGSMAGALLA MATTGATTPM

WGIVLIVLWL GVGIGLSQTV ITSPMQNSAP KSQLGVANGA SGLCRQIGGS

TGIAVLFSVM FAVALGRLAD LLHTPRYERL LTDPAITGDP ANHRFLDMAE

SGQGAGINLD DTSLLNGIDA RLMQPVTDSF AHGFHIMFLA GGVVLLAGFV

MTWFLRELQE ETAPEEERPA ESGAGAKNGP LPASDA

Fig.3.9

Shows the putative OtrB transmembrane regions (boxed). Residues in bold type belong to conserved motifs of the MFS family (Paulsen *et al*, 1996).

```
                1                                                    50
     OTRB  ---------- ---------- ---------- MSG.LLLAVF LAALDQTVIA
CTRB-NEND  MANATSQTGE AVADEAGGPA GFTHRQIITA LSG.LLLAVL LAALDQTIVS
   SVORF4  ---MTDTANR RSATPDSGAD AVDKRRIRLI MTGLLLLGLF MAALDQTIIS
     CMCT  ----------M TSVRGASKTG RTSKTSTATT ALVLACTAHF LVVFDTSVIT
     PUR8  ----MARKPD ISAVPVESAA CQGPDPRRWW GLVVILAAQL LVVLDGTVVN
    ACTVA  ----------M TANPGRPGGP ADQGHPRRWA ILGVLVLSLV GIILDNTVLN
      PTR  ----------M TATTTETSKA PSGGHPQRWL ILGVICLAQL TVLLDNTVLS
     TCMA  --MSTETHDE PSGVAHTPAS GLRGRP..WP TLLAVAVGVM MVALDSTIVA
      MMR  ---MTTVRTG GAQTAEVPAG GRRDVPS.GV KITALATGFV MATLDVTVVN
                                                          LDxTvxn

                51                                                  100
     OTRB  TAMRTIADDL HGQ....TEQ AWATTGYLIA SVLAMPFYGK LSDIYGRKPM
     CTRB  TALRTIGDQL HGQ....TVQ AWVITGYLVS STIAMPFYGK LSDIYGRKPL
   SVORF4  AALRTIADDL NGL....SEQ AWANTSYMIT SVIMTALYGK LSDIYGRRPV
     CMCT  VALPSV...R ADLGFAPASL QWVVNSYTLA FAGLLLFGGR LADIHGHRRV
     PUR8  IALPSV...Q RDLGMSDTSR QWVITAYTLA FGGLLLLGGR VADAFGRRRI
    ACTVA  VTLRTLTDPE QGLGASHSQV EWVLSAYTLA FAATLFTWGV LGDRLGRRRV
      PTR  VAIPSLT... RELHASTADI QWMINAYSLV QSGLLLTAGN AADRYGRRKM
     TCMA  IANPAI...Q QDLHASLADV QWITNGYLLA LAVSLITAGK LGDRFGHRQT
      MMR  VAGATI...Q ESLDTTLTQL TWIVDGYVLT FASLLMLAGG LANRIGAKTV
           vAIP M  (motif D)                          Gx LaDrxGrkxx

                101                                                 150
     OTRB  YLISIVVFIG GSVLCGTAGS MWELALFRAV QGLGGGGLMS LPTAVVADLA
     CTRB  YLAAIAVFIV GSAACAMANS METLAIARVL QGFGGAGLMS LPTAVIADLA
   SVORF4  YCTAVGVFVL GSVLCGLAQS MTTLAVFRG. QGIGAGGLMS LAFAILTDLV
     CMCT  FLGGLAVFTL TSLIGGLATS PASLIAARAG QGAGAAVLAP LAVTMLTTSF
     PUR8  FAVGILGFGL ASLLGGAAPD PGTLFLARAL QGVFAAALAP AALALINTLF
    ACTVA  LLLGLGLFGL SSLAGAYAGS PEQLIAARAC MGVSGAAVLP STLATIAAVF
      PTR  LVAGLALFGI GSLAAGLAQT SGQLIAARAG MGVGGALLMT TTLAVVVQVF
     TCMA  FLVGVAGFAV TSAAIGLSGS VAAIVVFRVL QGLFGALMQP SALGLLRVTF
      MMR  YLWGMGVFFL ASLACALAPT AETLIAARLV QGAGAALFMP SSLSLLVFSF
           xL (motif A)                    lxxxRx qGxgaa (motif B)

                151                                                 200
     OTRB  PV.RERGRYF AFLQMAWVVA SVAGPLAGGF FAEAGQVFGI DGWRWVFLLN
     CTRB  PV.RERGRYF SYLMMAWVAA SVLGPLVGGL FAGAGEILGV TGWRWAFLIN
   SVORF4  PL.AERSRYQ AWFGAVFGVS AVVGPVAGGF FAGLDSFLGA SGWRWAFFIN
     CMCT  AEGPRRTRAL TISTAVALVG GASGNLLGGV FTEF..... LSWRSVLLVN
     PUR8  TEPGERGKAF GVYGAVSGGG AAVGLLAGGL LTEY..... LDWRWCLYVN
    ACTVA  PL.RERPKAL GIWAASVGFA LGIGPVTGGI LLAH..... FWWGSVLLVN
      PTR  DE.TERVKAI ALWSTVSSLG FAAGPLIGGV MLEH..... YWWGAIFLIN
     TCMA  PPGK.LNMAI GIWSGVVGAS TAAGPIIGGL LVQH..... VGWEAVFFIN
      MMR  PEKRQRTRML GLWSAIVATS SGLGPTVGGL MVSA..... FGWESIFLLN
                                g xxxxGPxxGG l(motif C)    WxwxFlIN

                201                                                 250
     OTRB  VPLGLLALVT VRKALNLPHE R...REHRMD VLGAAALALF LVPLLI.VAE
     CTRB  VPLGLVALLS VRKALNLPHR R...VDHPID FRGALTLALC LVPLLI.VAE
   SVORF4  VPIGVAGALI IAALVRVQPG R...AQHKFD IAGLLALIAC LLPLLF.AVE
     CMCT  VPIGIPVLFL AARVLAGPRK RPWGR.VRLD LPGAVLATAG LTLLTLGVSQ
     PUR8  APVALLAL.L GCRLL..PRD RRTGRAVRLD LPGTLLGCGG LVAIVYAFAE
    ACTVA  VPLMAGCL.V AVVLVVPETR GTAGR..RVD AAGLLLSIAG VVPLVYAIIE
      PTR  IPVAVIGL.V AVVLLVPESK NPQGD..RPD LLGAVLSTIG MTAVVYAII.
     TCMA  VPVGLAALVA GLVILTDARA ERAPK..SFD VSGIVLLSGA MFCLVWGLIK
      MMR  LPIG..AIGM AMTYRYIAAT ESRAT..RLA VPGHLLWIVA LAAVSFALIE
           vPig (motif H)                        D xxGxxL (motif E)

                251                                                 300
     OTRB  QGRTWGWGSP AALALFALGA AGLAVFIPVE LRRGDEAILP LGLFRRGSIA
     CTRB  EGLDWGWGSA RSLTLFAVSL IGLVLFVLAE RARGLEAMVP LRLFRRGGIT
   SVORF4  QGPAWGWTDP GTLVLFGLGA VGLVLFSLAQ KRAADAALLP TPMFRNPLFS
     CMCT  T.HEHGWGEA AVAVPLAGGL LALLAFVVVE ARFAASPLIP PRLFGLPGVG
     PUR8  A..ESGWGDP LVVRLLVLGV LMLVAFALVE RR.VQDPLLP PGVVAHRVRG
    ACTVA  AGRSGGVTRP AVWAAGLAGL GLLLVFLWHE RR.TPEPSLE LGFFRMKAFS
      PTR  SGPDHGWTST QVLASAALGA LFLGAFVFWE LR.IPYPMLD MHFFRNQRFI
     TCMA  .APAWGWGDL RTLGFLAAAV LAFAGFTLRE SR.ATEPLMP LAMFRSVPLS
      MMR  .GPQLGWTAG PVLTAYAVAV TAAALLALRE HR.VTNPVMP WQLFRGPGFT
```

```
          301                                                       350
  OTRB    LCSAVNFTIG VGIFGTVTTL PLFLQMVQGR TPTQAGLVVI PFMLGTIASQ
  CTRB    MTTAVNFTIG VGIFGTVSTL PLFLQLVQGR SATVAGLVII PVMTGAIVSQ
SVORF4    VYNGVNVLVG AAVFGALSVL PLYLQMVKGL SPTQAGLMML PQTLGIVVAG
  CMCT    WGNLAMLLAG ASQVPVWFFL TLSMQHVLGY SAAQAGLGFV PHALVMLVVG
  PUR8    GSFLVVGLPQ IGLFGLFLFL TYYLQGILDY SPVLTGVAFL PLGLGIAVGS
 ACTVA    TAVAAVGFVS FAMMGFLFFS AFYLQSVRGY TPLQAGGCTV ALAVANVVCG
   PTR    GAVAGAILVA FGMTGSLFLL TQHLQFVLGY GALDAGLRTA PLALTVVALN
  TCMA    AGTVLMVLMA FSFIGGLFFV TFYLQNVHGM SPVESGVHLL PLTGMMIVGA
   MMR    GANLVGFLFN FALFGSTFML GLYFQHARGA TPFQAGLELL PMTIFFPVAN

          351                                                       400
  OTRB    MVS.GKLIAS SGRFKKLAIV GLGSMAGALL AMATTGATTP MWGIVLI...
  CTRB    TIC.AKIIKK WNRYKKPAIV GLGSMAGALL SLSAAGADTP LAVIVVI...
SVORF4    RIA.RPYVTR TGRYKAVLLS GLVLMVVATF WFGVLTADTS LWQTGVA...
  CMCT    LR.VVPWLMR HVQARVLIAA GAAIGALGFW WQSLLTPDS. .AYLGGILGP
  PUR8    SL.IAARLLP RTRPRTLIVG ALLAAAAGMA LLTRLEPDTP QVYLTHLLPA
 ACTVA    ..PLSTVLVR SIGPRNVCAA GMLAVTASLC GVTFVTQHAP VWLILVLF..
   PTR    LTGMGARLLR IFGTPVTIAV GMALVSGGLA AIAVLGAGGY NGMLLGL...
  TCMA    P..VSGIVIS RFGPGGPLVV GMLLTAASLW GMSTLEADSG MGITSLWF..
   MMR    I..VYARISA RFSNGTLLTA FLLLAGAASL SMVTITASTP YWVVAVAV..

          401                                                       450
  OTRB    VLWLGVGIGL SQTVITSPMQ NSAPKSQLGV ANGASGLCRQ IGGSTGIAVL
  CTRB    AAWLGFGIGL SQTVITLAIQ SSAPKSELGV ANAASGLFRQ LGGTSGAAVF
SVORF4    AGVMGFGIGL CWQVMLIAIQ TGVAPQYMGA GMGSFTFFRQ IGGTGRHRVF
  CMCT    AVLISIGGGL VGTPLARTVT SGVGPLDAGA ASGLMNTTRQ FGGAFGLAVL
  PUR8    QILIGLGIGC MMMPAMHTAT ARVAPHEAGA AAAVVNSAQQ VGGALGVALL
 ACTVA    .AALGAGVAC VMPTAAVSIM NAIPREKAGV ASAMNNTVRQ LGGALGVAVL
   PTR    .VVMGAGVAL SMPAMANAIM SAIPPEKAGV GAGINGTLAE FGNGLGVAVL
  TCMA    .VLLGLGLAP VMVGTTDVIV SNAPAELAGV AGGLQQSAMQ VGGSLGTAVL
   MMR    .GVANIGAGI ISPGMTAALV DAAGPENANV AGSVLNANRQ IGSLVGIAA
                                                                  LgxxxGxavx

          451                                                       500
  OTRB    FSVMFAVALG RLA....DLL HTPRYERLLT DPAITGDPAN HRFLDMAESG
  CTRB    MSVLFGVAAG RLD....GAD PDEAVRRALS DPDSTGG.LS ASAVDAFTSG
SVORF4    LSMFFGAVGE KVAAAYRGAA SDPAYTAAVN DPAVTGQAAN KVLL.....G
  CMCT    LTVTGSGTSG SPAELASHYG DAFVGIAVFM LAIAVLTPVL PALARSTPPG
  PUR8    NTVSTGATAA YLADHGTSPA .ATVDGTVHG YTVAIAFAVG VLLLTAVLAW
 ACTVA    GSLMGAAYRR GIEDELAVLP PSARHQAGES LDATLLAATR LGESGLVGPA
   PTR    GAVLNSRF.. ......AALV PAA..VGATS LPAALAAADD AGERARISDA
  TCMA    GVLMASRVGD VFPDKWAEAN LPRVGPREAA AIEDAAEVGA VPPAGTLPGR
   MMR    GVVLHSTSD. ..WDHGAAIS FLAVGLAYLL GGLSAWRLIA RPERRSAVTA
          gxl (Motif F)

          501                                                       550
  OTRB    QGAGINLDDT SLLNGIDARL .MQPVTDSFA HGFHIMFLAG GVVLLAGFVM
  CTRB    FDTMFLVGGL ILAVGFLLTF PLRELRDEE~ ~~~~~~~~~~ ~~~~~~~~~~
SVORF4    PGRRVGLDNS SFLDHADPRL A.KPFLEGLA EAMQTVFVVS GVMLLIALVL
  CMCT    VIHVSPVAR~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  PUR8    VLIDSRTEAA DETGSASVTP ARPR~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 ACTVA    RQAFLDAMHL AAGAAAAVAL VGALAVLRWL PSSVTTPTPP AGAVPGREHS
   PTR    FASGLETSQL V..GAVAVLA GGLLAALLLR RAERADPPAK A~~~~~~~~~
  TCMA    HAGTLSEVVH SSFISGMGLA FTVAGAVALV AAAVALFTRK AEPDERAPEE
   MMR    AT~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
```

```
          551                                                    585
  OTRB    TWFLRELQEE TAPEEERPAE SGAGAKNGPL PASDA
  CTRB    ---------- ---------- ---------- -----
SVORF4    AAVPKEKPPR AGRRFAGRRK ETGKQPAAKK -----
  CMCT    ---------- ---------- ---------- -----
  PUR8    ---------- ---------- ---------- -----
 ACTVA    DHLKVQGS-- ---------- ---------- -----
   PTR    ---------- ---------- ---------- -----
  TCMA    FPVPASTAGR G--------- ---------- -----
   MMR    ---------- ---------- ---------- -----
```

Fig 3.10

Alignments between a series of Streptomycete 14 TMH drug, exporters. CmcT, the cephamycin transporter from *Nocardia lactamdurans* is also included.

Conserved motifs of the MFS family are denoted in bold type below the PileUp. Upper case residues denote the amino acid is conserved in over 70% of proteins, lower case letters denote conservation of the residue in over 40% of proteins. (Motif concensus sequences taken from Paulsen & Skurray, 1996) Residues in bold type within the PileUp diagram are motifs found to be well conserved amongst the drug transporters of *Streptomyces* spp.

For accession numbers and references for these proteins, see Table 3.1.

```
          1                                                      50
TetK   MFSLYKKFKG LFYSVLFWLC ILSFFSVLNE MVLNVSLPDI ANHFNTTPGI
TetL   MNTSYSQSNL RHNQILIWLC ILSFFSVLNE MVLNVSLPDI ANDFNKPPAS
OtrB   ~~~~~~~~~~ ~~~~~MSGLL LAVFLAALDQ TVIATAMRTI ADDLHGQTEQ

          51                                                    100
TetK   TNWVNTAYML TFSIGTAVYG KLSDYINIKK LLIIGISLSC LGSLIAFIGH
TetL   TNWVNTAFML TFSIGTAVYG KLSDQLGIKR LLLFGIIINC FGSVIGFVGH
OtrB   A.WATTGYLI ASVLAMPFYG KLSDIYGRKP MYLISIVVFI GGSVLCGTAG

          101                                                   150
TetK   NHFFILIFGR LVQGVGSAAF PSLIMVVVAR NITRKKQGKA FGFIGSIVAL
TetL   SFFSLLIMAR FIQGAGAAAF PALVMVVVAR YIPKENRGKA FGLIGSIVAM
OtrB   SMWELALF.R AVQGLGGGGL MSLPTAVVAD LAPVRERGRY FAFLQMAWVV

          151                                                   200
TetK   GEGLGPSIGG IIAHY..... ...IHWSYLL ILPMITIVTI PFLIKVMVPG
TetL   GEGVGPAIGG MIAHY..... ...IHWSYLL LIPMITIITV PFLMKLLKKE
OtrB   ASVAGPLAGG FFAEAGQVFG IDGWRWVFLL NVPLGLLALV TVRKALNLPH

          201                                                   250
TetK   KSTKNTLDIV GIVLMSISII CFML..... ...FTTNYNW TFLILFTIFF
TetL   VRIKGHFDIK GIILMSVGIV FFML..... ...FTTSYSI SFLIVSVLSF
OtrB   ERREHRMDVL GAAALALFLV PLLIVAEQGR TWGWGSPAAL ALFALGAAGL

          251                                                   300
TetK   VIFIK.HISR VSNPFINPKL GKNIPFMLGL FSGGLIFSIV AGFISMVPYM
TetL   LIFVK.HIRK VTDPFVDPGL GKNIPFMIGV LCGGIIFGTV AGFVSMVPYM
OtrB   AVFIPVELRR GDEAILPLGL FRRGSIALCS AVNFTIGVGI FGTVTTLPLF

          301                                                   350
TetK   MKTIYHVNVA TIGNSVIFPG TMSVIVFGYF GGFLVDRKGS LFVF.ILGSL
TetL   MKDVHQLSTA EIGSVIIFPG TMSVIIFGYI GGILVDRRGP LYVL.NIGVT
OtrB   LQMVQGRTPT QAG.LVVIPF MLGTIASQMV SGKLIASSGR FKKLAIVGLG

          351                                                   400
TetK   SISISFLTIA FFVEFS.MWL TTFMFIFVMG GLSFTKTVIS KIVSSSLSEE
TetL   FLSVSFLTAS FLLETT.SWF MTIIIVFVLG GLLFTKTVIS TIVSSSLKQQ
OtrB   SMAGALLAMA TTGATTPMWG IVLIVLWLGV GIGLSQTVIT SPMQNSAPKS

          401                                                   450
TetK   EVASGMSLLN FTSFLSEGTG IAIVGGLLSL QLINRKLVLE FINYSSGVYS
TetL   EAGAGMSLLN FTSFLSEGTG IAIVGGLLSI PLLDQRLLPM EVDQSTYLYS
OtrB   QLGVANGASG LCRQIGGSTG IAVLFSVMFA VALGRLADLL H....TPRYE

          451                                                   500
TetK   NILVAMAILI ILCCLLTIIV FKRSEKQFE~ ~~~~~~~~~~ ~~~~~~~~~~
TetL   NLLLLFSGII VISWLVTLNV YKHSQRDF~~ ~~~~~~~~~~ ~~~~~~~~~~
OtrB   RLLTDPAITG DPANHRFLDM AESGQGAGIN LDDTSLLNGI DARLMQPVTD

          501                                                   550
TetK   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
TetL   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
OtrB   SFAHGFHIMF LAGGVVLLAG FVMTWFLREL QEETAPEEER PAESGAGAKN
```

```
551
TetK  ~~~~~~~~~
TetL  ~~~~~~~~~
OtrB  GPLPASDA
```

Fig 3.11

Alignment of OtrB with TetK and TetL, two tetracycline exporters from Gram-positive bacteria. For accession numbers and references for these proteins, see Table 3.1.

```
                1                                                   50
TetA(A)     MKPNIPLIVI LSTVALDAVG IGLIMPVLPG LLRDLVHSND VTAHYGILLA
TetA(C)     MKSNNALIVI LGTVTLDAVG IGLVMPVLPG LLRDIVHSDS IASHYGVLLA
TetA(G)     --VRSSAIIA LLIVGLDAMG LGLIMPVLPT LLRELVPAEQ VAGHYGALLS
TetA(B)     --MNSSTKIA LVITLLDAMG IGLIMPVLPT LLREFIASED IANHFGVLLA
TetA(D)     --MNKPAVIA LVITLLDAMG IGLIMPVLPS LLREYLPEAD VANHYGILLA
TetA(E)     --MNRTVMMA LVIIFLDAMG IGIIMPVLPA LLREFVGKAN VAENYGVLLA
OtrB        ---------- MSGLLLAVFL AALDQTVIAT AMRTIADDLH GQTEQAWATT

               51                                                  100
TetA(A)     LYALVQFACA PVLGALSDRF GRRPILLVSL AGATVDYAIM ATAPFLWVLY
TetA(C)     LYALMQFLCA PVLGALSDRF GRRPVLLASL LGATIDYAIM ATTPVLWILY
TetA(G)     LYALMQVVFA PMLGQLSDSY GRRPVLLASL AGAAVDYTIM ASAPVLWVLY
TetA(B)     LYALMQVIFA PWLGKMSDRF GRRPVLLLSL IGASLDYLLL AFSSALWMLY
TetA(D)     LYAVMQVCFA PLLGRWSDKL GRRPVLLLSL AGAAFDYTLL ALSNVLWMLY
TetA(E)     LYAMMQVIFA PLLGRWSDRI GRRPVLLLSL LGATLDYALM ATASVVWVLY
OtrB        GYLIASVLAM PFYGKLSDIY GRKPMYLISI VVFIGGSVLC GTAGSMWELA

              101                                                 150
TetA(A)     IGRIVAGI.T GATGAVAGAY IADITDGDER ARHFGFMSAC FGFGMVAGPV
TetA(C)     AGRIVAGI.T GATGAVAGAY IADITDGEDR ARHFGLMSAC FGVGMVAGPV
TetA(G)     IGRLVSGV.T GATGAVAAST IADSTGEGSR ARWFGYMGAC YGAGMIAGPA
TetA(B)     LGRLLSGI.T GATGAVAASV IADTTSASQR VKWFGWLGAS FGLGLIAGPI
TetA(D)     LGRIISGI.T GATGAVAASV VADSTAVSER TAWFGRLGAA FGAGLIAGPA
TetA(E)     LGRLIAGI.T GATGAVAAST IADVTPEESR THWFGMMGAC FGGGMIAGPV
OtrB        LFRAVQGLGG GGLMSLPTAV VADLAPVRER GRYFAFLQMA WVVASVAGPL

              151                                                 200
TetA(A)     LGGLMGGFSP HAPFFAAAAL NGLNFLTGCF LLPESHKGER RPLRREALNP
TetA(C)     AGGLLGAISL HAPFLAAAVL NGLNLLLGCF LMQESHKGER RPMPLRAFNP
TetA(G)     LGGMLGGISA HAPFIAAALL NGFAFLLACI FLKETHHSHG GTGKPVRIKP
TetA(B)     IGGFAGEISP HSPFFIAALL NIVTFLVVMF WFRETKNTRD NTDTEVGVET
TetA(D)     IGGLAGDISP HLPFVIAAIL NACTFLMVFF IFKPAVQTEE KPADE..KQE
TetA(E)     IGGFAGQLSV QAPFMFAAAI NGLAFLVSLF ILHETHNANQ VSDELKNETI
OtrB        AGGFFAEAGQ VFGIDGWRWV FLLNVPLGLL ALVTVRKALN LPHERREHRM

              201                                                 250
TetA(A)     LSFVRWARGM .TVVAALMAV FFIMQLVGQV PAALWVIFGE DRFHWDATTI
TetA(C)     VSSFRWARGM .TIVAALMTV FFIMQLVGQV PAALWVIFGE DRFRWSATMI
TetA(G)     FVLLRLDDAL .RGLGALFAV FFIIQLIGQV PAALWVIYGE DRFQWNTATV
TetA(B)     QSNSVYI.TL FKTMPILLII YFSAQLIGQI PATVWVLFTE NRFGWNSMMV
TetA(D)     SAGISFI.TL LKPLALLLFV FFTAQLIGQI PATVWVLFTE SRFAWDSAAV
TetA(E)     NETTSSIREM ISPLSGLLVV FFIIQLIGQI PATLWVLFGE ERFAWDGVMV
OtrB        DVLGAAALAL F..LVPLLIV AEQGRTWGWG SPAALALFAL GAAGL.AVFI

              251                                                 300
TetA(A)     GISLAAFGIL HSLAQAMITG PVAARLGERR ALMLGMIADG TGYILLAFAT
TetA(C)     GLSLAVFGIL HALAQAFVTG PATKRFGEKQ AIIAGMAADA LGYVLLAFAT
TetA(G)     GLSLAAFGAT HAIFQAFVTG PLSSRLGERR TLLFGMAADG TGFVLLAFAT
TetA(B)     GFSLAGLGLL HSVFQAFVAG RIATKWGEKT AVLLGFIADS SAFAFLAFIS
TetA(D)     GFSLAGLGAM HALFQAVVAG ALAKRLSEKT IIFAGFIADA TAFLLMSAIT
TetA(E)     GVSLAVFGLT HALFQGLAAG FIAKHLGERK AIAVGILADG CGLFLLAVIT
OtrB        PVELRRGDEA ILPLGLFRRG SIALCSAVNF TIGVGIFGTV TTLPLFLQMV

              301                                                 350
TetA(A)     RGWMAFPIMV LLASGGIGMP ALQAMLSRQV DEERQGQLQG SLAALTSLTS
TetA(C)     RGWMAFPIMI LLASGGIGMP ALQAMLSRQV DDDHQGQLQG SLAALTSLTS
TetA(G)     QGWMVFPILL LLAAGGVGMP ALQAMLSNNV SSNKQGALQG TLTSLTNLSS
TetA(B)     EGWLVFPVLI LLAGGGIALP ALQGVMSIQT KSHQQGALQG LLVSLTNATG
TetA(D)     SGWMVYPVLI LLAGGGIALP ALQGIISAGA SAANQGKLQG VLVSLTNLTG
TetA(E)     QSWMVWPVLL LLACGGITLP ALQGIISVRV GQVAQGQLQG VLTSLTHLTA
OtrB        QGRTPTQAGL VVIPFMLGTI ASQMVSGKLI ASSGRFK... .KLAIVGLGS

                351                                                 400
TetA(A)     IVGPLLFTAI YAASITTWNG WAWIAGAALY LLCLPALRRG LWSGAGQRAD
TetA(C)     ITGPLIVTAI YAASASTWNG LAWIVGAALY LVCLPALRRG AWSRATST--
TetA(G)     IAGPLGFTAL YSATAGAWNG WVWIVGAILY LICLPILRRP FATSLV----
TetA(B)     VIGPLLFAVI YNHSLPIWDG WIWIIGLAFY CIIILLSMTF MLTPQAQGSK
TetA(D)     VAGPLLFAFI FSQTQQSADG TVWLIGTALY GLLLAICLLI RKPAPVAATC
TetA(E)     VIGPLVFAFL YSATRETWNG WVWIIGCGLY VVALIILRFF HPGRVIHPIN
OtrB        MAGALLAMAT TGATTPMWGI VLIVLWLGVG IGLSQTVITS PMQNSAPKSQ
```

```
                401                                                        450
TetA(A)         R---------- ---------- ---------- ---------- ----------
TetA(C)         ---------- ---------- ---------- ---------- ----------
TetA(G)         ---------- ---------- ---------- ---------- ----------
TetA(B)         QETSA----- ---------- ---------- ---------- ----------
TetA(D)         ---------- ---------- ---------- ---------- ----------
TetA(E)         KSDVQQRI-- ---------- ---------- ---------- ----------
OtrB            LGVANGASGL CRQIGGSTGI AVLFSVMFAV ALGRLADLLH TPRYERLLTD

                451                                                        500
TetA(A)         ---------- ---------- ---------- ---------- ----------
TetA(G)         ---------- ---------- ---------- ---------- ----------
TetA(C)         ---------- ---------- ---------- ---------- ----------
TetA(B)         ---------- ---------- ---------- ---------- ----------
TetA(D)         ---------- ---------- ---------- ---------- ----------
TetA(E)         ---------- ---------- ---------- ---------- ----------
OtrB            PAITGDPANH RFLDMAESGQ GAGINLDDTS LLNGIDARLM QPVTDSFAHG

                501                                                        550
TetA(A)         ---------- ---------- ---------- ---------- ----------
TetA(G)         ---------- ---------- ---------- ---------- ----------
TetA(C)         ---------- ---------- ---------- ---------- ----------
TetA(B)         ---------- ---------- ---------- ---------- ----------
TetA(D)         ---------- ---------- ---------- ---------- ----------
TetA(E)         ---------- ---------- ---------- ---------- ----------
OtrB            FHIMFLAGGV VLLAGFVMTW FLRELQEETA PEEERPAESG AGAKNGPLPA

                551
TetA(A)         ---
TetA(G)         ---
TetA(C)         ---
TetA(B)         ---
TetA(D)         ---
TetA(E)         ---
OtrB            SDA
```

Fig 3.12

Alignment of OtrB with 12 TMH Tetracycline resistance proteins from Gram-negative bacteria. The 12 TMH's of the TetA group are marked, residues in bold type indicate conservation between OtrB and the TetA group. Underlined regions denote the putative 14 TMH's of OtrB
For accession numbers and references for these proteins see Table 3.1.

PILEUP for OtrB and Tetracycline resistance proteins with 14 TMH's from other Gram-positive bacteria. It should be noted that all tet exporters from Gram-positives are so far thought to be made up from 14 TMH's whilst those from Gram-negatives seem to consist of 12 TMH's. Fig 3.11 shows that although OtrB and the TetK, TetL proteins both contain 14 TMH's , TetK and TetL are much more closely related to each other than they are to OtrB. OtrB shows much more homology with the proteins in the previous PileUp.

OtrB and some Gram-negative TET exporters are compared in Fig 3.12. The proteins from Gram-negatives show very high homology to each other, the positions of the 12 TMH's of the TetA group are marked, it can be seen that all but two of the residues conserved between OtrB and the TetA group occurs in the first 6 helices of the proteins. The remaining two conserved residues occur within TMH10 of both proteins. It is possible that this motif, present as ALQ in the TetA group and as ASQ in OtrB are linked to substrate recognition as there is no conservation of these residues within the Streptomycete transporter group whose members transport varied substrates. However if this were the case, the motif might be expected also to appear in the TetK and TetL proteins, no such motif can be observed.

### 3.6.3 Conserved motifs

Within the members of the MFS, certain conserved regions known as motifs have been identified (Henderson & Maiden, 1990; Griffith *et al* ,1992; Marger & Saier, 1993; Paulsen & Skurray, 1993). Some of these motifs are conserved throughout the entire superfamily (motifs A and B) whilst another is found only in the 12 and 14 TMH families (motif C and D, although there is some variation in motif D between the 2 families). Other motifs were identified as being exclusive to either the 12 TMH (motif G) or the 14 TMH families (motifs E and F) (Paulsen & Skurray, 1993). In a review by Paulsen *et al* (1996) these motifs have been refined using extensive multiple sequence alignments. Also presented in that review is an additional motif (motif H). This is present in the 14 TMH family but can also be seen in a somewhat diverged format in some 12 TMH transporters. The consensus sequences of motifs presented here are taken from Paulsen *et al*, (1996).

Many of these motifs can be clearly identified in OtrB, ( Fig 3.10).

Motif A

**GxLaDrxGRkxxxL**

Initially identified as a motif from certain sugar and tetracycline transporters, by Rouch *et al* (1990) as GXXSDR/KSGR/ER/K and by Henderson & Maiden (1990) as RXGRR, motif A is located in the cytoplasmic loop between the second and third TMH's. The motif is conserved in all four families of the MFS, has been postulated to act as a substrate entrance gate (Yamaguchi *et al*, 1990) and is thought to form a β turn of the protein. In OtrB the motif has been totally retained and can be seen as GkLsDiyGRKpmyL where upper case letters denote residues in agreement with the motif sequence (Fig 3.10). The residues of motif A, which is the most highly conserved loop in the N-terminal half of the MFS proteins, have been well-characterised and the two glycine residues are essential for function. Studies of TetB have led to the postulation that they are involved with a gating function (Yamaguchi *et al*, 1990; Yamaguchi *et al*, 1992a). It is also possible that they may act to promote global conformational changes in the protein that enable the substrate to cross the membrane. The aspartic acid residue and arginine residue are also essential most likely playing a structural role, (Jessen-Marshall *et al*, 1995). This second postulation is supported by the fact that second site suppressor mutations which retain function have been identified and these mutations occur at various sites all over the protein (Yamaguchi *et al*, 1995, Someya *et al* 1995). The SD (serine, aspartate) motif within motif A shows absolute conservation in tetracycline transporters of the 12 and 14 TMH families (Chopra, 1986). This observation, made more than 10 years ago, has been borne out in sequences published since that time.

Motif C

**gxxxGPxxGGxl**

This motif , located in TMH 5 of the 12 & 14 TMH antiporter families is not found in the symport proteins of the other MFS families. This suggests a role in the linking of proton translocation to the antiport but not the symport of a substrate (Griffith *et al*, 1992; Paulsen & Skurray, 1993). Mutagenesis studies of the glycine residue at position 5 in the motif in TetC have been carried out and it is suggested by Varela *et al* (1995) that the residues of motif C form a bend in TMH5. This invites the

speculation that motif C may determine the positioning of the unoccupied substrate binding site and therefore control transport direction of the substrate.

In OtrB, motif C is conserved as avsaGPlaGG (Fig 3.10).

Motif Di

**lDxTvxnvAIP**

As with motif C, motif D is only found within the 12 & 14 TMH families and it is slightly different in each of these. The motif is known as Di in the 14 TMH proteins and Dii in the 12 TMH proteins. The conserved residues which make up the motif occur within the first TMH of the transporter. The role of this motif has not as yet been investigated. Motif Di appears in OtrB as, LDqTViatA. (Fig 3.10).

Motifs E and F

**DxxGxxL and IgxxxGxavxGxl**

Motifs E and F are only conserved among members of the 14 TMH family. At this time there are no published data regarding their function although it should be noted that within motif E there is a highly conserved intramembranous charged residue, namely the aspartate at position 1 of the motif. Motif E sits within the TMH7 whilst motif F is positioned within TMH13.

In OtrB, both motifs have been conserved to some extent. Motif E is seen as DvlGaaa and motif F as iGgstGiAVlfsv (Fig 3.10)

Motif H

**WxwxFLINPig**

This motif occurs in the sixth helix of the 14 TMH exporters but can also be recognised in a divergent form in some of the members of the 12 TMH family. Like motifs E and F, there is little known of its functional significance. In OtrB the motif is very well conserved and appears as WcWvFLINVPIG. (Fig 3.10)

There are a number of other well conserved residues among the 14 TMH family, identified by Paulsen *et al.*, (1996) which are not present as part of motifs, the OtrB protein contains many of these conserved residues. (Fig 3.10).

PILEUP analysis has also shown up two regions that are extremely well conserved amongst the Streptmoycete drug transporters. A conserved Glu (E) residue at the end of the eighth TMH and a conserved Leu-Gln (LQ) motif in the cytoplasmic loop connecting TMH's 9 and 10. The conserved E is present in 8 out of the 8 *Streptomyces* drug transporters analysed and the LQ motif present in 7 of the 8. In Mmr, the methylenmycin transporter from *Streptomyces coelicolor*, the L is replaced by an F (Phe), both are nonpolar amino acids.

### 3.6.4 Relationship to other MFS members

From the results of the PILEUP analysis, it can be seen that OtrB shares more similarity with other streptomycete transport proteins than it does with tetracycline pumps from other bacteria, be they Gram-positive or negative. Although the OtrB protein is more closely related in its topology to the Gram positive than the Gram negative Tet exporters, the results of these PILEUP's show that more residues are conserved between TetA(B) (Gm-ve) and OtrB than between TetK and OtrB.

A different method of assessing "relatedness" of proteins is to use the TblastN search from the NCBI (National Center for Biotechnology Information). This tool enables the peptide sequence of a query protein to be compared against various databanks of nucleotide sequences and outputs a list of the most closely related peptides by virtue of the number of conserved residues between the two. A table of relatedness (Table 3.2) was compared using data from such an analysis on the *otrB* nucleotide sequence When the OtrB sequence was looked at in this way, the two most closely related peptides were, CtrB, the chlortetracycline determinant from *S. aureofaciens* and the product of *SvOrf4* from the granatacin cluster of *S. violaceoruber*, for which the substrate is probably granatacin The third most closely related peptide was not from a Streptomycete nor was it a single drug transporter but the macronuclear DNA from *Bacillus subtilis* that encodes a protein responsible for multidrug resistance in that organism.

Table 3.1, Shows names, substrates, sources, accession numbers and references of proteins used in PILEUP analyses.

| Protein | Organism | substrate | Accession no. | References |
|---|---|---|---|---|
| OtrB | *Streptomyces rimosus* | OTC | AF061335 | This work |
| CtrB | *Streptomyces aureofaciens* | CTC | GB I32939 | Ryan *et al*, 1997 |
| SvOrf4 | *Streptomyces. violaceoruber* | Granaticin (?) | GB L37334 | Bechthold *et al* unpublished. cited in Paulsen *et al*, 1996 |
| CmcT | *Nocardia lactamdurans* | Cephamycin | SWQ04733 | Coque, *et al* 1993 |
| Pur8 | *Streptomyces. lipmanii* | puromycin, N-acetyl puromycin | GBX76855 | Tercero *et al*, 1993 |
| ActVa | *Streptomyces. coelicolor* | Actinorhodin | GBX58833 | Cabellero *et al*, 1991b |
| Ptr | *Streptomyces. pristinaespiralis* | Pristinamycin I, II, rifampin | GBX84072 | Blanc *et al*, 1995 |
| TcmA | *Streptomyces. glaucesns* | Tetracenomycin C | GBM80674 | Guilfoile & Hutchinson, 1992 |
| Mmr | *Streptomyces. coelicolor* | Methylenomycin A | GBM18263 | Neal & Chater, 1987 |
| | | | | |
| TetK | *Staph. aureus* | TET | EMM16217 | Noguchi *et al*, 1986 |
| TetL | *Bacillus Stearothermophillus* | TET | SWP07561 | Hoshino *et al*, 1985 |
| | | | | |
| TetA (A) | *Escherichia coli* | TET | EMX00006 | Waters *et al*, 1983 |
| TetA (B) | *Escherichia coli* | TET | EMJ01830 | Nguyen *et al*, 1983 |
| TetA (C) | *Pseudomonas aeruginosa* | TET | EMJ01749 | Peden, 1983 |
| TetA (D) | *Salmonella ordonezi* | TET | EMX65876 | Allard *et al*, 1993 |
| TetA (E) | *Escherichia coli* | TET | SWQ07282 | Allard & Bertrand, 1993 |
| TetA (G) | *Vibrio anguilarium* | TET | GBS52437 | Zhao & Aoki, 1992 |
| TetA (H) | *Pateurella maltocida* | TE | GBU00792 | Hansen *et al*, 1993 |

As the nucleotide databases are continually updated, slightly lower scoring peptides with similarity to OtrB are often displaced down the list by new additions to the growing family of MFS members. The data presented here were obtained by comparing *otrB* to all non redundant GenBank translations plus the PDB, Swissprot PIR protein databases using the basic (un-gapped) blast search, TblastN on 13.10.97. The fourth most significant alignment was with a presumed multidrug transporter protein from *Mycobacterium tuberculosis* cosmid Y180 and the next but one highest scorer was the ErmB homologue from *B. subtilis*. None of the remaining members of the "top ten" were from bacteria but were all from yeasts. Indeed, 11 out of the top 20 scores are proteins from these eukaryotic organisms.

Other streptomycete exporter genes are brought up by the search . These are the actinorhodin exporter *actIIOrf4* from *S. coelicolor* at position 17, *tcmA* from *S. glaucesens* at position 22, *actVa* also from *S. coelicolor* at 29, *lmrA* which encodes lincomycin resistance in *S. lincolnensis* at position 39. Finally the gene encoding *S. alboniger pur8* protein, a puromycin exporter, which is even less well related was at position 83.

It is pertinent to note that other tetracycline resistance determinants come fairly far down the list. Only TetA(B) from Tn*10* is present in the first 100, at position 65. This is interesting as the TetK protein from *S. aureus* does not appear in the top 100 most closely-related proteins even though it is made up from 14 TMH's (Guay et al, 1993a, Ginn *et al*, 1997) as is OtrB and is encoded by a Gram-positive organism whilst TetA(B) is a member of the 12TMH subfamily and is encoded by a Gram-negative organism, *E.coli*. These findings will be discussed further in section 3.7.

Most of the similarities between closely-related members of MFS families and subfamilies occur in the N-terminal portion of the protein rather than in the C-terminal portion (Griffith *et al* 1992, Marger & Saier, 1993, Paulsen & Skurray, 1993, Rouch *et al*, 1990). It has been suggested therefore that the N-terminal halves of these proteins are primarily concerned with energisation of transport whilst the more diverged C-terminal regions are involved in recognition of the wide variety of substrates transported.

Table 3.2

20 most high scoring nucleotide coding sequences to *otrB*. Data from TblastN analysis performed 13/10/97.

| | Accession No. | Organism | gene / protein | Score |
|---|---|---|---|---|
| **1** | gbD38215 | *Streptomyces aureofaciens* | *ctrB* | 1057 |
| **2** | gbL37334 | *Streptomyces violaceoruber* | *svOrf4* | 352 |
| **3** | dbjD50098 | *Bacillus subtillus* | MDR * | 331 |
| **4** | embZ97193 | *Mycobacterium tuberculosis* | MDR* | 289 |
| **5** | embZ68144 | *Schizosaccharomyces pombe* | MDR* | 342 |
| **6** | dbjD50453 | *Bacillus subtillis* | ErmB homologue | 311 |
| **7** | embZ36162 | *Saccharomyces cerevisiae* | unknown function | 229 |
| **8** | gbZ49259 | *Saccharomyces cerevisiae* | unknown function | 234 |
| **9** | gbL11640 | *Saccharomyces cerevisiae* | unknown function | 229 |
| **10** | gbL24961 | Yeast | MDR* | 229 |
| **11** | embX77765 | *Saccharomyces cerevisiae* | NOR1 | 229 |
| **12** | gbU02077 | *Saccharomyces cerevisiae* | unknown function | 229 |
| **13** | gbU25841 | *Saccharomyces cerevisiae* | unknown function | 229 |
| **14** | embZ28330 | *Saccharomyces cerevisiae* | unknown function | 209 |
| **15** | embZ80226 | *Mycobacterium tuberculosis* | unknown function | 191 |
| **16** | embZ73009 | *Saccharomyces cerevisiae* | unknown function | 216 |
| **17** | gbM64683 | *Streptomyces ceolicolor* | *ActII* Orf4 | 159 |
| **18** | gbU32771 | *Haemophilus influenzae* | unknown function | 183 |
| **19** | embX87941 | *Saccharomyces cerevisiae* | unknown function | 216 |
| **20** | gbU51164 | *Propionibacterium freudenreichi* | unknown function | 152 |

* MDR denotes gene encoding putative multidrug resistance protein.

## 3.7 Discussion.

### 3.7.1 Sequencing results.

The nature of the *Streptomyces* genome (>70% G+C), has the consequence that manual DNA sequencing of S*treptomyces* genes can prove difficult. Compressions of bands in the polyacrylamide gel occur due to the formation of secondary structures by the DNA GC clamps and hairpin loops occur often with G+C rich DNA. This may offer an explanation for the discrepancies between the *otrB* sequences presented by MacGregor-Pryde (1995), Reynes *et al*, (1988) and that presented in this thesis, which was obtained using automated DNA sequencing methods. The advent of automated sequencing introduced computerised technology for the analysis of sequencing gels. This, coupled with the increase in the length of DNA that can be sequenced in one run and the reduction in the time taken to obtain the final results, means that multiple runs can be performed, analysed and compared with each other in shorter periods of time and with considerably less expense than traditional manual radiolabelling methods.

### 3.7.2 OtrB within theMajor Facilitator Superfamily

The volume of information on the MFS and its members has grown considerably in the last few years which has enabled a thorough analysis of the OtrB peptide sequence with respect to other like proteins.

The analysis undertaken in this chapter gives much evidence to support the designation of OtrB in the 14 TMH subfamily of the Major Facilitator Superfamily. A phylogenetic tree, taken from Paulsen *et al,* (1996) shows that within the 14 TMH group, several distinct clusters can be discerned (Fig 1.13). OtrB (shown as Tet347) is present within a small cluster containing itself and SvOrf4. CtrB is also presumed to be a member of this cluster. As previously mentioned, most of the similarity between these transporters is in the N-terminal region, and is believed to indicate that this region is involved with energisation of membrane transport. Through the work of Yamaguchi and colleagues, the TetA(B) protein of *E. coli* has been studied extensively by mutation analysis, and essential and functionally important residues identified ( 1.6.2 ). Although TetA(B) is not the most closely related protein to OtrB, it is interesting to investigate which, if any, of these residues are conserved.

Some of the motifs are conserved fairly well across the whole MFS superfamily and are believed to be functionally important with respect to transport. There are also some residues identified by Yamaguchi *et al* which are conserved across the Gram negative Tet exporters. Some of these observations do not apply to other families such as OtrB, or TetK/TetL but others can be applied. The aspartate (D) residue within motif A (located between helix 2 and 3) is conserved across the whole MFS and the preceding residue is often an alanine (A) residue. However, in the Gram-negative Tet exporters this residue is conserved as a serine (S). This is also the case in the Gram positive Tet exporters, TetK, TetL and OtrB. Yamaguchi *et al* have proposed that these two residues may act as a substrate gate (Yamaguchi *et al.*, 1990) and not as a substrate binding site as originally suggested by Chopra (1986). The work of the Yamaguchi group has also shown that the other serine residue in this loop towards the N-terminus is also well conserved among Gram-negative Tet exporters and is thought to be important in formation of a gating channel in TMH 3 along with S91 and D84 (Yamaguchi *et al*, 1992a).

A histidine residue H257, located in TMH 8 of TetA(B) was demonstrated by Yamaguchi *et al* (1991) to be essential for proton translocation. When mutated to alanine (A) or asparagine (N) the protein loses its antiporter activity and becomes a uniporter. As OtrB has 14 TMH's this observation cannot be extended to it.

It would be more interesting to look at those residues believed to be involved in substrate recognition in TetA(B). The aspartate residue D84 was thought to play some role in substrate recognition in TetA(B) along with D15 and D285 by virtue of their negative charges (Yamaguchi *et al*, 1992b) as it was conserved in all the Gram negative Tet proteins , but as this aspartic acid has now been shown to be part of a motif observed in all families of the MFS, the hypothesis that it is involved in substrate specificity seems somewhat outdated. Two residues have been implicated in substrate recognition. Guay *et al* (1994) showed that mutations in two codons of TetA(B), namely codon W231 and codon L308, resulted in increased resistance to the novel tetracycline 9-(dimethylgycylamido)-minocycline, (DMG-Mino) with a concurrent decrease in tetracycline resistance. Both amino acids altered in DMG-Mino resistant mutants are located in transmembrane regions of the protein, W231 in TMH9 and L308 in TMH11. It is proposed that the side chains at positions 231 and

308 point into a channel that is involved in tetracycline recognition (Guay *et al*, 1994). It is also suggested is that these two residues may interact, as they are positioned at around the same depth within the helix. The fact that both mutations are needed to severely affect tetracycline resistance is consistent with the hypothesis that both residues are important. Both residues deemed to be important for substrate recognition are located in the C-terminal half of TetA(B). This is consistent with the fact that transport proteins show most homology with each other in their N-terminal portions which may be involved in energisation of the transport function whist the C terminal halves, being less similar may have evolved to specifically recognise a particular substrate.

The comparison analysis shown in this Chapter has indicated that OtrB is less closely related to most other tetracycline pumps than it is to the proton/drug antiporters from other antibiotic producing Streptomycetes. Also, OtrB is more closely related to some multidrug exporters than it is to these narrow specificity proton antiporters. This, coupled with the fact that tetracycline exporters appear in different branches of the MFS phylogeny, suggests that tetracycline transporters have evolved separately in the various lineages of the MFS. It has been suggested that the Gram negative TET resistance pumps evolved from a shared ancestor in the producer organisms, *Streptomyces* ( Benveniste & Davies, 1973).

Multidrug resistance within the MFS has occurred independently on a number of occasions (Paulsen *et al* 1996). This conclusion is drawn as the broad specificity members of the MFS are no more related to each other than they are to other members of their respective families (Lewis, 1994). OtrB seems to share more sequence similarity with various multidrug exporters including many from yeasts than it does with TET proteins from other species of bacteria. The upstream region encodes the putative repressor protein OtrR, which has more in common with the upstream repressors of multidrug transporters than it does to those of the TET family. Whether or not OtrB can transport more diverse compounds has yet to be investigated. At this point we should remember that *Streptomyces* were once classified alongside the fungi and share some characteristics of fungal organisms:- complex life-cycle, production of secondary metabolites etc. So perhaps this relationship is not as strange as it first appears.

It can be speculated that TET resistance has evolved in other species of bacteria, especially the medically important bacteria, as a result of the huge use of the drug world-wide. This selective pressure may have driven the spread of TET resistance genes via horizontal transfer, resulting in a large and wide-ranging population of bacteria resistant to the drug. Also it has been shown that OtrB will give resistance against a wider range of substrates when expressed in *E. coli* than is observed for *S. rimosus* (4.4). So possibly the host strain also exerts some evolutionary pressure on the protein because the kinetics of drug binding may be different in the Gram-negative and Gram-positive strains owing to their physiological differences. Over time, residues which result in different kinetic constants for drug recognition and transport maybe selected for.

## 3.8 Future Work

In order to know more about the topology of OtrB, its functional motifs, its position within the MFS and the complex evolutionary history thereof, there are several questions that should be asked and experiments that should now be performed.

1) Regarding the 14 TMH's presumed to make up the OtrB protein, the presence and exact positioning of the transmembrane regions could be determined by employing a fusion strategy, using fusions to reporters such as alkaline phosphatase or beta-lactamase which show different phenotypes in periplasmic and non periplasmic domains. This approach has been used successfully on TetA(C) (Allard & Bertrand, 1993) and recently on TetK (Ginn *et al*, 1997).

2) Site-directed substitution mutagenesis to elucidate which residues are essential and to give clues as to their function.

3) Resistance and specificity studies of CtrB, in order to investigate its relationship to OtrB.

4) Studies to investigate any cross-resistance between OtrB and the SvOrf4 gene product using Granatacin and Oxytetracycline as the substrates, also further studies

on the substrate specificity of OtrB in order to assess its possible function as a multidrug transporter.

The specificity of OtrB when expressed in *E. coli* was broader than when the plasmid-borne *otrB* gene was present in *Streptomyces* spp. that did not contain the OTC cluster. ( Section 4.2). This observation became the basis for the specificity studies carried out in Chapters 4 and 5 of this thesis. The observation that Gram negative TET resistant bacteria have efflux pumps of 12 TMH's and that 14 TMH s are present in the TET efflux pumps of Gram positive bacteria was also followed up. This work is detailed in Chapter 5 and involves deletion of the two central putative helices of OtrB and the expression of the truncated gene in *E. coli.*

# CHAPTER 4

## Expression of *otrB* in *Streptomyces* and *E. coli*

# CHAPTER 4

## Expression of *otrB* in *Streptomyces* and *E. coli.*

### 4.1 Introduction.

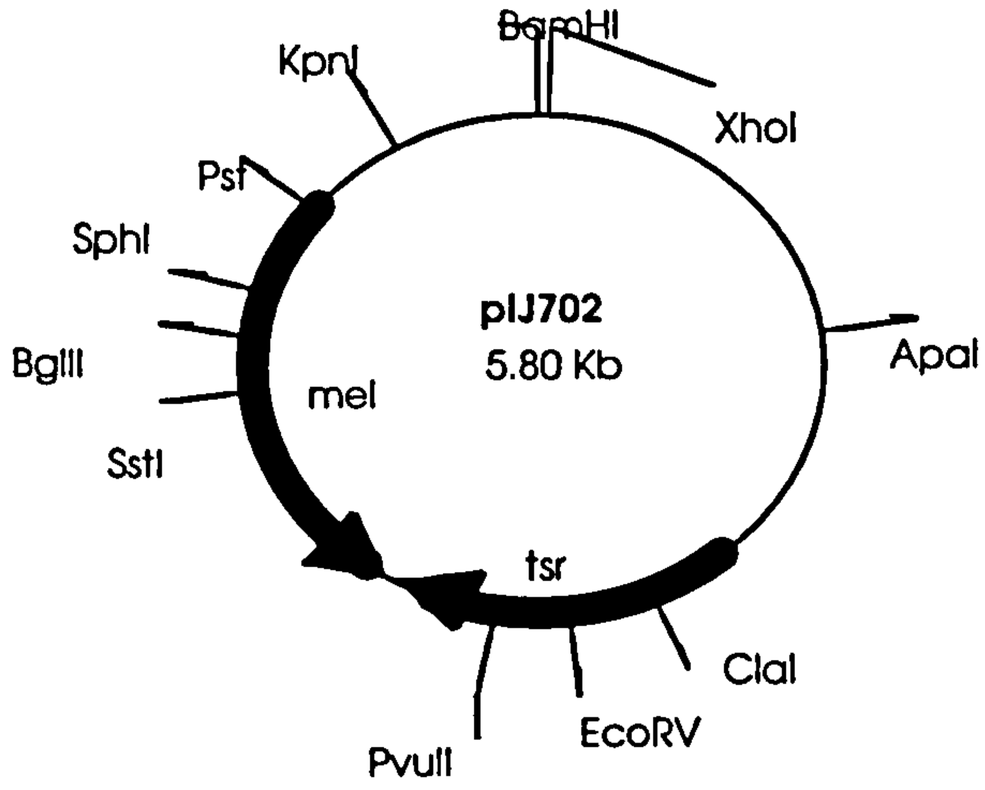
OtrB has been known to be an oxytetracycline resistance determinant since initial work on the cluster in the 1980's (Ohnuki *et al.*, 1985, Butler *et al.*, 1989). Since then, OtrB, (as Tet347 (Reynes *et al*, 1988) has been classified as a member of the rapidly expanding Major Facilitator Superfamily of transmembrane transport proteins (Paulsen *et al* 1996; Paulsen & Skurray, 1993).

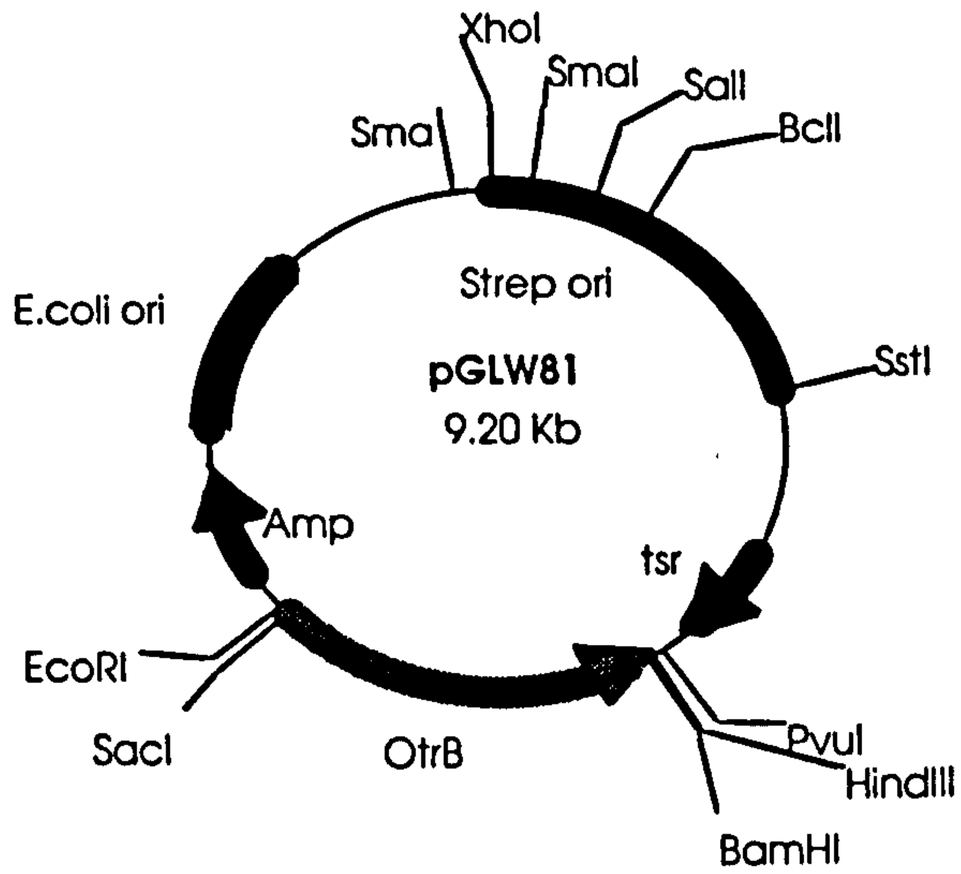
The host organism, *Streptomyces rimosus* possesses at least two other resistance determinants, *otrA* (Ohnuki *et al,* 1985) which modifies the ribosome and so prevents translational arrest ( Doyle *et al,* 1991), and *otrC* (Hunter, unpublished results) about which little is known at this time. It is also reported that there is evidence for a second "silent" OTC cluster on the *Streptomyces rimosus* chromosome (Marsic (this laboratory) unpublished results). For these reasons, the level of resistance to oxytetracycline offered by OtrB cannot be measured in the host background.

The *otrB* gene has previously been subcloned and sequenced in this laboratory, (MacGregor-Pryde, 1995). Two plasmids were used in this study, pGLW81 ( K.J. Linton, this laboratory) contains the *otrB* gene on a *Streptomyces-E. coli* shuttle vector and pGLW12 ( Fig 4.1) which contains *otrB* and part of *otrR* , a regulatory gene upstream of *otrB* that is thought to be a repressor (MacGregor-Pryde, 1995).

To determine the contribution of OtrB to the cell's overall OTC resistance, the plasmids mentioned above were transformed separately into the heterologous host *Streptomyces albus,* which is not a producer of oxytetracycline and has not been shown to contain OTC production or resistance genes. The plasmids were also transformed into *S. rimosus* strain 15883s. This strain is deleted for the OTC cluster including both *otrA* and *otrB* (Butler *et al*, 1989), enabling evaluation of OTC resistance in the host background.

Expression of the *otrB* gene in *E. coli* was important for two main reasons, the first being that there are a number of well-characterised, easily available systems for the



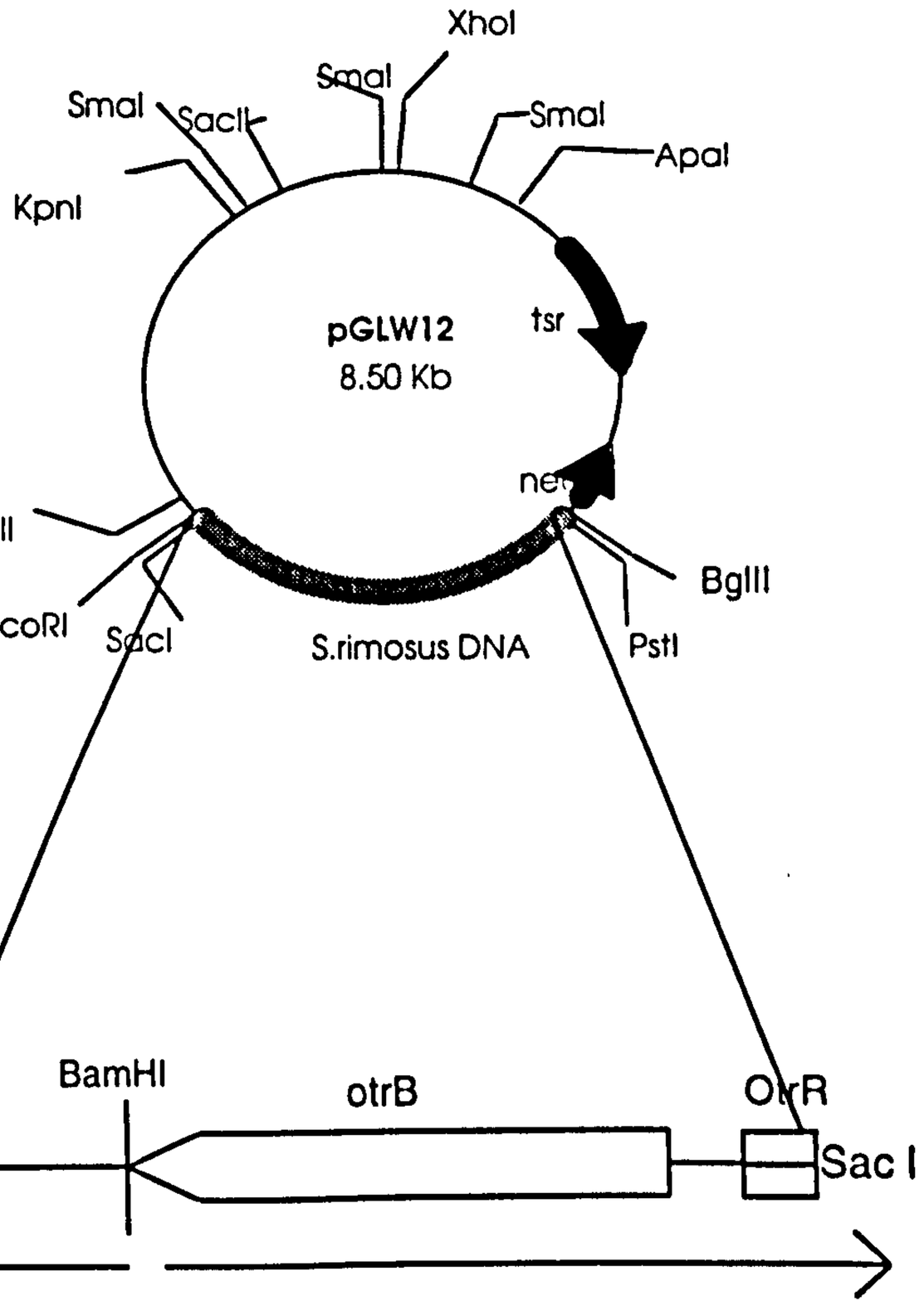




Fig 4.1.

Plasmids pIJ702, pGLW12 and pGLW8. References are given in Table 2.2

over-expression of foreign proteins in this host. By developing a system which produces reasonable quantities of the protein it should be possible in the future to use immunoblotting techniques to quantify the amount of OtrB produced under different conditions and to set up systems to investigate the transport of OTC and presumably protons across the *E. coli* membrane. By over-expressing *otrB*, purification of the protein product for structural determination should be achievable. Secondly, by expressing the OtrB protein which has the characteristic 14 putative transmembrane helices of Streptomycete transport proteins, in this Gram-negative background , a comparison can be drawn between it and the Tet proteins of Gram-negative bacteria. These proteins consist of 12 such helices (Eckert & Beck, 1989). The best studied of these proteins is TetA from transposon Tn*10* of *E. coli.*

## 4.2 Aims of this chapter

1. To determine the level of resistance to oxytetracycline provided by *otrB in Streptomyces rimosus 15883s,* and *Streptomyces albus.*

2. To express the *otrB* gene in the heterologous host, *Escherichia coli* and determine the level of resistance provided in this foreign host.

3. To investigate the specificity of OtrB for oxytetracycline and whether it is able to provide any resistance against the related compounds tetracycline and chlortetracycline.

## 4.3 Initial expression of *otrB* in *Streptomyces.*

### 4.3.1 Plasmid constructs containing *otrB*.

Both plasmid pGLW12 and pGLW81 contain the part of the region of DNA known to encode the divergent *otrB/otrR* promoter region (Fig 4.1). pGLW12 contains DNA encoding 86 of the 151 amino acids of the putative regulator protein OtrR and the promoter region. It maybe that the expression of *otrB* from this promoter will be affected in some way by a product of the partial *otrR* gene upstream. pGLW81 contains the promoter region and only 24 bases of *otrR* meaning that gene expression from this construct is probably not subject to regulation by OtrR. This could have an effect on the level of resistance shown by organisms containing this plasmid.

### 4.3.2 Expression in *S. rimosus* strain 15883s.

Plasmids pGLW12 and pGLW81 were transformed into protoplasts of *S. rimosus* strain 15883s ( 2.3.10 for methodology). A spore suspension of the strain was serially diluted and 100µl aliquots spread on to TSB plates. The dilution which gave a countable number of colonies, i.e. between 50 and 200, was chosen for the resistance tests. The spore suspension was streaked onto TSB plates containing increasing concentrations of oxytetracycline. Colonies growing on the OTC plates were counted and compared to the number of colonies on the control plate. The concentration at which no colonies were visible was taken as the MIC (minimum inhibitory concentration). Hence the concentration below this was the equivalent to the level of resistance provided by *otrB* in any particular strain. As a control, plasmid pIJ702 (Fig 4.1) which does not contain the *otrB* gene was transformed into both *S. rimosus* 15883s and *S. albus* and tests carried out as before (Table 4.1). Tests were also carried out on plates containing tetracycline and chlortetracycline to investigate the specificity of OtrB. Results are given in Table 4.1

### 4.3.3 Expression of OtrB in *S. albus*

As mentioned above, there is evidence to suggest that there is a third oxytetracycline resistance determinant *otrC*, in *S. rimosus*. For this reason it was pertinent to repeat the tests carried out above in a strain negative for OTC production. Plasmids pGLW12 and pGLW81 were transformed into protoplasts of *S. albus* and the tests performed as for *S rimosus* strain 15883s. Control plasmid pIJ702 was included as in the previous section (Table 4.1).

Plasmid pGLW81 conferred the highest levels of OTC resistance when present in both *S. rimosus* 1583s and *S. albus*, giving resistance up to 12 and 10µg/ml respectively. The resistance conferred by pGLW12 was lower at 5µg/ml in both *S. rimosus* 15883s and in *S. albus*. As expected, the control plasmid pIJ702 which does not contain *otrB* confers no OTC resistance on either host. pGLW12 contains part of the putative repressor gene *otrR*. However, pGLW81 contains only a few bases of this gene. It is conceivable that the absence of the repressor deregulates *otrB* expression, resulting in elevated levels of OTC resistance. There was no resistance to

tetracycline or chlortetracycline by any of the strains tested, indicating high specificity of OtrB for oxytetracycline when expressed in a Streptomycete background.

Table 4.1 Results of streak tests in *Streptomyces sp.*

| OTC conc. µg/ml | *S. rimosus* 15883s | | | *S. albus* | | |
|---|---|---|---|---|---|---|
| | pGLW12 | pGLW81 | pIJ702 | pGLW12 | pGLW81 | pIJ702 |
| 0 | +* | + | + | + | + | + |
| 2 | + | + | - | + | + | - |
| 4 | + | + | - | + | + | - |
| 5 | + | + | - | + | + | - |
| 6 | -** | + | - | - | + | - |
| 8 | - | + | - | - | + | - |
| 10 | - | + | - | - | + | - |
| 12 | - | + | - | | - | - |
| 15 | - | - | - | - | - | - |

* + denotes growth

** - denotes lack of growth

## 4.4. Expression of *otrB* in *E.coli.*

### 4.4.1 The T7 expression system.

The T7 expression system works by expressing the gene of choice from the viral T7 promoter, present on an ampicillin-resistant vector. This is recognised by viral T7 RNA polymerase, hence the T7 RNA polymerase gene must be present in the host cell. The strain used as host is *E. coli* BL21 (DE3). As a B strain, BL21 has the potential advantage that it is deficient in the *lon* protease and the OmpT outer membrane protein that can degrade proteins during purification. DE3 is a lambda derivative that contains a DNA fragment containing the *lacI* gene, the *lac*UV5 promoter and the gene for T7 RNA. polymerase (Studier & Moffat, 1986). This fragment is inserted into the *int* gene, because when this gene is then inactivated the

lysogen is very stable and requires a helper for excision from the chromosome. The only promoter which directs transcription of the T7 RNA polymerase gene in this strain is the IPTG inducible *lac*UV5 promoter. The expression of the gene of choice is induced by the addition of 0.4 mM IPTG to a log phase culture. Some proteins may be toxic to the cell if expressed at a high level. This is especially true for membrane proteins. If a toxic protein is expressed constitutively, albeit at low levels, the host cells become "sick" and growth is affected. To alleviate this problem, T7 lysozyme may be expressed in the cell. This enzyme inhibits transcription by T7 RNA polymerase (Inouye *et al,* 1973; Moffat & Studier, 1987; Studier, 1991). When produced from the cloned gene, relatively high levels of T7 lysozyme can be tolerated by *E. coli*, apparently because the protein is unable to pass through the inner membrane to reach the peptidoglycan layer. Any basal levels of T7 polymerase present in the cell prior to induction will be prevented from transcribing the gene of choice by the action of the lysozyme. However, when the cells are induced, the levels of T7 RNA polymerase are increased so that the lysozyme is swamped, allowing expression of the chosen gene. The lysozyme gene is present on a plasmid, pACYC184, (Chang & Cohen, 1978) which also carries chloramphenicol resistance. There are two versions of this plasmid, pLysE and pLysS. pLysE expresses T7 lysozyme from the pACYC184 *tet* promoter. Cells carrying this plasmid accumulate fairly high levels of T7 lysozyme. Cells carrying pLysE display a reduced growth rate. A plasmid containing the fragment in the opposite orientation where it is transcribed from the $\phi$3.8 promoter is known as pLysS; cells carrying this plasmid accumulate much lower levels of T7 lysozyme and do not show any decrease in growth rate. There are a number of vectors commercially available for this system (pET system from Novagen, Qiaexpress system from Qiagen).

### 4.4.2 Construction of expression plasmid pJJ4.

pJJ4 is based on the expression vector pT7-7 (Tabor & Richardson, 1985).

The *otrB* gene encoding the oxytetracycline efflux protein from *S. rimosus* (on a 2.3kb *EcoR* I - *Hind* III fragment from pGLW81) was cloned into pUC18 to create pJJ1. This was the template in the polymerase chain reaction using a mutagenic primer to introduce an *NdeI* site (CATATG ) at the N-terminus. The product of this

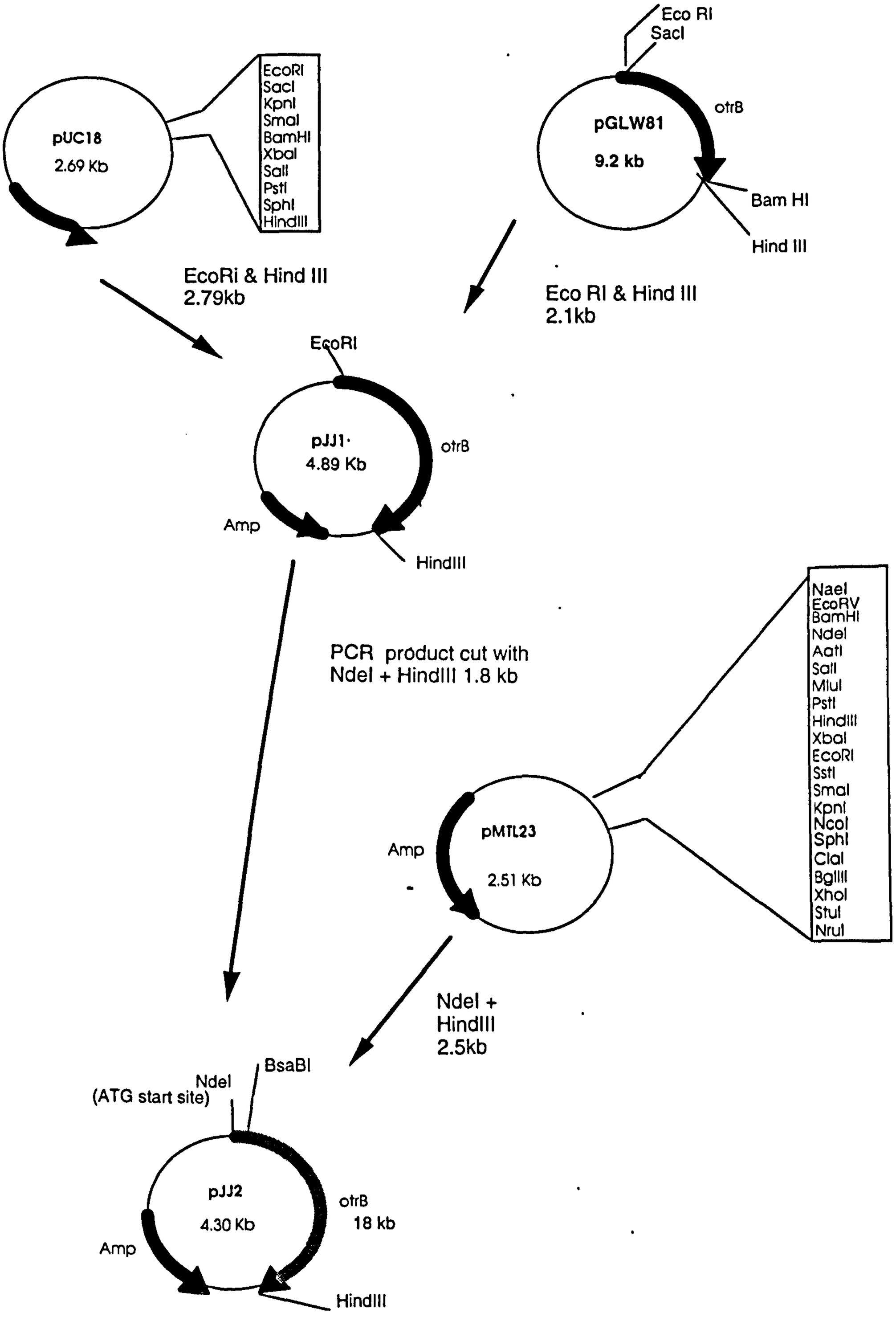
reaction with engineered *Nde* I site was digested with *Nde* I and *Hind* III and ligated into similarly digested pMTL23 to create pJJ2.

The first 250 bases of the *otrB* PCR product were sequenced from pJJ2. To compare these sequence data with wild type *otrB* sequence, a 0.46kb *Apa* I fragment from pJJ1 beginning 70bp upstream of the *otrB* start site was cloned into pIBI24 to create pJJIBI and sequenced using a ABI automated sequencer and universal forward primer. Both the PCR product and the wild type sequence showed differences to the *otrB* sequence presented by MacGregor-Pryde (1995). This is discussed fully in Chapter 3. (3.4.1). To circumvent mutations due to *Taq* polymerase further downstream of the sequenced region, pJJ2 isolated from CB51 cells (*dam*$^-$) was restricted with *Bsab*I and *Hind*III to remove the region which had not been sequenced. pJJ1 was similarly restricted and the C-terminal 1.669kb fragment of wild type *otrB* was ligated with the 2.7kb pJJ2 (vector and small insert) to create pJJ6. pJJ6 was restricted with *Nde*I and *Hind*III and the 1.8kb *otrB* fragment ligated with similarly-cut pT7-7 to create pJJ4 (Fig 4.2).

A second expression construct, was created using the 5.4 Kb Novagen vector pET21b, the *Nde*I - *Hind*III fragment from pJJ4 was ligated with similarly cut vector to create pJJ5. (Fig 4.3). pJJ5 does not express the HisTag present on this vector.

### 4.4.3 Expression of pJJ4 and pJJ5.

The plasmids were transformed into three strains of *E. coli* BL21(DE3), BL21(DE3) pLysE and BL21(DE3) pLysS. All induction experiments were carried out as described in 2.3.6.1. Streaks of induced and uninduced culture were plated on L-agar containing increasing concentrations of OTC either with or without the addition of 0.4mM IPTG to the cooled molten agar. Growth on oxytetracycline was taken as a measure of the level of expression of the *otrB* gene. Results of the tests for pJJ4 and pJJ5 are presented in Figs 4.4 and 4.5 respectively.


pUC18
2.69 Kb
EcoRI
SacI
KpnI
SmaI
BamHI
XbaI
SalI
PstI
SphI
HindIII
EcoRi & Hind III
2.79kb
Eco RI
SacI
pGLW81
9.2 kb
otrB
Bam HI
Hind III
Eco RI & Hind III
2.1kb
EcoRI
pJJ1·
4.89 Kb
otrB
Amp
HindIII
PCR product cut with
NdeI + HindIII 1.8 kb
NaeI
EcoRV
BamHI
NdeI
AatI
SalI
MluI
PstI
HindIII
XbaI
EcoRI
SstI
SmaI
KpnI
NcoI
SphI
ClaI
BglII
XhoI
StuI
NruI
pMTL23
2.51 Kb
Amp
NdeI +
HindIII
2.5kb
BsaBI
NdeI
(ATG start site)
pJJ2
4.30 Kb
otrB
18 kb
Amp
HindIII

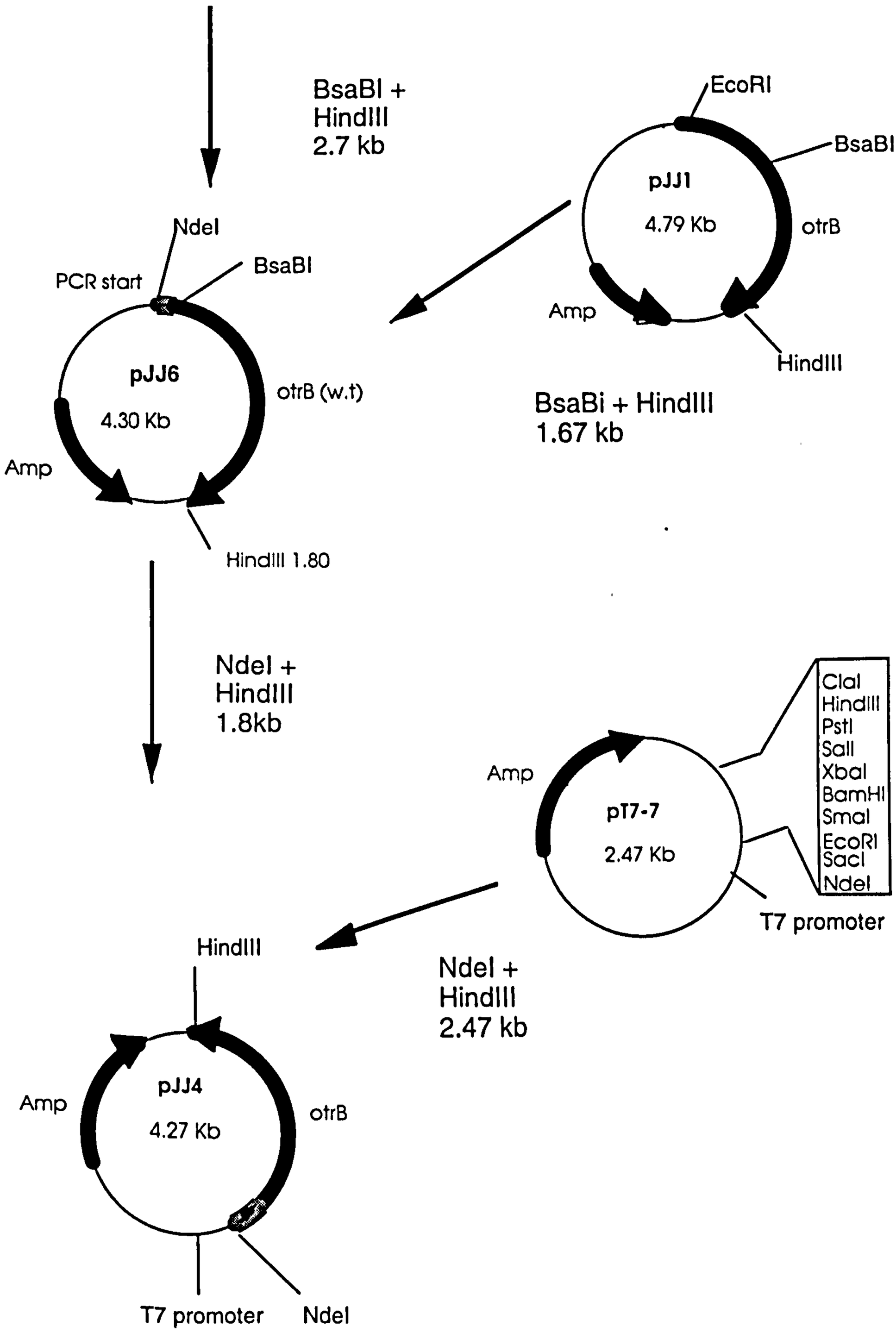



Fig 4.2

The construction of plasmids pJJ1 to pJJ4

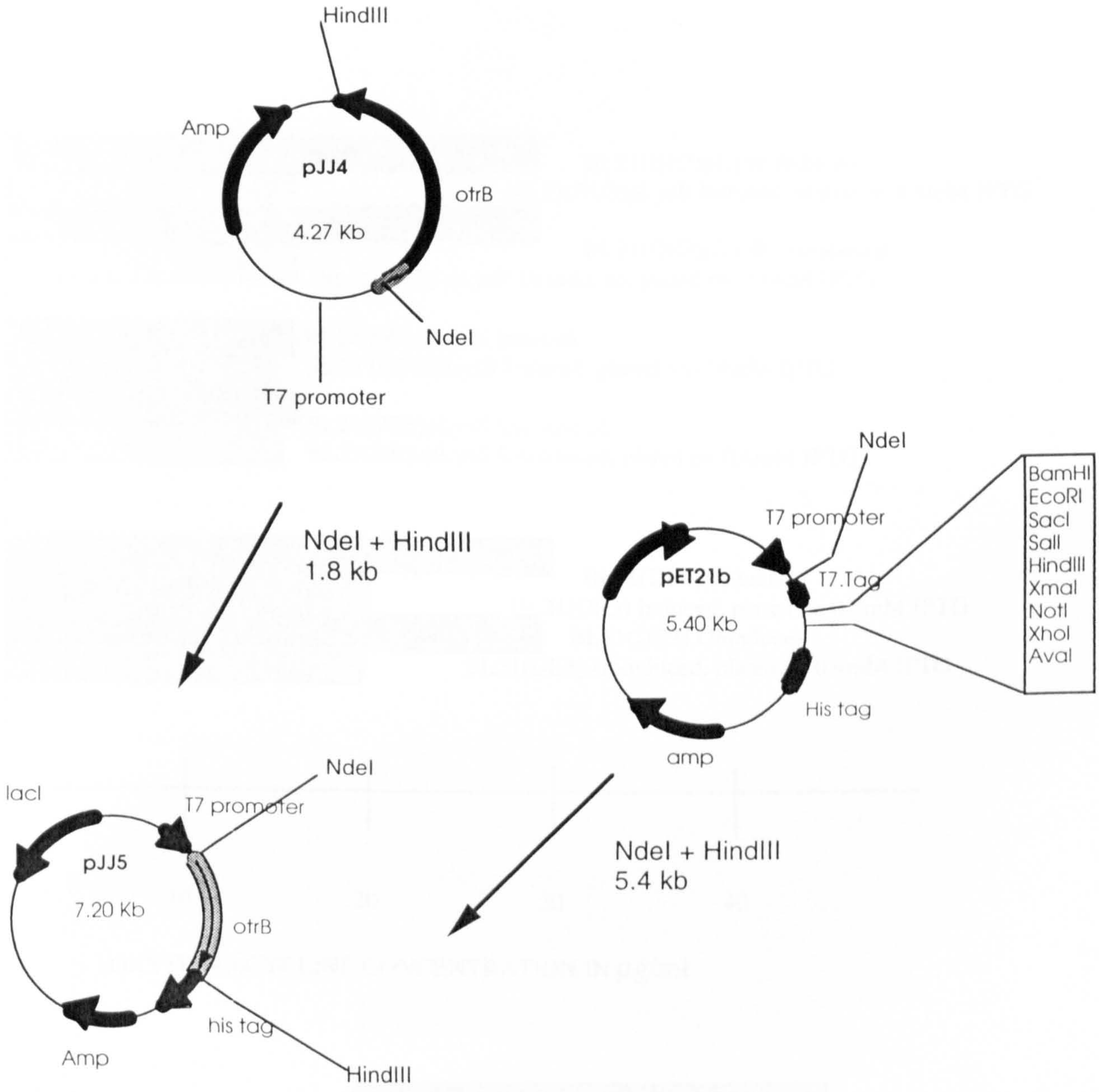



Fig 4.3

The construction of plasmid pJJ5.

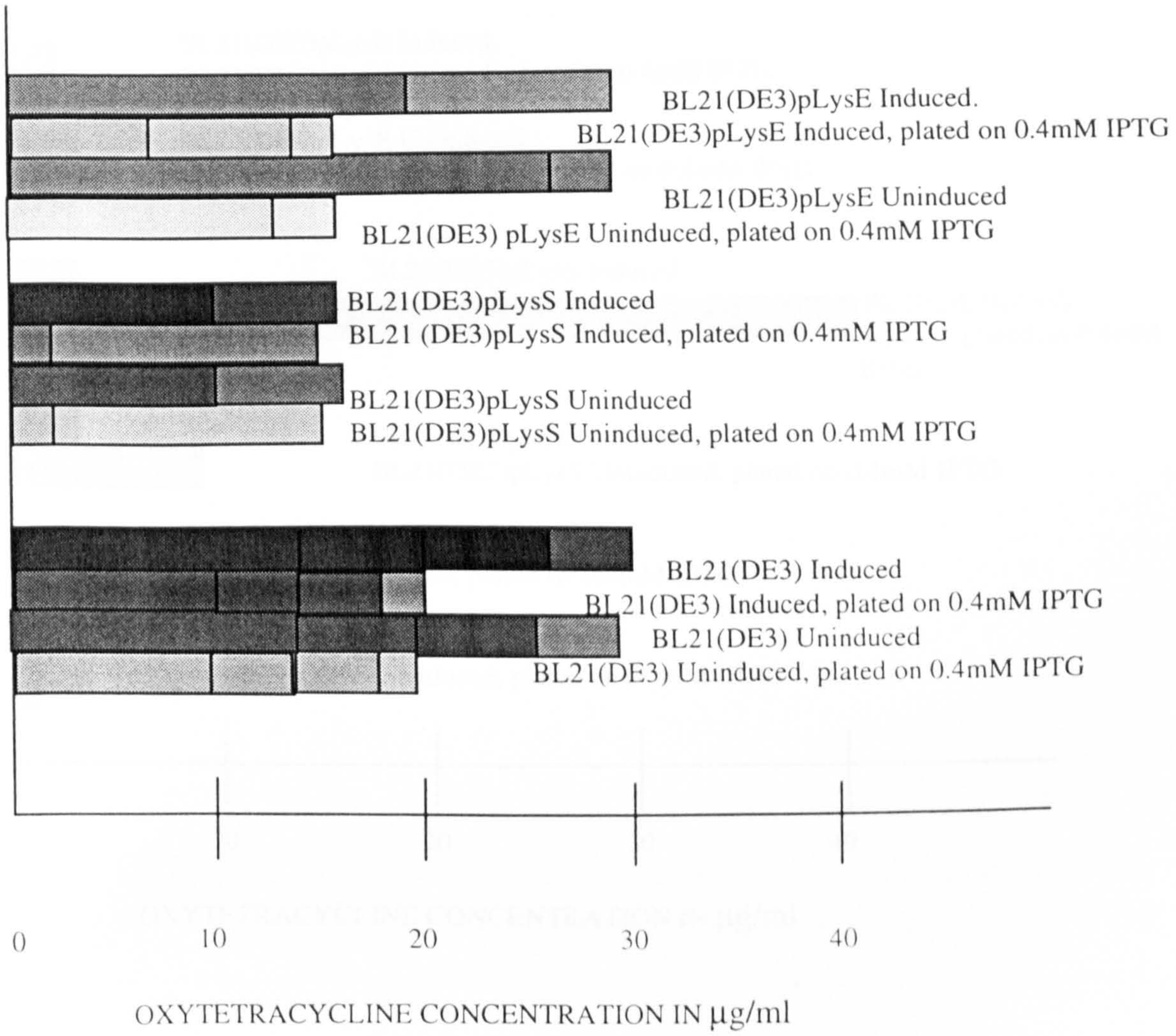



Fig 4.4

OTC resistance profiles of pJJ4.

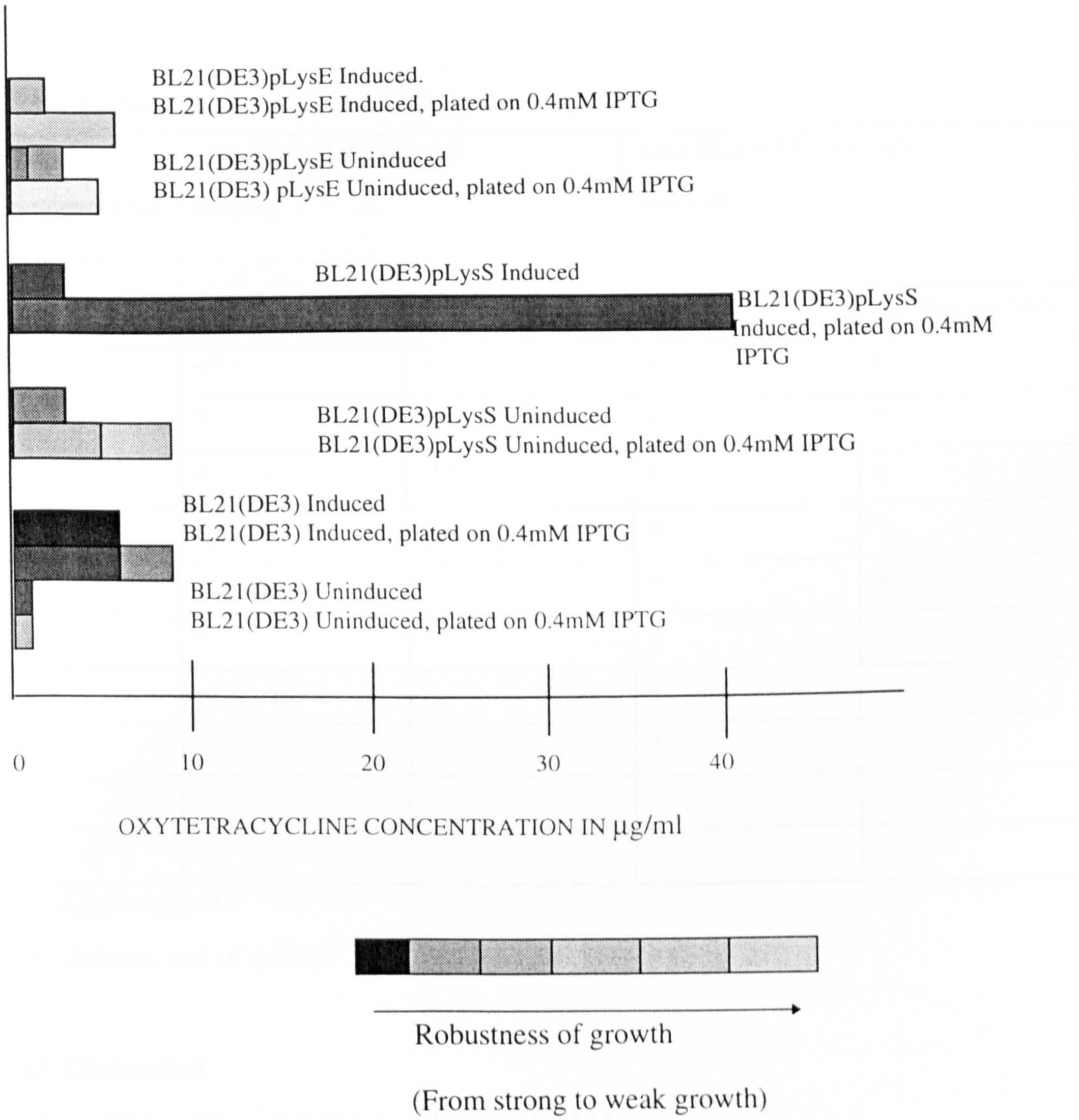



Fig 4.5

OTC resistance profiles of pJJ5.

Table 4.2 Results of specificity tests

| drug concentration in µg/ml | pJJ4 BL21(DE3)pLysS induced | | pJJ5 BL21(DE3)pLysS induced | |
|---|---|---|---|---|
| | tetracycline | chlortetracycline | tetracycline | chlortetracycline |
| 0 | +* | + | + | + |
| 2.5 | + | + | + | + |
| 5 | + | -** | + | + |
| 7.5 | + | - | + | + |
| 10 | + | - | + | + |
| 12 | - | - | + | - |
| 15 | - | - | + | - |
| 20 | - | - | + | - |
| 25 | - | - | + | - |
| 30 | - | - | - | - |

* + denotes growth

** - denotes lack of growth

## 4.5 Discussion.

### 4.5.1. Discussion of resistance levels of pJJ4

The pT7-7 based plasmid, pJJ4 shows more robust growth on plates that do not contain IPTG than on those with IPTG. This is true for induced and uninduced cultures, whether or not OTC is present. This second point shows poor control of expression of the target gene in this vector. The fact that all cultures containing pJJ4 show more robust growth on plates without the inducer could be because the culture is already expressing the OtrB protein at the maximal level which can be tolerated by the cells' physiology and further increasing amounts of the inducer cause expression of the protein to reach levels that are lethal to the cell.

When streaked onto plates that contained 0.4mM IPTG, cultures of BL21(DE3)/pJJ4 , induced or uninduced, did not grow at oxytetracycline

concentrations above 20µg/ml whilst strains BL21(DE3) pLysS and BL21(DE3) pLysE transformed with pJJ4 both had a minimum inhibitory concentration of 12µg/ml induced or uninduced. The difference between strains can be explained by the fact that the pLys plasmids encode T7 lysozyme which inhibits transcription of the *otrB* gene leading to reduced levels of the protein in the cells.

When streaked onto plates that did NOT contain the inducer IPTG, strains BL21(DE3)/pJJ4 and BL21(DE3) pLysE /pJJ4 both grew at 30µg/ml. However when the strains were induced, BL21(DE3) pLysE was not able to grow at OTC concentrations above 15µg/ml, once again this is probably due to expression levels becoming toxic when IPTG was added to the culture. Strain BL21(DE3) pLysS/pJJ4 would not grow at concentrations higher than 15µg/ml whether induced or uninduced.

### 4.5.2 Discussion of OTC resistance levels conferred by pJJ5

The unsatisfactory control of expression using T7-7 as an expression vector lead to the use of another expression system. The pET vector series is widely used in the expression of prokaryote and eukaryote proteins in *E. coli* (Aukhil *et al*, 1993; Campbell, 1992; Dietrich *et al*, 1991) With the pET system the control of expression should be more tightly regulated than when using the pT7-7 vector. There are a number of different pET vectors suitable for different purposes. The vector chosen for this work was pET21-b, this vector gave the option of creating a C-terminal polyhistidine fusion to OtrB.

pJJ5 was a more suitable construct for evaluating levels of OTC resistance provided by OtrB. Maximal resistance with this construct was to more than 40 µg/ml, higher than the maximum levels observed for pJJ4 which showed some growth at 30 µg/ml but none at 35 µg/ml. The regulation was well controlled as there was always little or no resistance in the uninduced samples from all three strains. Growth of induced samples was not observed on plates containing oxytetracycline without IPTG. This shows that in the absence of the inducer, expression of *otrB* is switched off. Little or no growth of uninduced samples was observed on plates containing oxytetracycline whether or not they also contained 0.4 mM IPTG. Uninduced cells have not expressed *otrB* and so are killed by the oxytetracycline before any IPTG that

may be present in the agar can activate T7 polymerase and turn on the expression of *otrB*.

For samples that have been induced with IPTG and grown for some time to allow expression of the OtrB protein, the subsequent levels of resistance to OTC depend on the strain. In this situation, BL21(DE3) pLysS was found to be the optimal background, giving resistance to up to 50 μg/ml of the drug. Presumably this is because levels of OtrB in BL21(DE3) become too toxic for the cell after some time and in BL21(DE3) pLysE, the level of *otrB* expression is reduced considerably by the inhibitory actions of the T7 lysozyme. As T7 lysozyme is expressed from a *tet* promoter in pLysE, it could be that this promoter is further activated by the presence of OTC in the media leading to very high levels of T7 lysozyme expression, which then inhibit T7 polymerase even when the cells are induced.

### 4.5.3 Discussion of specificity tests.

All specificity tests were carried out in *E. coli* BL21(DE3) pLysS as this gave the highest resistance to oxytetracycline. Results of the tests are shown in Table 4.2.

The pJJ4 construct gave resistance to tetracycline (TET) to the same levels as with the natural substrate oxytetracycline (10 μg/ml). However, pJJ4 conferred resistance to only 1/4 of the concentration of chlortetracycline (CTC) (2.5μg/ml). These levels were the same whether the cells were induced with IPTG or not.

pJJ5 conferred most resistance to oxytetracycline (50 μg/ml), giving resistance to half this level of tetracycline and one quarter of this level of chlortetracycline. Uninduced cells would not grow in the presence of any of the three drugs.

When these results are compared to the tests performed with pGLW12 and pGLW81 in *Streptomyces* sp. it should be noted that the OtrB gene product is not as substrate specific when expressed in *E. coli* as it is when expressed in its natural host. There are a number of possible explanations for this.

(1) The differences could be due to the variation in the structure of the cytoplasmic membrane between Gram positive and negative bacteria which may then affect the conformation which OtrB adopts. The $K_M$ of OtrB maybe more similar for the three different drugs when it is present in the *E. coli* membrane whereas the configuration

of OtrB in *Streptomyces* could mean the $K_M$ for OTC is significantly smaller than for CTC or TET

(2) It is possible that in the natural host, the protein is produced at much lower levels than it is in the *E. coli* expression system which is, by its design, maximising the levels of protein produced. Higher copy number of the protein may lead to an increased rate of transport, even for a poor substrate.

(3) The metal dication ($M^{2+}$) component of the transported complex may be different between the two types of bacteria, this would cause a change in the character of the substrate which could possibly affect the specificity of the protein.

# CHAPTER 5

## Construction of OtrB-Polyhistidine fusion, expression and attempted isolation.

# CHAPTER 5

## Construction of polyhistidine fusion, expression and isolation.

### 5.1 Introduction

Fusion proteins containing a peptide tag can be used to facilitate purification of proteins of interest from prokaryotic and eukaryotic organisms. A fragment of DNA containing the gene of interest is cloned into a plasmid vector with a strongly regulated promoter. The gene is cloned in frame with a sequence of nucleotides encoding a protein tag. The tag can be positioned at the N- or C-terminus of the protein. The tag should exhibit a strong affinity to the complementary ligand used in the main separation process. If the protein of interest is to be secreted, a signal sequence should be included. Also a peptide cleavage site may be introduced between the protein of interest and the tag. This enables the desired protein to be obtained in its native state. Whereas the purification tag must be a peptide or protein, the complemntary ligand may be derived from any biological or chemical source. Examples of such pairings are antigen & antibody, enzyme or enzyme binding site & substrate or substrate analogue, hormone & receptor, binding protein & binding ligand, polyhistidine & metal chelate ligand and charged peptide & charge ligand or ion exchangers. The ideal affinity tag should be able to form a stable complex with an inexpensive, chemically stable and easily available ligand with an absolute specificity, yet should dissociate under mild conditions. Commonly used tags are, maltose binding protein, the FLAG ™ tag and poly-histidine peptide.

<u>FLAG ™ Fusions</u>

This affinity tag developed by International Biotechnologies Inc. consists of an 8 amino acid sequence, AspTyrLysAspAspAspAspLys which is used as an N-terminal fusion (Hopp *et al*, 1989). The first four residues make up the antigen by which the fusion protein is recognised by the mouse anti-FLAG monoclonal antibody. The last four amino acids make up the cleavage site for enterokinase which can be used to remove the tag from the purified protein. Although separation is achieved by use of an antibody-antigen complex (usually very strongly associated), dissociation is easily

achieved because the interaction depends on the presence of calcium ions. When calcium ions are withdrawn the complex dissociates.

Maltose Binding Protein (MBP) Fusions

Fusions with MBP are commonly used to induce secretion of the protein into the *E. coli* periplasm. The fusion protein adsorbs onto cross-linked starch. Elution may be acheived by adding maltose which binds with a higher specificity than the starch and so displaces it. A disadvantage of this system is the high MW of MBP (about 41kDa). However for large-scale purifications it remains a good system due to its use of an inexpensive ligand.

Poly-Histidine Fusions

Chelating ligands used with the polyhistidine system include IDA (iminodiacetic acid), TED (N,N,N'-Tris (carboxymethyl) ethylene diamine and NTA (nitrilotriacetic acid) (Flaschel & Friehs, 1993). Metal ions which have been used include $Cu^{2+}$, $Ni^{2}$, $Co^{2+}$ or $Zn^{2+}$. The Ni (II)-NTA group is suited to complexing proteins containing a run of several histidine residues (Hochuli, 1988) whilst other metal-chelate combinations best complex proteins with only one or two histidine residues. Histidine is a rare amino acid in many proteins, therefore a run of several histidines in sequence is extremely uncommon. The Ni (II)-NTA system has thus become the system of choice for purification of poly-histidine fusions, the strategy being to separate the tagged protein of interest from the bulk of other bacterial proteins by metal affinity chromatography ( Hochuli *et al*, 1988 ). This is done by passing a cell extract through a column that contains NTA-nickel-agarose. The tagged protein binds to the matrix through interactions with the divalent cations and can then be eluted with high levels of imidazole, competing metal ions or by lowering the pH to a value where binding cannot occur.

The proteins produced by this method are often pure enough to be subjected to further analyses such as NMR and CD studies that can determine structural identity. The histidine tag could also be used to determine the level of expression of the protein of interest. It is known that for the *E. coli* TetA(B) protein, the level of tetracycline resistance supplied by the gene on a plasmid is not always directly

related to the level of gene expression (Walmsley pers. com.). If this were also to be the case for OtrB, HisTagging would be a suitable method of monitoring expression levels by the use of anti-His antibodies. Fusions at the amino terminus although proven useful in the isolation of many cytoplasmic fusion proteins, may interfere with the folding of membrane proteins and may affect positioning in the bacterial membrane and hence the activity of the protein. Successful isolation of membrane proteins has been achieved using fusions to the C-terminus of the protein, (Aldema *et al.*, 1996) As OtrB is a membrane protein, the fusion product was designed with the HisTag at the carboxy terminus. A second construct with an amino terminal HisTag was made later to investigate any difference in activity with respect to positioning of the HisTag.

## 5.2 Construction of OtrB fusion plasmids

In order to construct a plasmid that would express *otrB* fused to the polyhistidine tag, it was first necessary to create an intermediate-stage construct from the cloning vector pMTL23 and the *otrB* PCR product containing a mutation at the C-terminal end which replaces the stop codon with a coding version. The mutated *otrB* was subcloned from this plasmid into the expression vector pET15b and the resulting fusion product expressed in *E. coli.*

### 5.2.1 Construction of pJJ7

pJJ7 was constructed by cloning the product of a PCR reaction that used pJJ6 as the template, the mutagenic oligonucleotide, primer 2 ( 2.4.14) at the carboxy terminus and universal reverse primer at the amino terminus (Fig 5.1).

The mutagenic primer changes the original *otrB* stop codon to TGG, which codes for tryptophan, following this are a *SacI* site, a TGA stop codon, an *Xho I* site and an *EcoRI* site. The PCR product was cut with *Nde* I and *Eco R* I and cloned into similarly cut pMTL23. The PCR reaction was carried out using *pfu* polymerase which has a fidelity of $1.3 \times 10^{-6}$, approximately six times less error prone than *Taq* polymerase (Cline *et al*, 1996). The 1.63 Kb PCR product was not expected to contain any mistakes, hence the PCR product was cloned in entirety ( Fig 5.1). As *otrB* activity is manifested as OTC resistance, the activity of the *otrB* PCR product

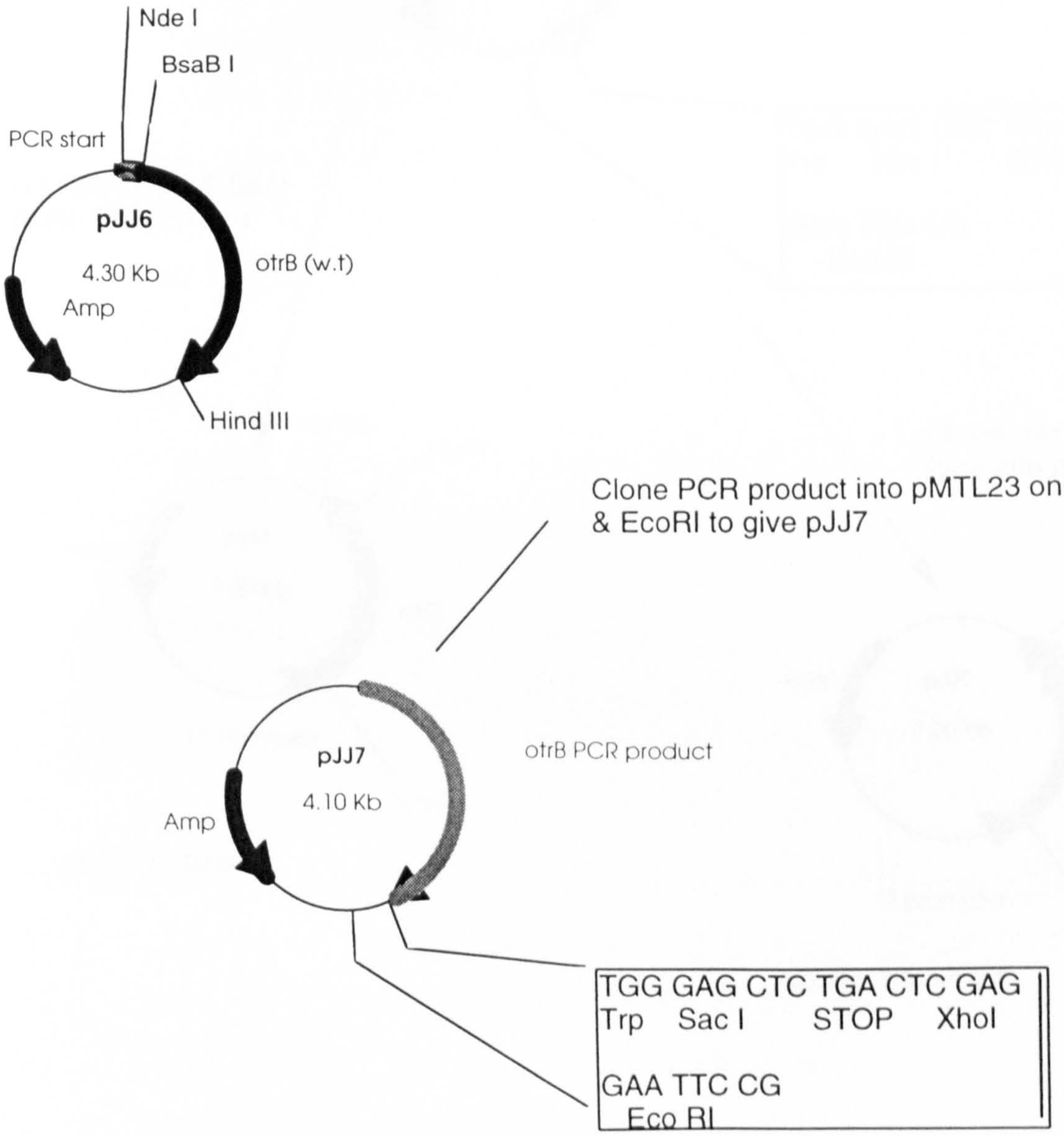



Figure 5.1

Construction of pJJ7.

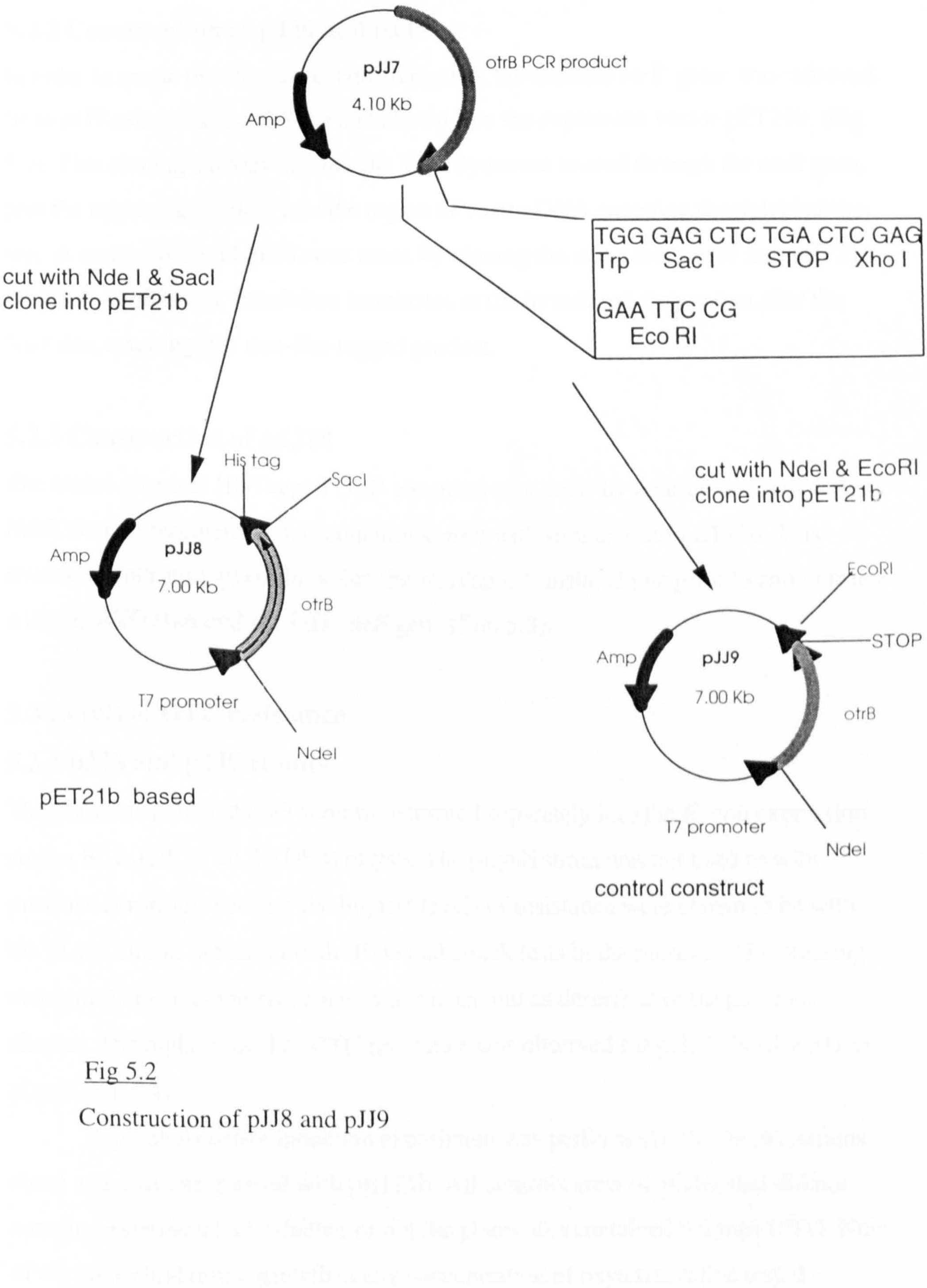



Fig 5.2

Construction of pJJ8 and pJJ9

can be measured as expression of the gene.

### 5.2.2 Construction of pJJ8 and pJJ9

In order to create the His fusion construct pJJ8, the mutated *otrB* gene was removed from pJJ7 using *Nde* I and *Sac* I and cloned into the expression vector pET21b. (Fig 5.2). This cloning strategy enables the T7 polymerase to read through the *otrB* gene, past the tryptophan codon, into the region of vector DNA encoding the polyhistidine tag. A control plasmid, pJJ9, was made by cloning the insert from pJJ7 into pET21b on *NdeI -Eco*RI ends, translation terminates at the introduced stop codon after the *SacI* site, resulting in a non-His-tagged product.

### 5.2.3 Construction of pJJ18

The amino terminal HisTagged OtrB construct was made by ligating the pJJ6 derived *Nde*I, BamHI fragment (3.3.1) containing *otrB* with similarly cut pET15b. This results in a plasmid, pJJ18 in which the HisTag is translated just prior to and in frame with the ATG start codon of the *otrB* gene (Fig. 5.3).

## 5.3 Levels of OTC resistance

### 5.3.1 pJJ8 and pJJ9 results.

The plasmids pJJ8 and pJJ9 were transformed separately into the *E. coli* expression strains BL21(DE3), BL21(DE3) pLysS, The pLysE strain was not used as with previous constructs because the highest levels of resistance were shown to be with the pLysS strain. Induction with IPTG and streak tests in the presence of increasing concentrations of oxytetracycline were carried out as described in the previous chapter. The highest level of OTC resistance was observed for pJJ8/9 in BL21(DE3) pLysS (Fig 5.4).

A parallel control induction experiment was performed with the two strains alone and also transformed with pET21b. All controls grew on plates that did not contain oxytetracycline, whether or not the plates also contained 0.4 mM IPTG. None of the controls showed growth at any concentration of oxytetracycline tested irrespective of the presence of IPTG, the lowest being 2.5 μg/ml.

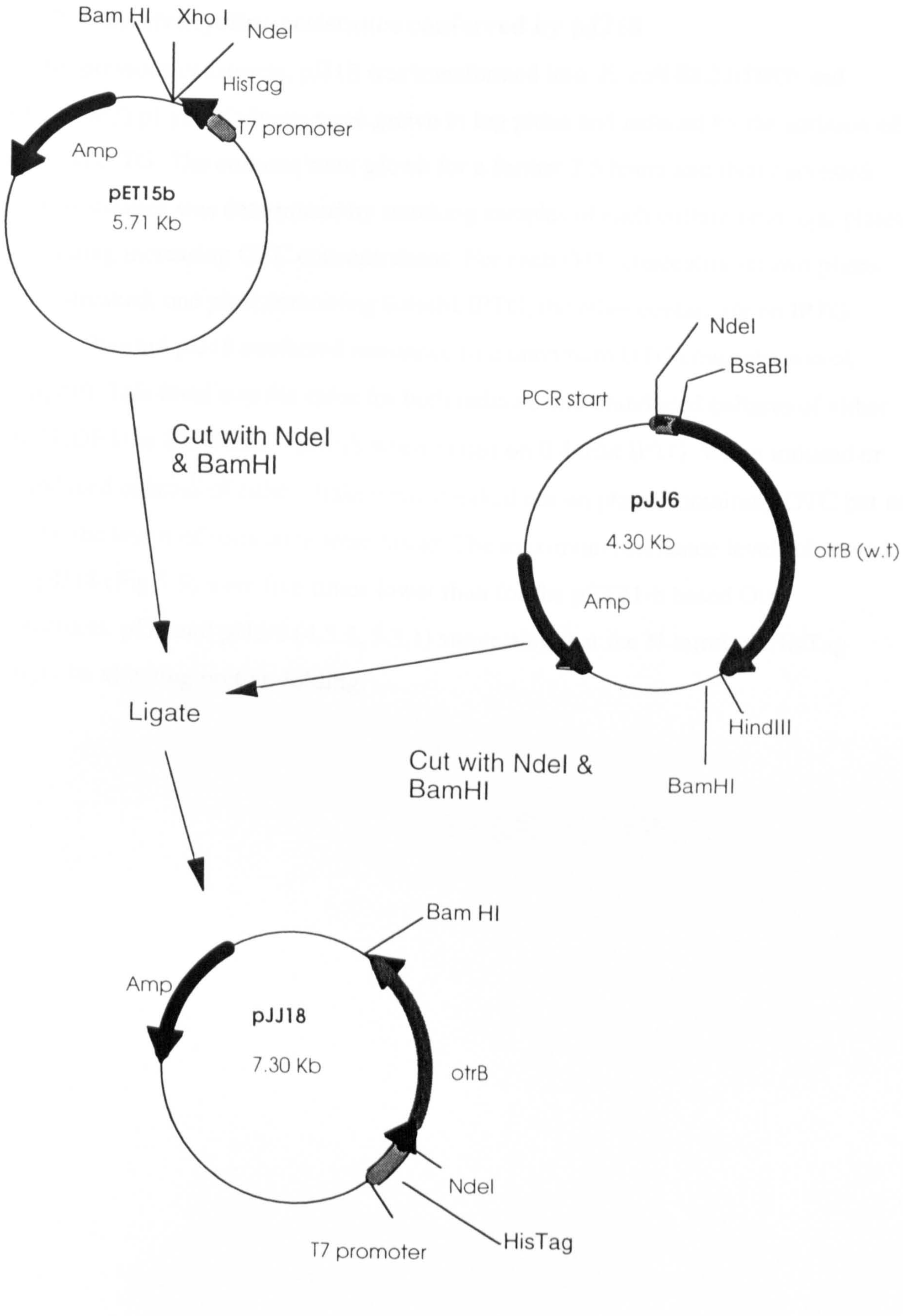



Fig 5.3

Construction of pJJ18

### 5.3.2 Oxytetracycline resistance conferred by pJJ18

As for previous constructs, pJJ18 was transformed into *E. coli* BL21(DE3) and BL21(DE3) pLysS. Cultures were grown to log phase and induced by the addition of 0.4 mM IPTG. The cultures were grown for a further 2.5 hours and then harvested. OTC resistance was determined by streaking samples of each culture onto agar plates containing increasing OTC concentrations. For each OTC concentration two plates were streaked, one plate containing 0.4mM IPTG, the other containing no IPTG.

Plasmid pJJ18 conferred resistance to a maximum OTC concentration of 10µg/ml. This level was the same for both induced and uninduced cultures of either BL21(DE3) or BL21(DE3) pLysS when plated on 0.4 mM IPTG. When induced or uninduced cultures of either strain were streaked out on plates containing OTC but no IPTG, the levels of resistance were lower. The maximum resistance levels observed for pJJ18 (Fig 5.5) were five times lower than for the pET21-b based OtrB constructs, pJJ5 and pJJ8/9 (4.5.2, 5.3.1) suggesting that the N-terminal HisTag might be affecting protein folding.

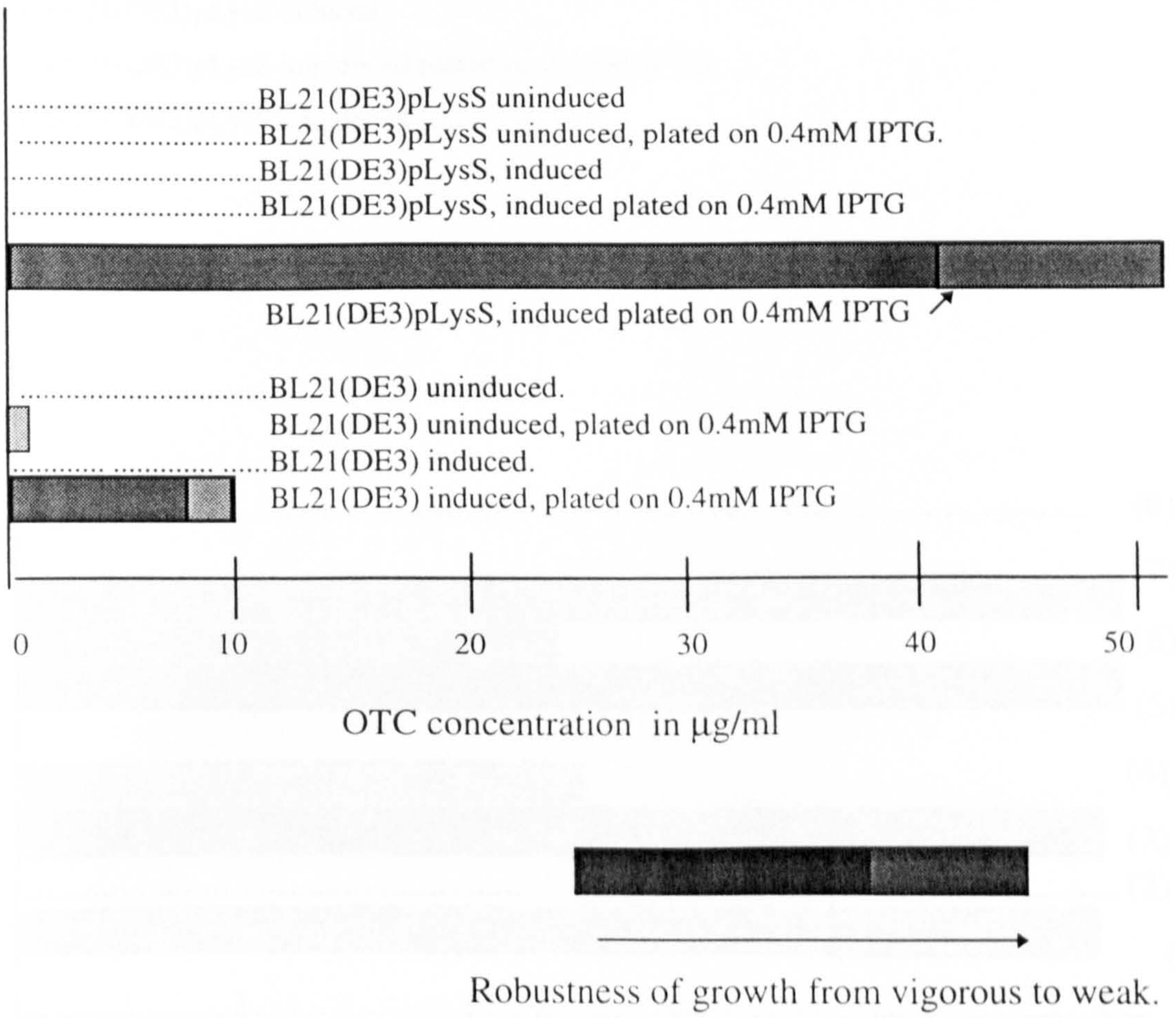



Fig. 5.4

Levels of OTC resistance conferred by plasmids pJJ8 and pJJ9

(1) BL21(DE3) induced plated on 0.4mM IPTG

(2) BL21(DE3) induced

(3) BL21(DE3) uninduced, plated on 0.4mM IPTG

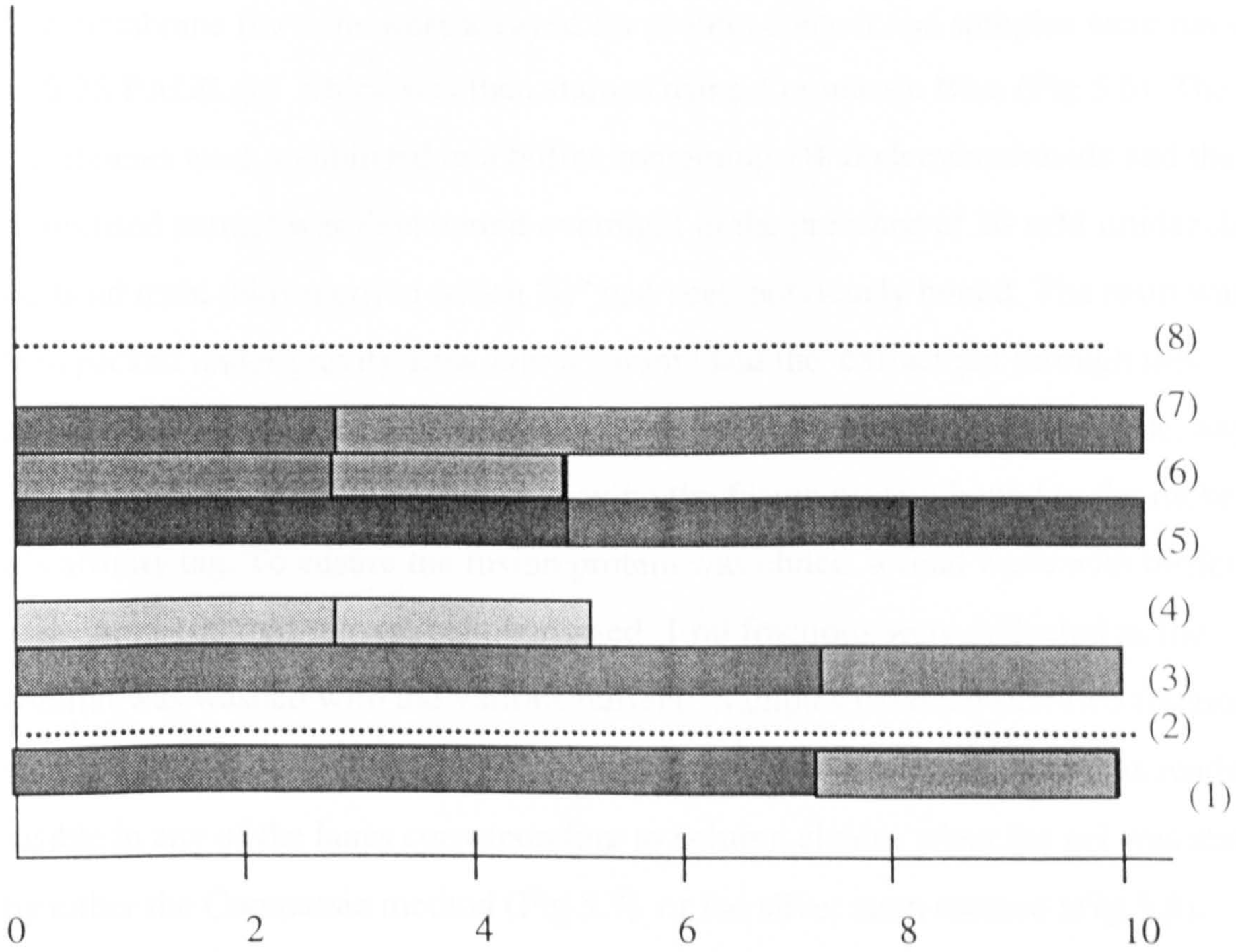
(4) BL21(DE3) uninduced

(5) BL21(DE3)pLysS induced, plated on 0.4mM IPTG

(6) BL21(DE3)pLysS induced

(7) BL21(DE3)pLysS uninduced plated on 0.4mM IPTG

(8) BL21(DE3)pLysS uninduced




..... Indicates no growth at 2.5 µg/ml OTC

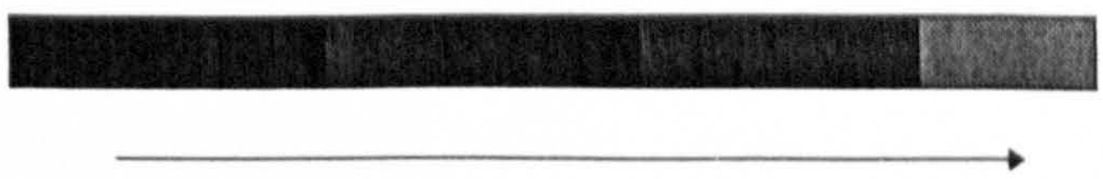
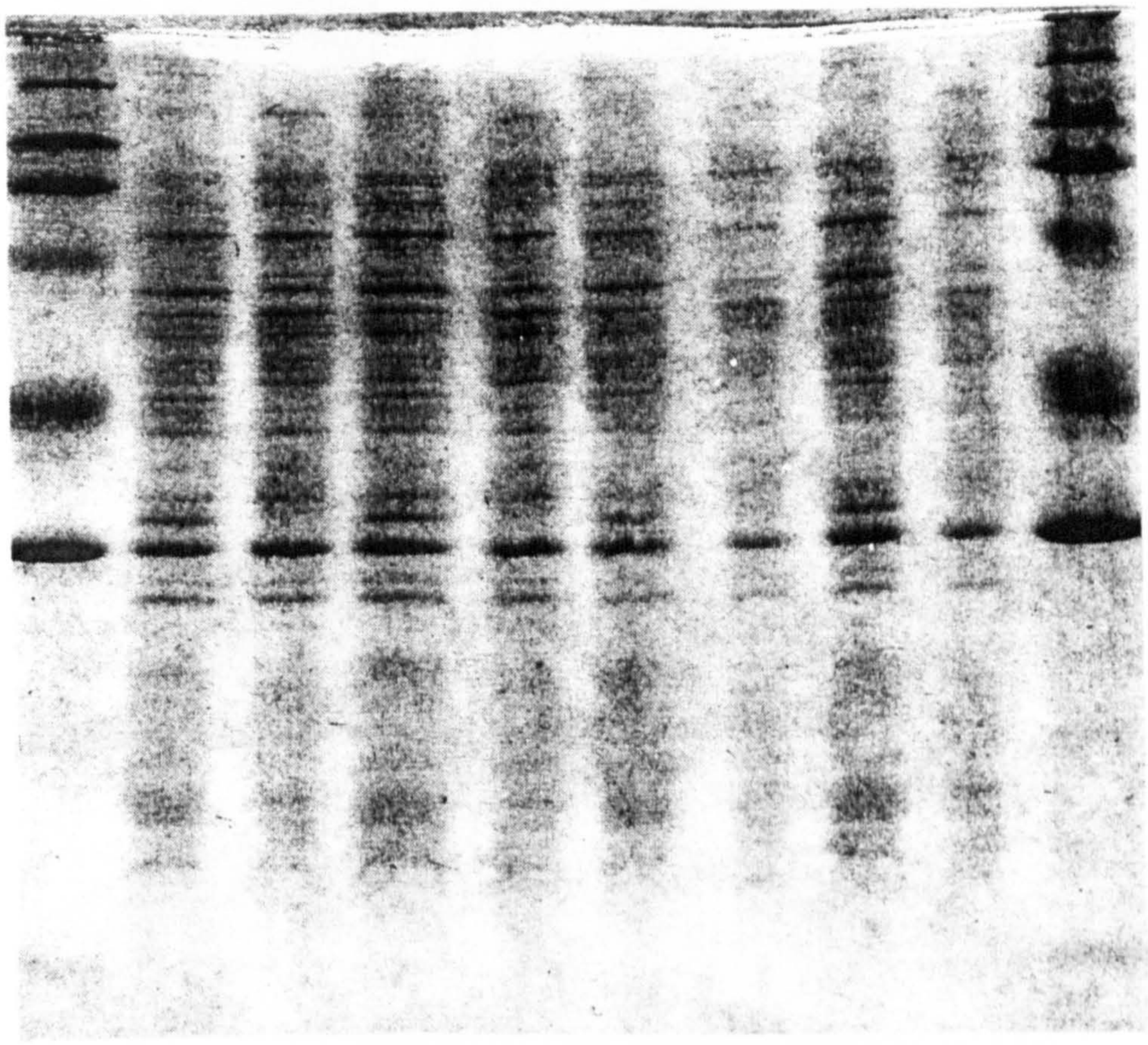



Fig 5.5

Levels of OTC resistance supplied by plasmid pJJ18

## 5.4 Attempts to isolate HisTagged fusion protein.

Isolation of the HisTagged fusion protein was attempted from BL21(DE3) / pJJ8. Although the pLysS strain had produced higher resistance levels, purification was not attempted from this strain. The T7 lysozyme produced by this strain causes the cells to lyse after thawing frozen pellets. Lysis produces a viscous sample that is unsuitable for use with the cell disrupter or French press. A large scale, (800 ml) induction run was carried out (2.3.6.1) and the membrane fraction prepared (2.5.2). The membrane fractions were assayed for protein content and samples were run on an SDS PAGE gel which was then stained using Coomassie Blue (Fig 5.6). The membranes were solubilised in a buffer containing 1% dodecylmaltoside and the solubilised extract was then bound overnight in the presence of 20 mM imidazole to HisBind resin (Novagen) to which $Ni^{2+}$ had been previously bound. The resin was then packed under gravity-flow into a column and the extract put through this column. By washing the column with increasing concentrations of imidazole, any unbound proteins should be eluted, leaving the fusion protein bound to the nickel by it's affinity tag. To ensure the fusion protein was eluted, a final wash with buffer containing 100 mM citrate, pH4 was used. 1 ml fractions were collected as the column was washed with the various buffers. Samples from the first two fractions were subjected to SDS PAGE, (2.5.5.3). No well over-expressed band was readily visible in any of the lanes corresponding to column eluates when the gel was stained by either the Coomassie method (Fig 5.7) or the silver stain method (Fig 5.8). Although this was the case, the OtrB fusion protein could well be present in low quantities. The initial gel (Fig 5.6) run using whole membrane fractions and stained by the Coomassie method had shown a number of bands present in the induced fractions which were not present in the non-induced samples.



7L Un Un Un Un Ind Ind Ind Ind 7L

Fig 5.6

SDS PAGE of solubilised whole membrane fractions from induced and uninduced (Ind, Un) samples of *E. coli* BL21 (DE3) transformed with pJJ8. Lanes 2-5 contain 10μl of membrane, lanes 6-9 contain 20μl of membrane.

Lanes 1 and 10 contain size markers at, 66,200 Da, 45,000 Da, 36,000 Da, 29,000 Da, 24,000 Da, 20,100 Da and 14,200 Da

The gel was stained by the Coomassie method (2.5.6).

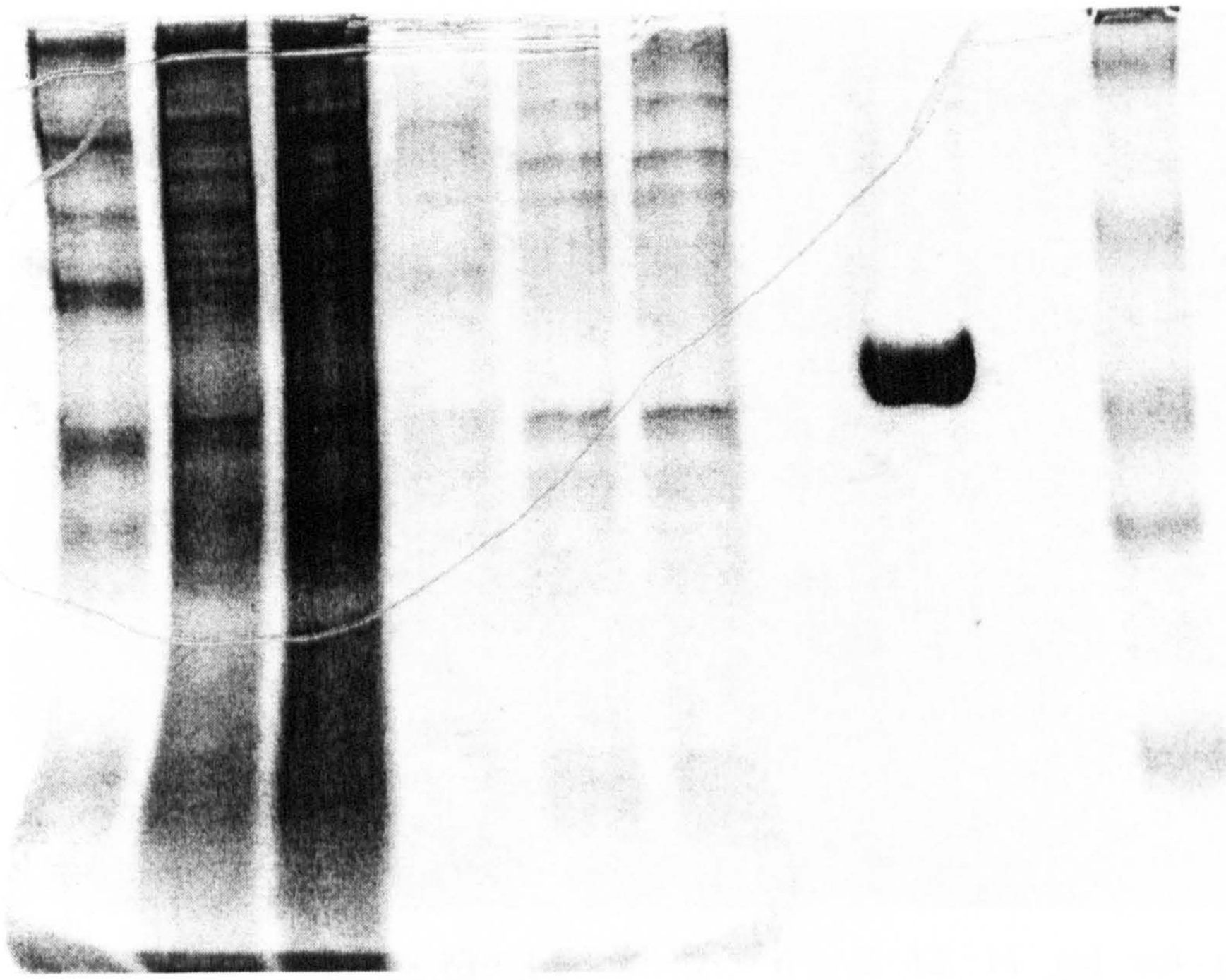



Fig 5.7

SDS PAGE of column eluates 1, 2 and 3 of OtrB-His from uninduced (lanes 1-3) and uninduced (lanes 4-6) samples of *E. coli* BL21(DE3) transformed with plasmid pJJ8. 15μl of each eluate were loaded per lane. Lane 8 contains 13.2μg AFAR-His as a positive control. Lane 10 contains size markers at 97,4000 Da, 66,200 Da, 45,000 Da, 31,000 Da, 21,500 Da,and 14,400 Da. The gel was stained by the Coomassie method (2.5.6).

Protein concentrations of the various eluates are given in Table 5.1.

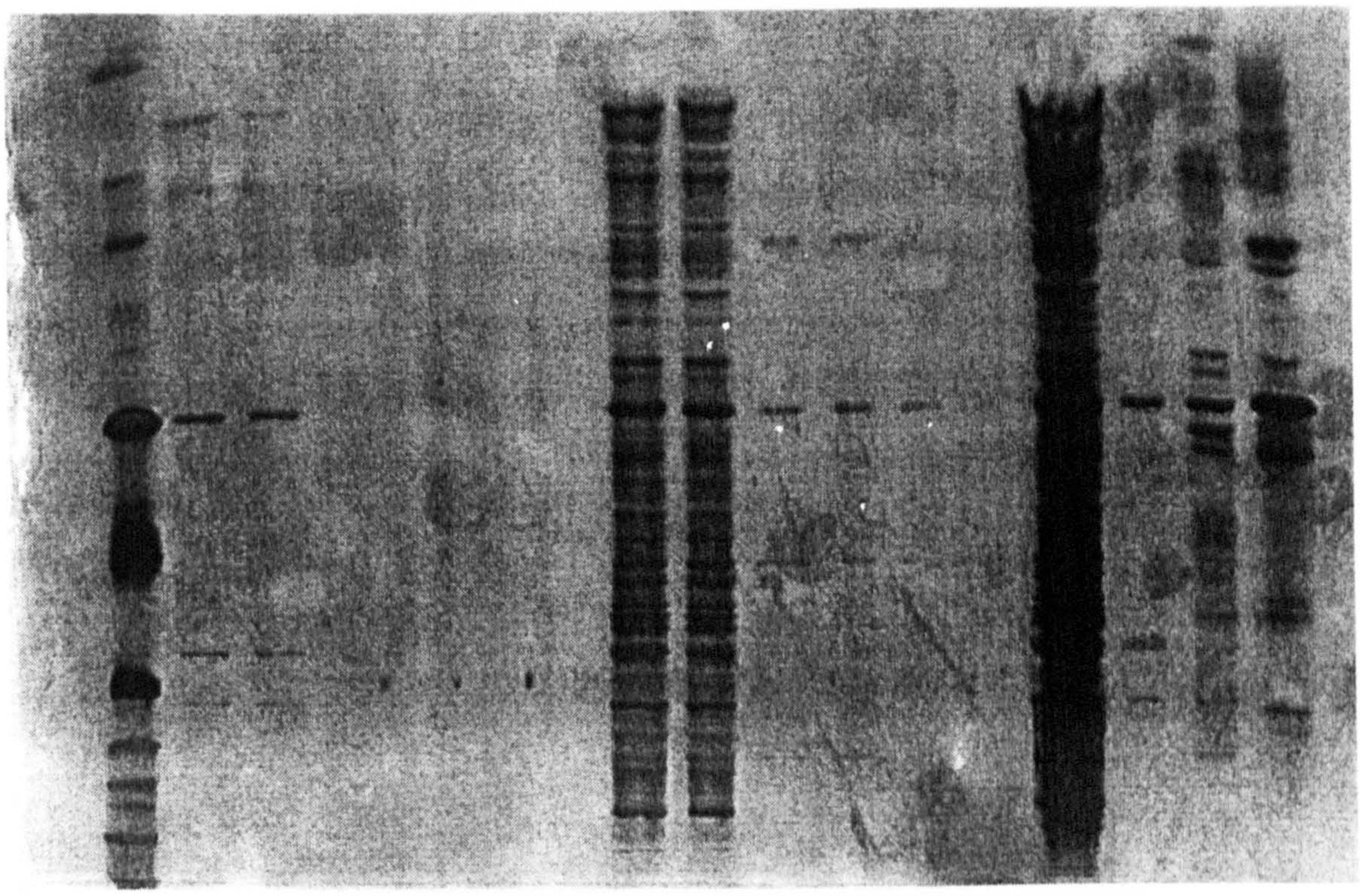



Fig 5.8

SDS PAGE of induced and uninduced column eluates from *E. coli* BL21(DE3) transformed with pJJ8. Underlined lanes indicate induced samples, those which are not underlined contain uninduced samples. Numbers refer to the fraction of eluate analysed. Lanes 14-16 contain either whole membrane fraction (sol) or a suspension of the pellet that remained after the solubilisation step (non). 10-15μg of protein was loaded per lane after BCA assay of the four samples.

Lane 1 contains size markers at 66,000 Da, 45,000 Da, 36,000 Da, 29,000 Da, 24,000 Da, 20,000 Da and 14,2000 Da.

The gel was stained by silver staining (2.5.6b)

Protein concentrations of the various eluates are given in Table 5.1

Table 5.1

Protein concentrations of column eluate fractions from induced BL21(DE3) /pJJ8

Determined by BioRad DC- Bradford assay.

| Fraction number | Wash buffer | µg /µl protein |
|---|---|---|
| 1 | 1 (20 mM imidazole) | 4.3 |
| 2 | 1 | 4.08 |
| 3 | 1 | 1.23 |
| 4 | 1 | 0.93 |
| 5 | 1 | 0.655 |
| 6 | 2 (200 mM imidazole) | 2.320 |
| 7 | 2 | 0.660 |
| 8 | 2 | 0.795 |
| 9 | 2 | 0.710 |
| 10 | 2 | 0.705 |
| 11 | 2 | 0.700 |
| 12 | 2 | 1.286 |
| 13 | 2 | not available |
| 14 | 3 (100 mM citrate) | 1.635 |
| 15 | 3 | 1.800 |
| 16 | 3 | 1.855 |
| 17 | 3 | 1.79 |
| 18 | 3 | 1.7 |

## 5.5 Assay for OtrB fusion using immunodetection

As the results of the affinity column did not demonstrate the presence of a purified fusion protein, the next step was to screen for OtrB-His using immunodetection techniques (Western blotting) as described in 2.5.7

Anti-His monoclonal antibodies raised against the hexahistidine tag (Qiagen, UK) were used to assay for low levels of OtrB-His that might be present in the eluates. These antibodies are available in three forms, 1) Anti- RGS-His, (this antibody is raised against proteins cloned using a Qiagen vector which encodes three amino acids, arginine, glycine and serine just prior to a HisTag. The "tail" is included to mobilise the HisTag in order that it may be well presented to the affinity resin) 2) Anti PentaHis and 3) Anti TetraHis. As it can not be predicted which of the three would give the best results, all three forms were used initially to assay for OtrB-His.

Samples of column eluates from the induced OtrB sample were subjected to SDS PAGE along with an N-terminal HisTagged positive control, Aflatoxin-Aldehyde-Reductase (AFAR-His) kindly donated by Dr. Elizabeth Ellis, University of Strathclyde. This control had been shown previously to be detected by the method described (E. Ellis pers. com.). After running the gel (Fig 5.6), the proteins were transferred to a nitrocellulose membrane and, after blocking, the membrane soaked in a solution containing the primary AntiHis antibody. At this stage, the antibodies should bind to any OtrB-His present on the membrane. Membranes were then washed to remove non-specific binding and soaked in a second antibody solution containing the secondary anti-mouse antibody ( Mannheim, Germany). The secondary antibody is conjugated to horseradish peroxidase, which in the presence of hydrogen peroxide, catalyses the oxidation of a diacylhydrazide substrate. An activated intermediate reaction product is formed which decays, emitting light. The light emission can be enhanced using 4-iodophenol which acts as a radical transmitter between the oxygen radical and the substrate. The presence of fusion protein is detected on photographic film as light emitted (Fig 5.9).

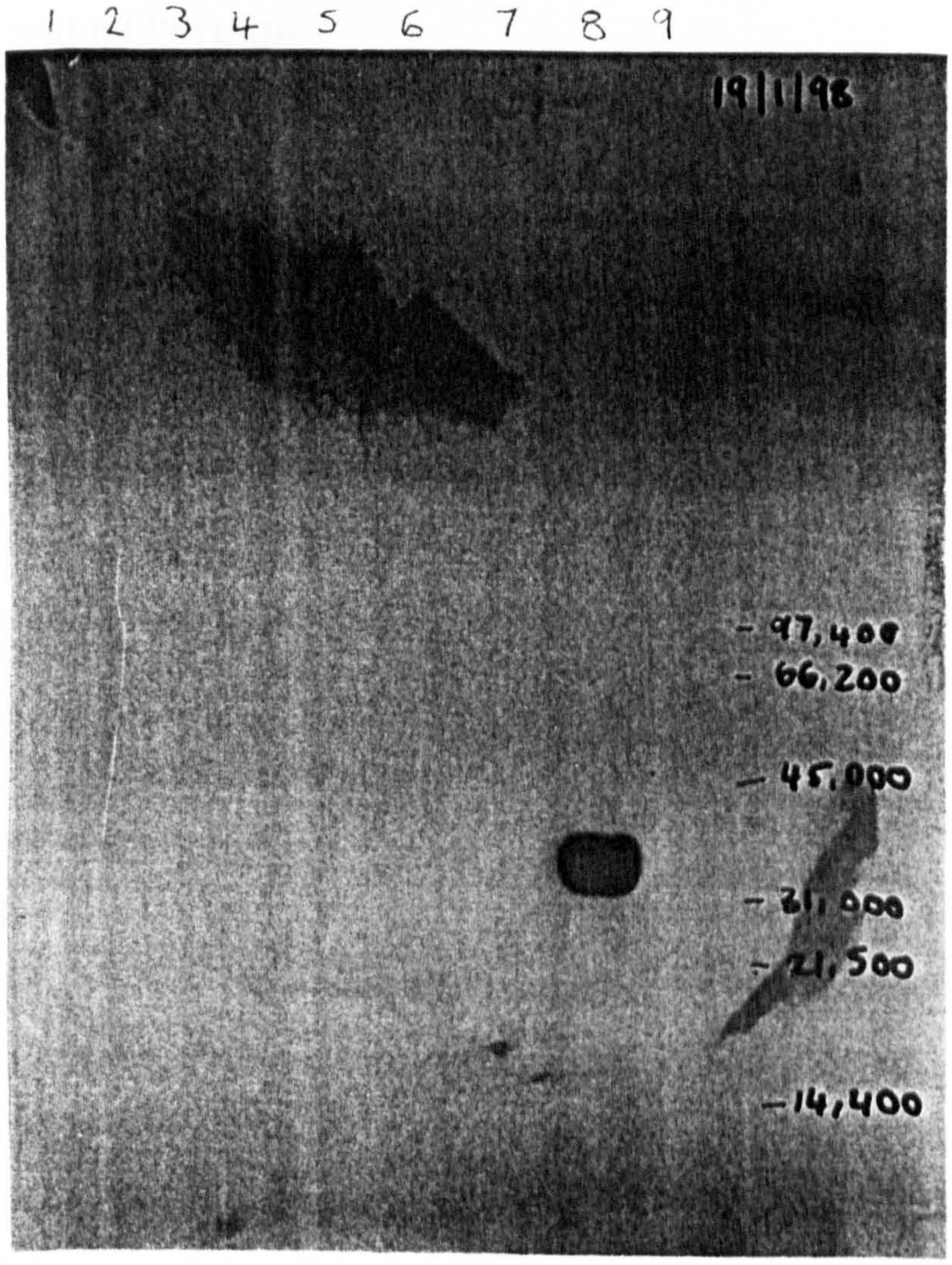



Fig 5.9

Fluorescent image of Western-blot (as gel shown in Fig 5.6) probed with Anti-His and Anti-mouse antibodies. Band visible in lane 8 is 4.4µg AFAR-His, the positve control. Lanes 7 and 9 were blank, lane contained markers, lanes 1-3 contained 10µl of fractions 1, 2, 3, from induced *E. coli* BL21(DE3)/ pJJ8 lanes 4-6 contained 10µl of corresponding fractions from uninduced sample.

Protein concentrations of the various fractions are given in Table 5.1.

## 5.6 Discussion

### 5.6.1 pJJ8 and pJJ9 results

The levels of expression of pJJ8 and pJJ9 were essentially identical, this indicates that the presence of the series of histidine residues at the C-terminus is not affecting the activity of the protein as it can still protect the cell from oxytetracycline. This implies that there are no problems with protein folding due to the presence of the HisTag. No difficulties were expected with the carboxy-terminal fusion.

The highest concentration of oxytetracycline at which strains containing these constructs would grow was 50 µg/ml for the pre-induced BL21(DE3) pLysS strain and 10 µg/ml for the pre-induced BL21(DE3) strain when both strains are growing on plates that contain 0.4mM IPTG in addition to the drug. The resistance of strains carrying pJJ5/ pJJ8/ pJJ9 to 50 µg/ml OTC is higher than resistance levels observed for the 12 transmembrane helix Tn*10* TetA protein (Walmsely pers. com.). This could feasibly relate to an increase in copy number, expression levels or activity of OtrB compared to TetA. The strains which were not preinduced are much less protected against oxytetracycline than their pre-induced counterparts, which shows that the control of gene expression is similar to the previous construct, pJJ5. Indeed, the plasmids, pJJ5 and pJJ8/9 have altogether similar resistance profiles. This is evidence that the PCR step which mutated the carboxy end of the gene in order to create the fusion product, did not introduce any other mutations to essential or important residues, i.e. the functional activity remains the same.

### 5.6.2 pJJ18 results

In direct contrast to the above findings, the results from the resistance tests of the induced pJJ18 strains show much lower resistance to oxytetracycline. The highest concentration at which growth was observed was 10µg/ml for both the pre-induced BL21(DE3) and BL21(DE3) pLysS strains, this is 20% of the resistance level of the carboxy tagged construct pJJ8, the control, pJJ9 and the original pET21b construct, pJJ5. The cultures of the two strains containing pJJ18 that were not pre-induced with

0.4mM IPTG in liquid culture show resistance at 10μg/ml which is a similar level to non pre-induced samples of the other three plasmids. The reason for the 80% reduction in resistance is presumed to be due to misfolding of the protein caused by the presence of the amino-terminal HisTag.

## 5.7 Deletion of the two putative central helices.

As previously discussed in Chapter 3, the OtrB protein is predicted to be formed from 14 transmembrane alpha helices connected by alternate cytoplasmic and periplasmic loops. The tetracycline resistance pumps from Gram-negative organisms have been shown to consist of only 12 such loops. The work of Yamaguchi and colleagues has given us much information about one such exporter, namely the TetA (B) / $H^+$ antiporter from transposon Tn*10* of *E. coli*. Conserved motifs have been identified which place Tn*10* in the Major Facilitator Superfamily as discussed in Chapter 3. Residues essential for tetracycline export have also been identified (Yamaguchi *et al,* 1990, 1992a, 1992b, 1993).

To investigate the importance of the two central helices present in the exporters of Gram-positive bacteria when expressed in a Gram-negative host, DNA encoding the two putative central helices from OtrB was deleted. The two halves of the gene were then joined using DNA encoding a large central cytoplasmic loop modelled on that of TetA(B). This "*E coli* friendly" protein was then expressed in the Gram-negative host and the resistance of cells expressing this mutated *otrB* gene was assessed.

## 5.8 Construction of truncated *otrB* gene

### 5.8.1 Construction of pJJ10 and pJJ11

To construct the shortened *otrB* gene, the two halves of the mutated gene were prepared separately by PCR based site directed mutagenesis. The template for both PCR reactions was plasmid pJJ7, (5.2.1).

For reaction I, the product of which was the DNA encoding the N-terminal half of OtrB, the primers used were universal forward primer and a custom reverse primer, primer 3 (2.4.14). The primer contains *EcoR*I and *Cla*I sites, and introduces

For PCR reaction II, which produced the DNA encoding the C-terminal half of the truncated OtrB, the primers used were universal reverse and a custom forward primer, Primer 4. This primer contains *Cla*I and NdeI sites and introduces two amino acids into the new cytoplasmic loop. The 0.5kb product of PCR reaction I and the 0.8 kb product of reaction II were cut with *Nde*I and *Eco*RI. The fragments were then ligated separately with pMTL23 that had also been restricted with *Nde*I and *Eco*RI. The ligation products were pJJ10, which contained PCR product I and pJJ11, which contained PCR product II ( Fig. 5.10).

### 5.8.2 Design of PCR primers

When designing PCR primers to create the cytoplasmic loop between the two halves of the OtrB deletion mutant, the basis of design was the TetA protein from transposon Tn*10* (Fig 5.11). This middle cytoplasmic loop consists of 33 amino acids, which makes it considerably longer than the other cytoplasmic and periplasmic loops in the protein, a longer middle loop appears to be characteristic for many members of the Multi Facilitator Superfamily.

In the *otrB* gene, the fragment of DNA encoding the two middle transmembrane helices, helices 7 and 8, begins at base 571 (G) and ends at base 781 (C). It was this DNA that was deleted to make the truncated *otrB* gene. In order to join the two PCR products a new cytoplasmic loop was constructed using the "tails" of the PCR primers As already stated, the TetA central cytoplasmic loop consists of 33 amino acids. After helices 7 and 8 have been removed from the *otrB* DNA, 28 residues remain to make up the new cytoplasmic loop, 11 from loop 6-7 and 17 from loop 8-9. Five extra amino acids were added by incorporating their codons into the primer tails, tails also had to contain restriction sites to clone them initially as two separate parts and then join them.

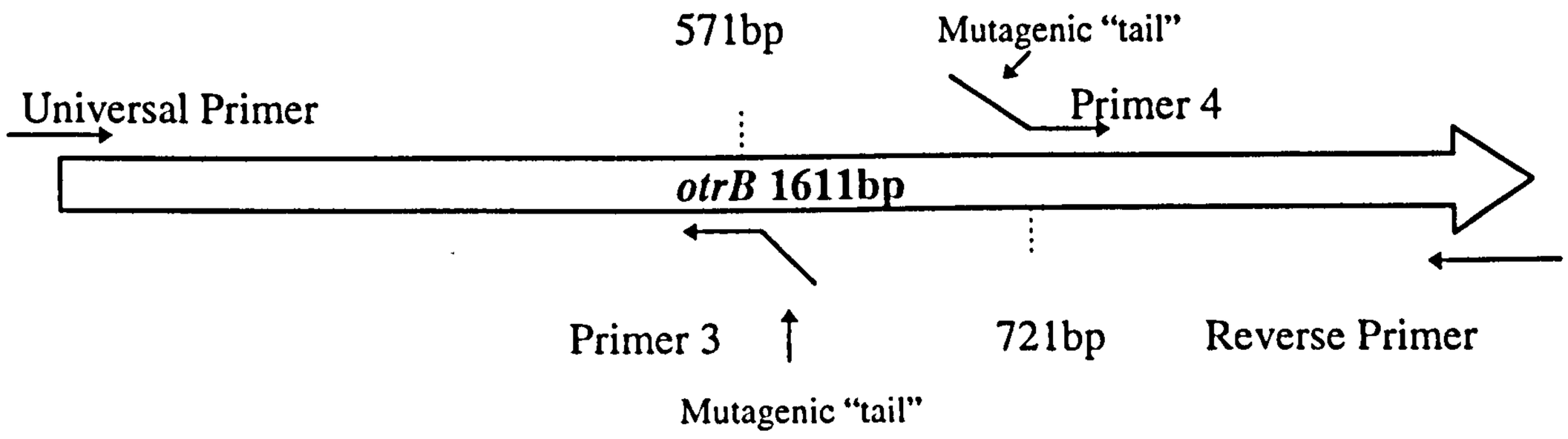
The middle five amino acids of TetA are, valine (V), glycine (G), valine (V), glutamic acid (E) and threonine (T). The DNA sequence which encodes these amino acids was incorporated into the design of the primers, taking account of the high G + C content, around 70%, of the Streptomycete genome. This was in order with the long-term aim to express *otrB* from an expression vector in *Streptomyces*. The codon

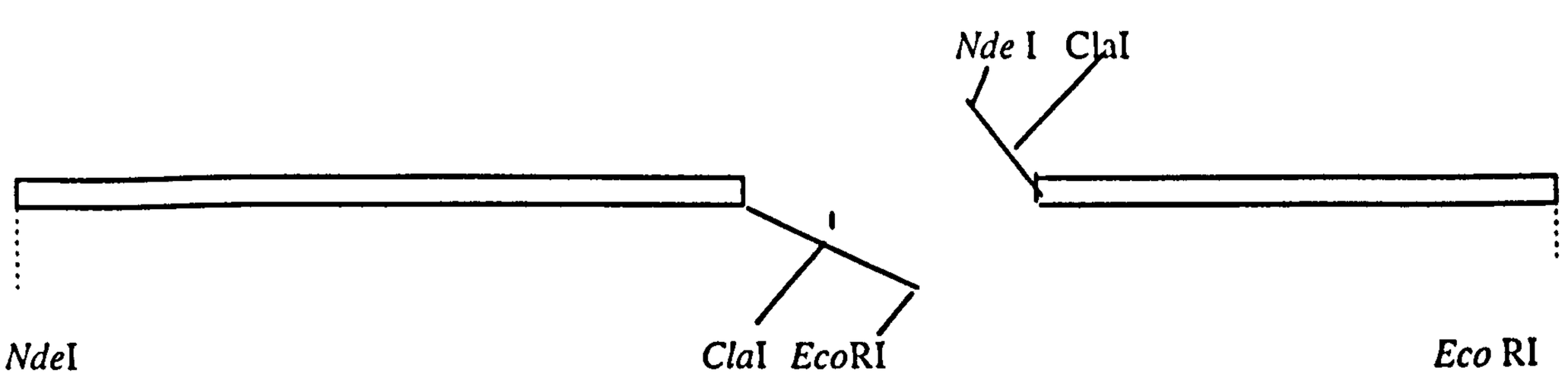
long-term aim to express *otrB* from an expression vector in *Streptomyces*. The codon bias figures used were compiled from 67 *Streptomyces* genes by Taylor, (1992) (Table 5.2).

From these figures, the DNA sequence, GTC GGC GTC GAG ACC was chosen to encode these amino acids in the primers, However, this sequence does not contain a suitable restriction site to join together the two PCR products. By changing the valine to isoleucine (I) and the glutamic acid to aspartic acid (D), the sequence is changed to GTC GGC ATC GAT ACC which introduces a *Cla*I site, (underlined) whilst retaining the character of the amino acids. Valine and isoleucine both have non-polar side chains, aspartic acid and glutamic acid both have negatively-charged side chains.

## 5.9 Construction of pJJ12, pJJ14 and pJJ17.

The N-terminal fragment of *otr*B was subcloned into pET21b from pJJ10 (5.8.1) to create pJJ12 (Fig 5.12). pJJ10 was restricted with *Nde*I and *Eco*RI and ligated with similarly cut pET21b. pJJ12 was then restricted with *Cla*I and *Eco*RI and ligated with the C-terminal fragment of *otr*B which had been removed on *Cla*I, *Eco*RI ends from pJJ11 (5.8.1). The product of this ligation was named pJJ14 (Fig 5.13). pJJ14 was therefore a pET21b-based plasmid which contained the truncated *otr*B gene. In pJJ14, the fragment containing *otrB* was cloned using *Eco*RI at the C-terminal end, thus the insert had a *Sac*I site, a stop codon and an *Xho*I site upstream of the vector DNA encoding the HisTag. Downstream of the *Eco*RI site, the vector polylinker contains another *Sac*I site. By cutting pJJ14 with *Sac*I and re-ligating, the small *Sac*I fragment was removed. This creates a fusion with the HisTag, pJJ17 (Fig 5.13). To demonstrate that this fragment had been removed, pJJ17 was restricted with *Eco*RI. As this site had been removed, the plasmid remained uncut. As a control, pMTL23 was included in the restriction mixture, this linearised with *Eco*RI, showing that the enzyme was active (Fig 5.14) and that the *Eco*RI site in pJJ17 had been removed.




PCR product 1

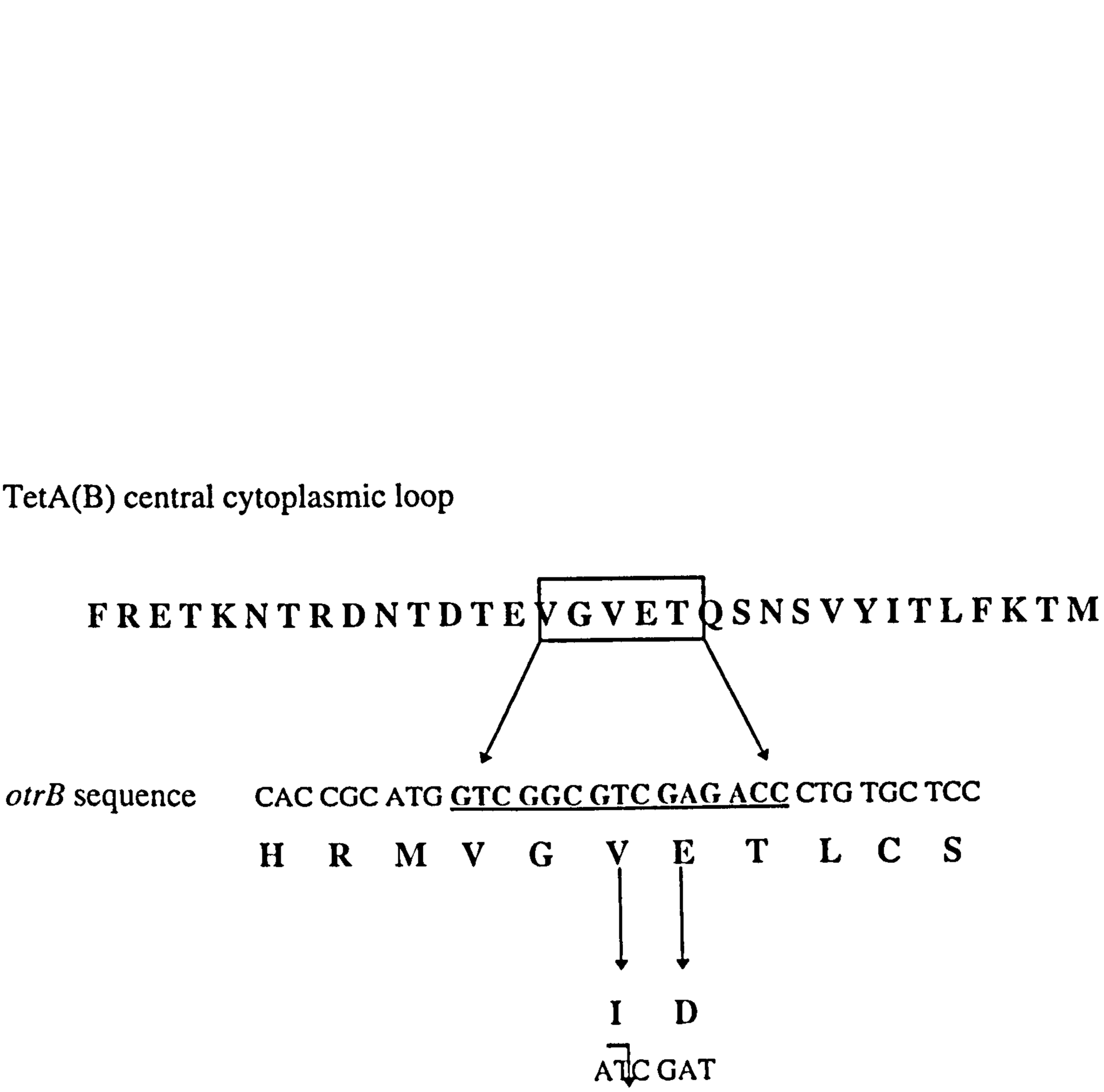
Encodes TMH 1-6

PCR product 2

Encodes TMH 9-12

Fig 5.10

Positioning of mutagenic, Universal and Reverse primers on the *otrB* template and the products of the PCR reactions.




Peptide sequence of new OtrB central cytoplasmic loop.

Fig 5.11

Cytoplasmic loop joining TMH6 and the six C-terminal TMH's of OtrB. showing the peptide sequence of TetA(B) central cytoplasmic loop, (top line), with the 5 residues to be conserved in the new OtrB loop shown within the box . The codons chosen from the codon bias table are shown in bold type on the second line with original *otrB* nucleotides in normal type. The two residues substituted to introduce the restriction site are shown by arrows along with the DNA that encodes them. The bottom line shows the final cytoplasmic loop introduced into OtrB in place of the deleted region.

Table 5.2

Codon preference table for 67 *Streptomycete* genes

| | G | | A | | T | | C | | |
|---|---|---|---|---|---|---|---|---|---|
| **G** | Gly | 19 | Glu | 81 | Val | 39 | Ala | 33 | **G** |
| | Gly | 10 | Glu | 19 | Val | 2 | Ala | 4 | **A** |
| | Gly | 8 | Asp | 4 | Val | 4 | Ala | 3 | **T** |
| | Gly | 64 | Asp | 96 | Val | 56 | Ala | 59 | **C** |
| **A** | Arg | 6 | Lys | 95 | Met | 100 | Thr | 30 | **G** |
| | Arg | 1 | Lys | 5 | Ile | 3 | Thr | 3 | **A** |
| | Ser | 3 | Asn | 4 | Ile | 5 | Thr | 2 | **T** |
| | Ser | 28 | Asn | 96 | Ile | 92 | Thr | 65 | **C** |
| **T** | Trp | 100 | END | 14 | Leu | 3 | Ser | 26 | **G** |
| | END | 83 | END | 3 | Leu | 0 | Ser | 3 | **A** |
| | Cys | 13 | Tyr | 5 | Phe | 1 | Ser | 1 | **T** |
| | Cys | 87 | Tyr | 95 | Phe | 99 | Ser | 39 | **C** |
| **C** | Arg | 36 | Gln | 93 | Leu | 55 | Pro | 52 | **G** |
| | Arg | 5 | Gln | 7 | Leu | 0 | Pro | 2 | **A** |
| | Arg | 7 | His | 6 | Leu | 2 | Pro | 3 | **T** |
| | Arg | 46 | His | 94 | Leu | 39 | Pro | 43 | **C** |

N.B. a) The first nucleotide of each codon is provided in the left hand column; the second nucleotide is shown in the top row; the third nucleotide is shown in the right hand column.

b) The figures show the relative codon uses (expressed as a percentage of each nucleotide triplet) in encoding each amino acid in the 67 genes analysed.

Data taken from Taylor, 1992.

## 5.10 Expression and Resistance results

### 5.10.1 Expression of pJJ14 and pJJ17

Plasmids pJJ14 and pJJ17 were transformed separately into *E. coli* BL21(DE3) and BL21(DE3) pLysS. Induction with IPTG was carried out as for other strains (2.3.6.1). Cultures were streaked onto agar plates containing increasing concentrations of oxytetracycline, both with and without IPTG. Controls were the same as in previous induction experiments. Untransformed and vector-only strains did not grow in the presence of 2 µg/ml OTC.

Table 5.3

Results of pJJ14 and pJJ17 resistance tests.

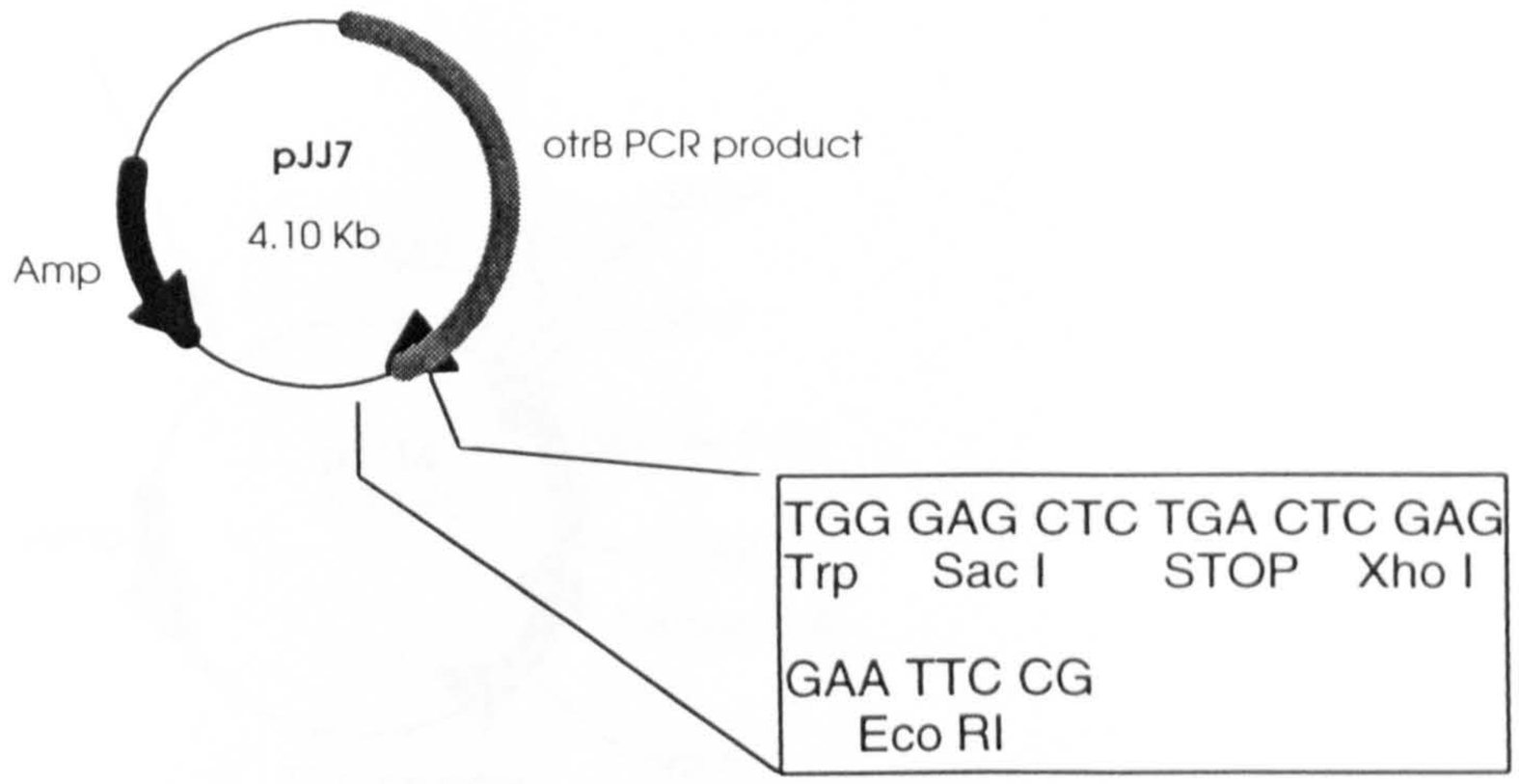
| OTC µg/ml | BL21(DE3)/pJJ14 or pJJ17 | | | | BL21(DE3) pLysS/pJJ14 or pJJ17 | | | |
|---|---|---|---|---|---|---|---|---|
| | pre-induced | | uninduced | | pre-induced | | uninduced | |
| | IPTG | no IPTG | IPTG | no IPTG | IPTG | no IPTG | IPTG | no IPTG |
| 0 | + | + | + | + | + | + | + | + |
| 3.0 | - | - | - | - | l.g. | l.g. | l.g. | l.g. |
| 5.0 | - | - | - | - | - | - | - | - |
| 10.0 | - | - | - | - | - | - | - | - |

+ indicates growth

- indicates no growth

l.g. indicates light growth

pJJ14 and pJJ17 exhibited similar oxytetracycline resistance profiles, both displaying resistance to a maximum of 3µg/ml OTC. No difference between the two strains was expected as the results for expression and resistance of earlier plasmids demonstrate that the HisTag does not appear to make any difference to the OTC resistance of the host. Resistance to OTC was only observed when the plasmids were expressed in *E. coli* BL21(DE3) pLysS. Strains lacking the pLysS plasmid could not grow in the presence of 2.0µg/ml OTC as is the case for the controls. The OTC resistance exhibited by the truncated protein was not increased by induction with IPTG during




PCR using 2 sets of primers, universal + Primer 3 and reverse + Primer 4
Each primer has a tail designed to mimic the central cytoplasmic loop from TetA(B)
Clone the two products separately into pMTL23 to create pJJ10 and pJJ11
Suclone from these plasmids into pET21b to create pJJ12 and pJJ13

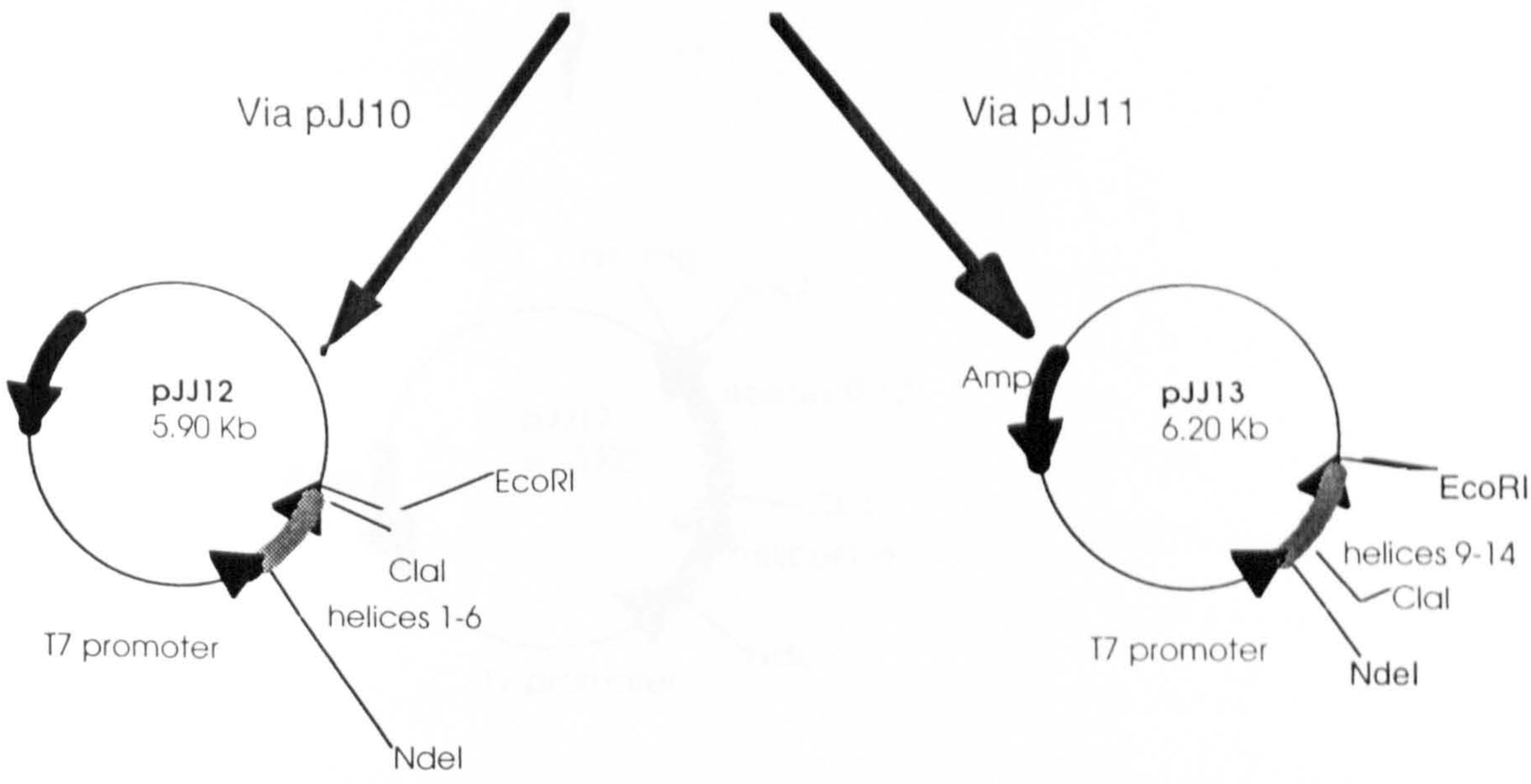
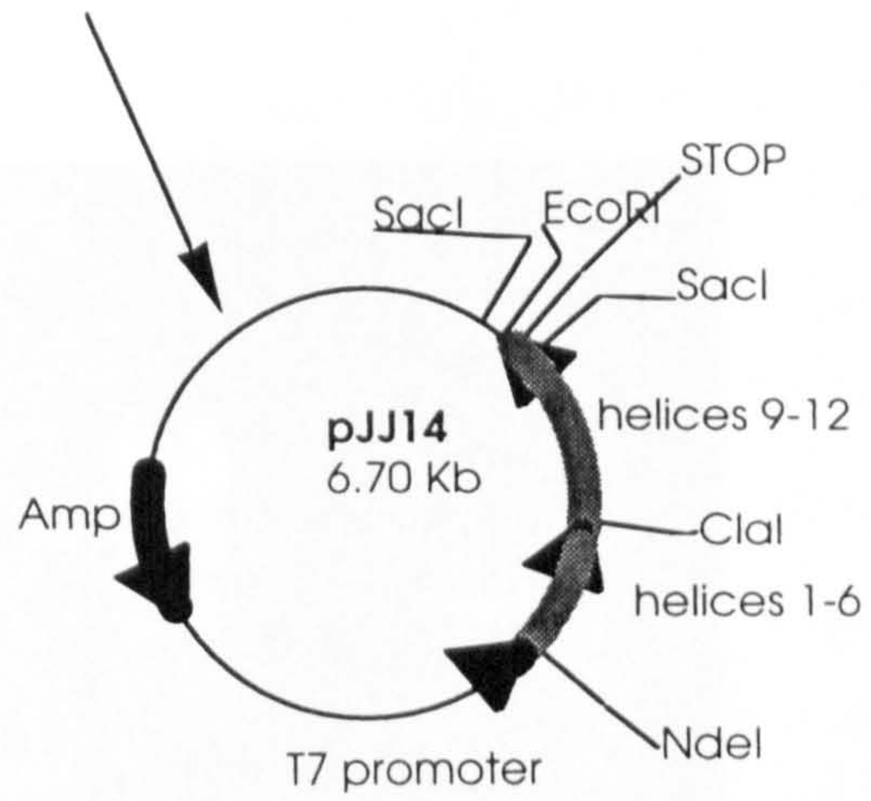



Fig 5.12
Construction of pJJ12 and pJJ13

Cut both pJJ12 and pJJ13 with ClaI & EcoRI. Ligate large fragment of pJJ12 to 0.8kb
fragment of PJJ13 to create pJJ14




Restrict withSacI and re-ligate
in order to remove STOP codon
and create polyhistidine fusion

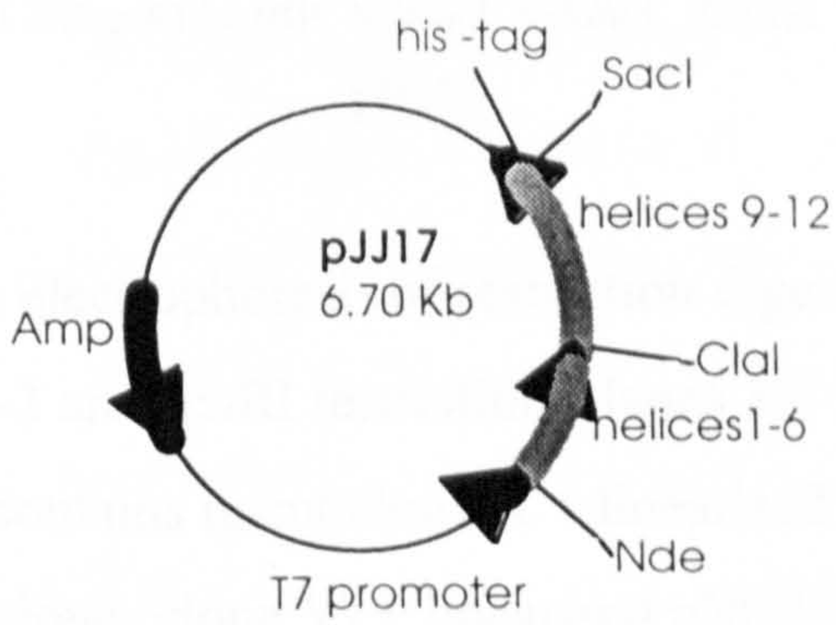
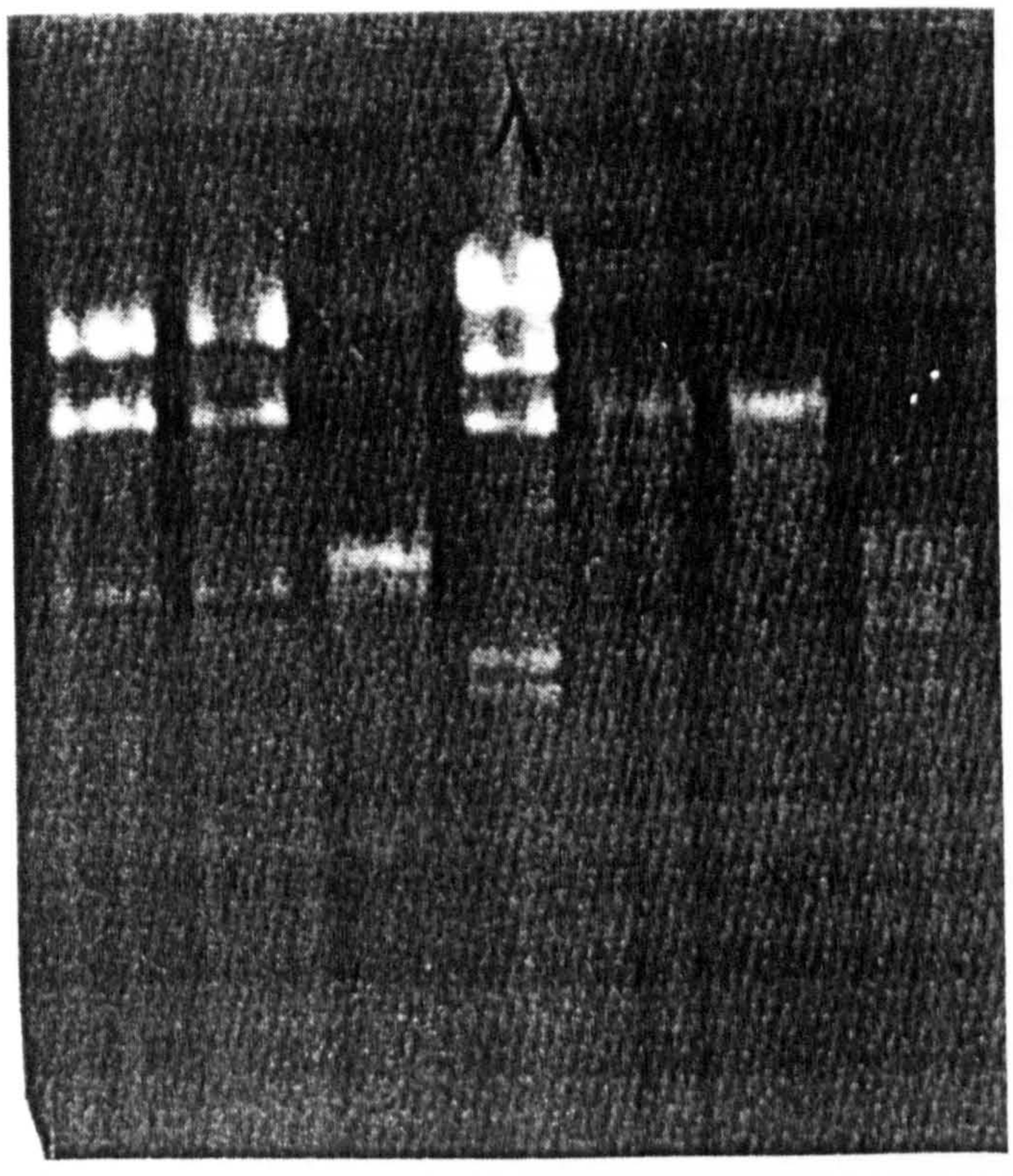



Fig 5.13

Construction of pJJ14 and pJJ17



X/*Eco*RI Y/*Eco* RI Z/*Eco* RI λ/ *H*III X/*Nde* I Y/*Nde* I Z/*Nde*I

pJJ17 pJJ17

Fig 5.14

Agarose electrophoresis of restriction digests to screen for pJJ17.

Lanes 1-3 are *Eco*RI restrictions, lanes 5-7 are *Nde* I restrictions.

Lane 1 contains uncut clone X + linearised pMTL23, lane 2 shows a second similar clone (clone Y) + linearised pMTL23, lane 3 shows plasmid from a third transformant(clone Z) + uncut pMTL23, Restriction of the pMTL23 control has not occurred in this sample so no conclusions can be drawn. Lane 4 contains λ/ *Hind III* size markers. Lanes 5-7 contain *NdeI* digests of clone X, clone Y and clone Z. Clone Z DNA was of poor quality, did not restrict with either enzyme and was discarded, clones X and Y showed no restriction with*EcoRI* . Clone X was named pJJ17.

growth of the inoculum and no difference was observed between samples plated in the presence or absence of 0.4 mM IPTG.

## 5.11 Discussion of resistance supplied by truncated OtrB

The OTC resistance exhibited by the two BL21(DE3) pLysS strains containing pJJ14 or pJJ17 is low, ca. 4% of that shown by plasmid pJJ8 in which the full length *otrB* gene fused to the polyhistidine tag is expressed. Nevertheless, some resistance is provided by the construct as cells containing no plasmid and those transformed with vector only are unable to survive at a lower concentration of 2 µg/ml OTC. Although resistance is not completely lost, the deleted region thought to contain transmembrane helices 7 and 8 must be important to the function of the protein as the greatly reduced resistance levels demonstrate.

It is highly probable that the reduced resistance activity is attributable to differences in the way the protein folds and assembles in the membrane when it is truncated in this way. What cannot be deduced from this study is whether the attempts to mimic the loop character of the Gram-negative TetA(B) protein make any positive contribution to the resistance levels in these strains or whether merely maintaining the length of the loop at 33 amino acids would have resulted in similar OTC resistance. The importance of the length of cytoplasmic loops is not known as regards insertion into the membrane. Although, length of periplasmic loops is thought to be important. Proteins containing cytoplasmic loops shorter than c.25 amino acids are thought to be able to insert into the inner membrane independently of the cellular *Sec* system (Kuhn, 1988, Andersson & Von Heijne, 1993). The *Sec* system is the machinery that enables secretion of certain proteins out of the cell. (Schatz & Beckwith, 1990). *Sec* independent proteins may insert into the membranes by a method known as the "helical hairpin" method, the helical hairpin structure being composed of a pair of anti-parallel alpha helices from neighbouring odd and even-numbered transmembrane regions. The reason for formation of such structures is thought to be a thermodynamic one as formation may increase the overall hydrophobicity of the sequence sufficiently enough to drive polar loop residues across the lipid bilayer and provide a way of masking interior hydrophilic residues via association of relatively polar faces of the helices (Engelman & Steitz, 1981). The

TetA(C) protein from pBR322 has been indicated to insert by this mechanism (Guo *et al*, 1996).The work presented by Guo and co-workers also shows that efficient insertion of odd numbered transmembrane helices requires the presence of the next even numbered segment.

As both the truncated and wild type form of OtrB contain a large putative periplasmic loop between transmembrane regions 13 and 14 it is unclear which method of insertion into the membrane is used. Proteins may insert by utilising both systems. *Lac* permease, a 12 segment membrane protein can assemble in the membrane as an active protein when contiguous fragments separated within periplasmic loops 1 and 4 are co-expressed in *E. coli* ( Zen *et al*, 1994). Possibly OtrB inserts in a similar manner as some transmembrane segments require the *sec* machinery whilst others form helical hairpins. The periplasmic loops of OtrB have not been affected by the deletion apart from the absence of loop 7-8. The loops that are present are intact and have not been mutated in any way. Insertion by helical hairpin formation should remain unaffected in the truncated protein. However, the deletion of TMH7 may affect the folding of the following six helices. It is thought that the N-terminal half of TetA(B) assembles in the membrane independently of the C-terminal half whilst the insertion of the C-terminal half is affected by the presence of the N-terminal portion (Yamaguchi *et al*, 1993). Vesicles containing the two halves of TetA(B) expressed separately, showed about 40% of the transport activity per C-terminal as those containing the wild type protein. These results suggest that the C-terminal half of the protein is important for correct insertion of TetA(B) into the membrane. This may also be the case in OtrB, removing TMH7 which is part of its N-terminal half could well result in incorrect folding of the C-terminal portion. Although the presence of the N-terminal region is important for insertion of the C-terminal, it is not essential as some C-terminal fragments were detected in the membrane. This is in contrast to similar work carried out using the *lac* permease where the presence of the C-terminal fragment was not detected in the membrane in the absence of the N-terminal fragment (Bibi & Kaback, 1990). It can be concluded that the low OTC resistance of *E. coli* strains expressing the truncated OtrB protein is probably due to misfolding of the C-terminal portion of the protein.

Yamaguchi *et al* ( 1993) also found that vesicles containing only one of the two

halves of TetA(B) had no active transport activity. This is also thought to be the case with OtrB as when pJJ12 or pJJ13 were expressed in BL21(DE3) or BL21(DE3) pLysS the level of OTC resistance observed was identical to that of untransfomed and vector-only controls.

## 5.12 Lack of detection of OtrB-His

Attempts to isolate OtrB-His by the NTA-nickel agarose column were assumed to be unsuccessful. This conclusion was reached after no over-expressed band was seen on SDS-PAGE gels of the membrane fractions or the column eluates and no signal was observed when a Western blot of column eluates was probed with mouse Anti-His monoclonal antibody and fluorescent anti-mouse IgG. The lack of a signal from the immunoblot could possibly have been due to the sensitivity of the system being too low to detect the level of OtrB present. The control protein, AFAR-His was detected from dot blots when a dilution of column eluate containing 125 ng of total protein was applied to the membrane. The AFAR eluate showed AFAR as the only major band when subjected to SDS-PAGE (E. Ellis pers. com). Column eluates of OtrB-His preparation contained between 0.6 and 4.3 µg/µl total protein, meaning that µg quantities of protein were loaded onto the gel. It is expected that one or two of the eluates would contain a very high proportion of OtrB-His. Initial thoughts as to the lack of detection are discussed below.

As OtrB is very hydrophobic it may be that the conformation of the protein in the membrane preparation is such that the C-terminal HisTag is not presented appropriately to the NTA-nickel resin. This is however an unlikely explanation as the membrane preparation and all subsequent steps in the purification procedure were carried out in denaturing conditions (in the presence of 8 M urea). The OtrB-His may not have bound or may have only weakly bound to the nickel and so may have been removed from the column by gravity flow as the agarose settled to form a compact bed. A small fraction (1 ml) of this flow-through was collected for analysis but unfortunately sample ran out before all analyses were performed, no sample was available for immunodetection. As the binding step took place in a 50 ml volume, the concentration of OtrB-His that may have been present in the sample would have been extremely low. Another possible explanation for the lack of detection of OtrB-His is

that the protein was cleaved or degraded by *E. coli* cell machinery or by a subsequent step in the procedure. As the HisTag system had been successful in isolating other membrane proteins from *E. coli* e.g. TetA(B) (Aldema *et al*, 1996), degradation may seem an unlikely explanation. However, it must be remembered that the TetA(B) protein originates from *E. coli* and that OtrB is a heterologous protein.

## 5.13 Further sequencing of pJJ8

When no over-expressed band could be detected using SDS PAGE it was suspected that constructs pJJ8 and pJJ9 (containing the TGA stop codon and therefore not tagged) had been confused with each other resulting in transformation and expression of the control construct. For this reason pJJ8 was sequenced in from the vector DNA using T7-terminator primer (5'GCTAGTTATTGCTCAGCGG3'). The nucletotide sequence data demonstrated the presence of a *Sac*I site but not an *Eco*RI site, therefore the *otrB/* vector junction was presumed intact.

After further attempts to detect OtrB-His using immunoblotting techniques had also failed the construct was resequenced. The resulting chromatograms were of much higher quality and more bases could be read. As a result of this, a C residue was discovered at position 1469 which was assumed to be a G. This difference from the Reynes/ MacGregor-Pryde sequence was either due to a mistake in both of these sequence or to an error by the *pfu* polymerase. If the former was the case other differences might also be present. To establish whether this inversion was the only deviation from the Reynes / MacGregor-Pryde sequence in the latter part of the gene, four sequencing primers were designed against *otrB* ( primers 4-7 Fig 3.7). These primers would sequence the region of *otrB* not previously sequenced using the ABI system. The original copy of *otrB* cloned into pJJ1 was used as the template in the sequencing reaction.

### 5.13.1 Discussion of T7 terminator primer sequencing results

On analysis of the sequence data from T7 terminator primer, the G to C inversion at position 1469 was found to be the only difference from the previous sequence, meaning that the reading frame would remain intact

However upon translation of the new sequence into peptide it was discovered that no

polyhistidine tail was present. The OtrB sequence was intact until the mutated tryptophan (codon TGG) residue that replaced the wild-type TGA stop codon. Sequence downstream of this position translated as out of frame vector sequence, resulting in a poly-proline tail (Fig 5.15). This was due to an error in the design of the reverse PCR primer. pET21a, pET21b and pET21c are related vectors differing from each other by either one or two nucleotides to allow phasing of the cloned gene. In the cloning strategy of pJJ8, a mistake was made when designing the inverse PCR primer (Primer 2), meaning that the *Sac*I site was not in frame with the HisTag-encoding DNA. In order to have built an expression construct in frame with a polyhistidine tag in pET21b, the primer should have been one base shorter, with the last G of the TGG tryptophan codon being the first G of the GAGCTC *Sac*I site.

i)

*Sac I*

**TGGGAGCT↓CCGTCGACAAGCTTGCGGCCGCACTCGAGCAC**

TrpGluLeuArgArgGlnAlaCysGlyArgThrArgAlaP

SerSerValAspLysLeuAlaAlaAlaLeuGluHis

**CACCACCACCACCACTGAGATTCGGCTGCTAA** a) DNA sequence

roProProProProLeuArgSerGlyCysEnd b) Actual protein seq.

HisHisHisHisHisEnd c) Expected protein sequence

ii) 5'CGG AAT TCA AGA GCT CCA GGC GTC CGA CGC GGG CAG CGG CCC G3'

Fig 5.15

I) Sequence of polypeptide tag fused to OtrB, from pJJ8 DNA sequenced using T7-terminator primer.

ii) DNA sequence of new PCR primer that will enable correction of fusion, stop codon is shown in bold type. *Eco*RI and *Sac*I sites are underlined

## 5.14 Future Work

1) The most important future work is obviously to re-construct pJJ8 using a PCR primer that results in translation of the His-tagged product. The expression and purification can then be repeated and hopefully OtrB will be isolated.

2) It would be interesting to express the truncated *otrB* gene in *S. albus* in order to assess the importance of these two TMH's in the Streptomycetes, as all MFS proteins isolated so far from these organisms are comprised of 14 TMH's.

3) In order to investigate the importance of residues within the central cytoplasmic loop, mutagenesis of individual residues should be carried out to change the character of the loop.

4) The effect of loop length should be investigated by studying the resistance levels of a series of deletion mutants.

# Chapter 6
# Concluding Remarks

## 6.1 Introduction

The aims of the work described in this thesis were;

**1)** To introduce a plasmid-born *S. rimosus otrB* gene into two non oxytetracycline producing Streptomycetes and investigate the level of OTC, CTC and TET resistance supplied by the gene.

**2)** To express *otrB* in *E. coli* using a well characterised, inducible expression system and investigate OTC resistance levels conferred on transformed cells.

**3)** To construct, express in *E. coli* and isolate a polyhistidine-OtrB fusion protein.

**4)** To produce a model for the 2-dimensional topology of OtrB using computer programs to analyse hydrophobicity and hydrophilicity.

**5)** To express a mutant OtrB protein from which the two central transmembrane helices had been deleted in *E. coli* and investigate the potential of this protein to confer OTC resistance upon *E. coli*.

During the course of these studies it also became necessary to determine the definitive nucleotide sequence of *otrB*.

## 6.2 Expression of *otrB* in *Streptomyces*

Protoplasts of *Streptomyces rimosus* 15883s and *Streptomyces albus*, both non OTC producers were transformed with one of two plasmids, both carrying *S. rimosus* DNA fragments encoding *otrB*. One of these plasmids also contained DNA encoding part of an divergently transcribed upstream gene *otrR* , thought to be the *otrB* repressor gene. The OTC resistance profiles of transformed strains were determined by growth on solid media containing increasing levels of OTC. Both plasmids conferred resistance upon the hosts. The highest OTC concentration at which growth occurred was observed for *S. rimosus* 15883s transformed with the plasmid that did not contain DNA encoding the putative repressor gene. It is thought that the absence of *otrR* in this construct leads to de-regulation of expression. When the resistance tests were repeated using CTC or TET in place of OTC, no growth was seen. This indicates that OtrB is substrate specific in these organisms.

In order to further investigate the drug resistance conferring properties of *otrB* in Streptomycetes, expression studies using an inducible *Streptomyces* expression vector such as pIJ4123 (Takano *et al* 1995) should be performed. In this vector, expression of the cloned gene is under control of the thiostrepton-inducible *tipA* promoter. It would be useful to use this vector to express a fusion protein for purification by affinity purification. Some initial attempts to set up this system to express HisTagged OtrB were carried out in *S. albus*, results indicated that this is an unsuitable host. pIJ4123 transformants must be selected for using kanamycin, unfortunately untransformed *S. albus* appears to be resistant to fairly high levels of the drug. In order to overcome these difficulties such studies could be undertaken in a different host background provided there was no other source of OTC resistance. Alternatively a modified vector could be constructed, utilising a more suitable resistance marker.

## 6.3 Expression of otrB in *E. coli*

Expression of *otrB* in *E. coli* was performed using expression vectors in which the gene is expressed under the control of the viral T7 promoter, expression is inducible by IPTG. Two vectors of this type were used for expression studies of *otrB*, one of which also provided the means to construct a polyhistidine fusion to OtrB. Using this vector, pET21b (Novagen, UK), transformed, induced cells of *E. coli* BL21(DE3) and BL21(DE3) pLysS exhibited high levels of OTC resistance while uninduced cells conferred little or no resistance on the host. Expression studies using the alternative vector, pT7-7 (Tabor & Richardson, 1985) gave unsatisfactory control of expression by the inducer, basal levels of OTC resistance were similar to the levels conferred when cells had not been induced to express the gene, maximum levels of OTC resistance were also lower using this vector.

Interestingly, when resistance tests were performed with *otrB* using pET21b, the induced *E. coli* displayed a much less substrate specific character than *S. albus* or *S. rimosus* 15883s carrying *otrB*, growing to some level when any of the three drugs was present. Possible causes for this phenomenon have been discussed (4.5.3). Future work in this area would be to investigate whether *otrB* can confer resistance against a wider range of drugs when expressed in *E. coli* and to ascertain the mechanism by

which the substrate specificity observed in the natural host organism is lost in *E.coli.* It should be noted at this point that *otrB* shows higher sequence similarity scores with multidrug resistance pumps than to other tetracycline/$H^+$ antiporters.

## 6.4 Construction, expression in *E. coli* and attempted purification of a HisTagged OtrB fusion protein.

A C-terminal fusion of *otrB* with DNA encoding a hexa-histidine peptide was constructed using pET21b. The plasmid was transformed into *E. coli* BL21(DE3), a transformant was grown and induced to express the protein by the addition of IPTG. Levels of OTC resistance in this strain were the same as that of cells expressing non tagged *otrB* from the same vector, indicating that addition of the HisTag did not interfere with the function of the protein. In contrast to these results, a construct in which *otrB* was fused at the N-terminus to the HisTag displayed very reduced levels of OTC resistance, this was presumed to be due to mis-folding of OtrB in the membrane directed by the modified N-terminal end of the protein.

The HisTag was added to OtrB with the intention of isolating the fusion protein by the principle of affinity chromatography using nickel-agarose as a separation matrix. Purification of OtrB from membrane fractions of induced cells was attempted, and the column eluates analysed by SDS PAGE. No bands characteristic of over-expressed proteins were observed on the gel when stained. Furthermore, no signal was seen when gels were Western-blotted and probed with AntiHis monoclonal antibody. The integrity of the fusion was checked by sequencing of the expression construct. On analysis of the sequence data it was discovered that an error had been made when designing the inverse PCR primer. This resulted in the junction being out of frame with the polyhistidine-encoding segment of the vector. The protein product therefore contained a tag that was made up largely of proline residues and so did not bind to the column and was not recognised in the immunodetection assay.

## 6.5 Construction and expression in *E.coli* of a partial deletion of *otrB*

The tetracyline resistance proteins of Gram-negative organisms consist of 12 transmembrane helices in contrast to the 14 such helices observed or predicted for TET and other drug exporters of Gram-positive organisms. This observation led to the construction of a mutant otrB gene that encoded only the first six and the last six putative transmembrane helices. The aim behind the construction of this mutant gene was to investigate it's expression in Gram-negative *E.coli*. Would the deletion of two helices of OtrB prevent the protein from functioning in *E.coli*? When the construct was transformed into BL21(DE3) pLysS, OTC resistance was observed but at a very low level compared to that of the wild type. In order to make this experiment more complete, a similar experiment must be performed using a streptomycete host, this would give an indication into the importance of the two central helices in Gram-positive organisms. Other experiments that could be performed to further investigate the "preference" of bacteria for 12 or 14 helix drug efflux proteins would be to attempt to express a TET resistance pump from a Gram-negative bacterium in a Gram-positive host. For example to transform the Tn*10 tetA* gene which encodes a 12 TMH protein into *S. aureus* which is the host of TetK. As well as going some way to investigate the choice of 12 or 14 helix tetracycline pumps the experiment performed using the mutant OtrB indicates that there are no residues absolutely essential for OTC resistance present in the deleted region.

## 6.6 Sequencing and protein structure analysis

During the course of working on experiments documented in this thesis it became apparent that the previous sequences reported for *otrB* contained some errors. In order to correctly translate the gene into amino acid sequence and use this to predict protein structure, the correct nucleotide sequence of the gene was of importance. The revised nucleotide sequence was then used in TblastN searched to find closely related proteins.

The amino acid sequence was itself compared against other proteins from computer databases to investigate the assignment of OtrB to a specific branch of the Major Facilitator Superfamily. Once the correct protein sequence had been compiled, a model of the membrane topology of OtrB was drawn up using computer-based

algorithms that produced hydropathy profiles of the protein. The model was then used to decide upon the exact position of the deletion discussed above. Although computer-based models of protein structure are useful they are not to be taken as absolute. Further experimental work is needed to determine the topology of OtrB within the membrane. This can be done by fusions of OtrB to reporters such as beta-lactamase or alkaline phosphatase or by protease digestion of the extruding loops from membrane vesicles. Future work to determine the essential and important functional residues of the protein would help to elucidate the precise mechanism of resistance of the protein.

## 6.7 Future Work

It is most important that the HisTag construct be re-made using the correct reverse PCR primer as detailed in Chapter 5. Once this plasmid has been built and transformed into *E. coli*, isolation of the fusion protein by the NTA- nickel agarose column method should be acheivable.

The threat of drug resistance in clinical micro-organisms is rapidly advancing, understanding of drug resistance mechanisms, especially those of producer organisms is important when designing drugs of the future. Much of the on-going work with *Streptomyces* sp. concerns the possibility that novel polyketide antibiotics may be synthesised by what is known as the hybrid genetics approach (Katz & Donadio, 1993; McDaniel *et al*, 1993; Burson *et al*, 1997). Recent years have seen the engineering of various hybrid polyketide synthases (Kramer & Khosla, 1998; Carreras & Khosla, 1998; McDaniel *et al*, 1997). It is envisaged that as understanding of the "programming" of such synthases becomes more advanced, novel compounds could be synthesised from a variety of genes chosen from a library of polyketide and post-polyketide biosynthesis genes. Understanding of the resistance and export mechanisms of producer organisms is critical if such an approach is to be successful.

# BIBLIOGRAPHY


Adamidis, T., Champness, W. (1992). Genetic analysis *of afsB,* a *Streptomyces coelicolor* locus involved in global antibiotic regulation. J. Bacteriol., **174**. pp4622-4628.

Adamidis, T., Riggle, P., Champness, W. (1990). Mutations in a new *Streptomyces coelicolor* locus which globally block antibiotic synthesis but not sporulation. J. Bacteriol., **172**. pp2962-2969.

Aldema, M. L., McMurray, L. M., Walmsley, A. R., Levy, S. B. (1996). Purification of the Tn*10* specified tetracycline efflux antiporter TetA in a native state as a polyhistidine fusion protein. Mol. Microbiol. **19**. pp187-195.

Allard, J . D., Bertrand, K. P (1993). Sequence of a class E tetracycline resistance gene from *Escherichia coli* and comparison to related tetracycline efflux proteins. J. Bacteriol., **175**. pp4554-4560.

Allard, J. D., Gibson, M. L., Vu, L. H., Nguyen, T. T., Bertrand, K. P. (1993) Nucleotide sequence of class D tetracycline resistance genes from *Salmonella ordonez*. Mol. Gen. Genet., **237**. pp301-305.

Amakasu, H., Suzuki, Y., Nishizawa, M., Fukasawa, T. (1993). Isolation and characterisation of SGE-1; a yeast gene that partially suppresses the *gal1* mutation in multiple copies. Genetics, **134**. pp675-683.

Andersson, H. & von Heijne, G. (1993). Sec dependent and sec independent assembly of *E. coli* inner membrane proteins: the topological rules depend on chain length. EMBO. J. **2**. pp683-691.

Aukhil, I. Joshi, P., Yan, Y. Erickson, H. P. (1993). Cell heparin binding domains of hexabrachion. J. Biol. Chem. **268**. pp2542-2553.

Babcock, M. J., Kendrick, K. E. (1990).Transcriptional and translational features of a sporulation gene of *Streptomyces griseus*. Gene, **95**. pp57-63.

Bainton, N. J., Stead, P., Chhabra, S. R., Bycroft, B. W., Salmond G. P. C., Stewart, G. S. A. B., Williams, P. (1992). *N*-(3-Oxohexanoyl)-L-homoserine lactone regulates carbapenem antibiotic production in Erwinia carotovora. Biochem. J. **288**. pp997-1004.

Baldwin, S. A., Henderson, P. J. F. (1989). Homologies between sugar transporters from prokaryotes and eukaryotes. Ann. Rev. Physiol., **51.** pp459-471.

Bascaran, V., Sandez, L., Hardisson, C., Brana, A. F. (1991). Stringent response and initiation of secondary metabolism in *Streptomyces clavuligerus*. Jnl. Gen. Microbiol., **137**. pp1625-1634.

Benveniste, R. Davies, J. E. (1973) Aminoglycoside antibiotic-inactivating enzymes in actinomycetes similar to those present in clinical isolates of antibiotic-resistant bacteria. Proc. Nat. Acad. Sci. USA., **70**. pp2276-2280.

Bibb, M. J., Biro, S., Motamedi, H., Collins, J. F., Hutchinson, C. R. (1989). Analysis of the nucleotide sequence of the *Streptomyces glaucesens tcmI* genes provides key information about the enzymology of polyketide antibiotic biosynthesis. EMBO. Jnl., **8**. pp2727-2736.

Bibb, M. J., Findlay, P. R., Johnson, M. W. (1984). The relationship bteween base composition and codon usage in bacterial genes and it's use for the simple and reliable identification of protein coding sequences. Gene, **30**. pp157-166.

Bibi, E. & Kaback, H. R. (1990) *In vivo* expression of the *lacY* gene in two segments leads to functional lactose permease. Proc. Nat. Acad. Sci. USA., **87**. pp4325-4329.

Binnie, C., Warren, M., Butler, M. J. (1989). Cloning and heterologous expresion in

*Streptomyces lividans* of *Streptomyces rimosus* genes involved in oxytetracycline biosynthesis. J. Bacteriol., **171**. pp887-895.

Birnboim, H. C. & Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucl. Acids. Res., **7**. pp1513-1523.

Blanc, V., Saleh-Bey, K., Folcher, M., Thompson, C. J. (1995). Molecular characterisation and transcription analysis of a multidrug resistance gene cloned from the pristamycin-producing organism, *Streptomyces pristinaespiralis*. Mol. Microbiol., **17**. pp989-999.

Bu'Lock, J. D. (1980). Mycotoxins as secondary metabolites. In *The biosynthesis of mycotoxins: A study in secondary metabolism*. Steyn, P.S. (Ed.) Academic Press. NY. pp1-16.

Burdett, V. (1986). Streptococcal tetracycline resistance mediated at the level of protein synthesis. J. Bacteriol., **165**. pp564-569.

Burdett, V. (1991). Purification and characterisation of TetM, a protein that renders ribosomes resistant to tetracycline. J. Biol. Chem., **266**. pp2872-2877.

Burson, K. K., Huestis, W. H., Khosla, C. (1997). Gene shuffling of bacterial aromatic polyketide synthases. Abstracts of papers of the American Chemical Society. **213**. Pt1. p83-BIOT.

Butler, M. J., Friend, E. J., Hunter, I. S., Kaczmarek, F. S., Sugden, D. A., Warren, M. (1989). Molecular cloning of resistance genes and architecture of a linked cluster involved in biosynthesis of oxytetracycline by *Streptomyces rimosus*. Mol. Gen. Genet., **215**. pp231-238.

Buttner, M. J., Chater, K. F., Bibb, M. J. (1990). Cloning disruption and transcriptional analysis of three RNA polymerase genes of *Streptomyces coelicolor*

A(3)2. Jnl Bacteriol, **172.** pp3367-3378.

Buttner, M. J. Lewis, C. G. (1992) Construction and characterisation of *Streptomyces coeliocolor* A(3)2 mutants that are multiply deficient in the non essential *hrd* encoded RNA-polymerase sigma factors. J. Bacteriol., **174.** pp5165-5167.

Cabellero, J. L., Malpartida, F., Hopwood, D. A. (1991)a. Transcriptional organisation and regulation of an antibiotic export complex in the producing *Streptomyces* culture. Mol. Gen. Genet., **278**. pp372-380.

Cabellero, J .L., Martinez, E., Malpartida, F., Hopwood, D. A. (1991b). Organisation and functions of the *act*VA region of the actinorhodin biosynthetic gene cluster of *Streptomyces coelicolor*. Mol. Gen. Genet., **230**. pp401-412.

Calcutt, M. J. Cundliffe, E. (1989). Use of a fractionated coupled transcription translation system in the study of ribosomal resistance mechanisms in the antibiotic producing *Streptomyces*. Jnl. Gen. Microbiol., **135**. pp1071-1081.

Campbell, W. H. (1992). Expression of *E. coli* cytochrome C reductase activity from a maize NADH: Nitrate reductase complementary DNA. Plant Physiol. **99**. pp693-699.

Carreras, C. W., Khosla, C. (1998). Purification and *in vitro* reconstitution of the essential protein components of an aromatic polyketide synthase. Biochem. **37**. pp2084-2088

Chambers, S.P., Prior, S. E., Barstow, D. A., Minton, N. P. (1988). The pMTL *nic* cloning vectors. 1. Improved pUC polylinker regions to facilitate the use of sonicated DNA for nucleotide sequencing. Gene, **68**. pp139-149.

Chang, A. C. Y. & Cohen, S. N. (1978). Construction and characterisation of amplifiable multicopy DNA cloning vehicles derived from the p15A cryptic

miniplasmid. J. Bacteriol. **134**. pp1141-1156.

Chater, K. F., Bruton, C. J., Plaskitt, K. A. Buttner, M. J., Mendez, C. Helmann, J. D. (1989) The developmental fate of *Streptomyces coelicolor* hyphae depends on a gene product homologous with the motility sigma factor of *Bacillus subtilis*. Cell, **59**. pp133-143.

Chater, K. F. & Merrick, M. J. (1979). In "*Developmental Biology of prokaryotes*". Parish, J. H. (ed). Blackwell Scientific Publications, Oxford. pp93-114.

Charpentier, E., Gerbaud, G., Courvali, P. (1993) Characterisation of a new class of tetracycline resistance gene, *tet(S)* in Listeria monocytogenes BM4210. Gene, **131**. pp27-34.

Chopra, I. (1986) Genetic and biochemical basis of tetracycline resistance. Jnl. Antimicro. Chemother., **18**. No. SC, pp51-56.

Chopra, I. (1994) Tetracycline analogues whose primary target is not the bacterial ribsome. Antimicrob. Agents. Chemother., **38**. pp637-640.

Cline, J., Braman, J. C., Hogrefe, H. H. (1996). PCR fidelity of *Pfu* DNA polymerase and other thermostable DNA polymerases. Nucleic Acids Research, **24**. pp3546-3551.

Coque, J. J., Liras, P., Martin, J. F. (1993). Genes for a β-lactamase, a penicillin binding protein and a transmembrane are clustered with the cephamycin biosynthetic genes in *Nocardia lactamdurans*. EMBO J., **12**. pp631-639.

Cortes, J., Haydock, S. F., Roberts, G. A. Bevitt, D. J. Leadlay, P. F. (1990). An unusually large multifunctional polypeptide in the erythromycin producing polyketide synthetase of *Saccharospora erythraea*. Nature, **345**. pp176-178.

Crameri, R., Davies, J. E. (1986). Increased production of aminoglycosides associated with amplified antibiotic resistance genes. Jnl. Antibiot., **39**. pp128-135.

Cuirale, M. S., McMurray, L. M., Levy, S.B. (1984). Intracistronic complementation of the tetracycline resistance membrane protein of Tn*10*. J. Bacteriol., **157**. pp211-217.

Cundliffe, E. (1987) On the nature of antibiotic binding sites in ribosomes. Biochimie, **69**. pp863-869.

Cundliffe, E. (1989). How antibiotic producing organisms avoid suicide. Ann. Rev. Microbiol., **43**. pp207-233.

Davies, J. E. (1990). What are antibiotics? Archaic functions for modern activities. Mol. Microbiol, **4**, pp1227-1232.

Davies, J. E., Vonahsen, U., Wank, H., Schroeder, R. (1992).Evolution of secondary metabolite production - potential roles for antibiotics as prebiotic effectors of catalytic RNA reactions. In "*Secondary metabolites, their function and evolution.*" Ciba Foundation Symposium. **171**. pp24-44.

Decker, H., Hutchinson, C. R. (1993). Transcriptional analysis of the *Streptomyces glaucesens* tetracenomycin-C biosynthesis gene cluster. Jnl. Bacteriol., **175.** pp3887-3892.

Demain, A. L. (1980). Do antibiotics form in nature? Search, **11**. pp148-151.

Dente, L., Cesareni, G., Cortese, R. (1983). pEMBL - A new family of single stranded plasmids. Nucleic Acid Res., **11**. pp1645-1655.

Devereux, J., Haeberli, P., Smithies, O. (1984). A comprehensive set of sequence analysis programs for the vax. Nucleic Acids Res. **12**. pp387-395.

Dietrich, J. B., Lorber, B., Kern, D. (1991). Expression of mammalian tyrosine aminotransferase in *Saccharomyces cerevisiae* and *Escherichia coli* : purification to homogeneity and characterisation of the enzyme overproduced in the bacteria. Eur. J. Biochem. **201.** pp399-407.

Dittrich, W. & Schrempf, H. (1992). The unstable tetracycline resistance gene of *Streptomyces lividans* 1326 encodes a putative protein with similarities to translational elongation factors and tet(M) and Tet(O) proteins. Antimicro. Agents. Chemother., **36**. pp1119-1124.

Distler, J., Ebert, A., Mansouri, K., Pissowotzki, K., Stockmann, M. (1987). Gene cluster for streptomycin biosynthesis in *Streptomyces griseus*: nucleotide sequence of three genes and analysis of transcriptional activity. Nucleic acids. Res., **15**. pp8041-8056.

Distler, J., Mansouri, K., Mayer, G., Stockmann, M. Piepersberg, W. (1992). Streptomycin production and it's regulation. Gene, **115.** pp105-111.

Doyle, D. (1987). Ph.D. Thesis, University of Glasgow.

Doyle, D., Butler, M.J. Hunter, I. S. (1988). Molecular analysis of the oxytetracycline resistance gene *otrA* of *Streptomyces rimosus*. Heredity, **61**. pp305

Doyle, D., McDowall, K. J., Butler, M. J., Hunter, I..S. (1991.) Characterisation of an oxytetracycline resistance gene *otrA* of *Streptomyces rimosus*. Mol. Microbiol., **5.** pp2923-2933.

Durckheimer, W. (1975). Tetracyclines: chemistry, biochemistry, and structure-activity relations. Angew. Chem. Int. Ed. Engl., **14**. pp721-734.

Eckert, B., Beck, C. F. (1989). Topology of the transposon Tn*10* encoded tetracycline

resistance protein within the inner membrane of *Escherichia coli*. J. Biol. Chem., **264**. pp11663-11670.

Engelman, D. M. & Steitz, T. A. (1981). The sponaneous insertion of proteins into and across membranes - The helical hairpin hypothesis. Federation Proc. **40**. pp1557.

Engelman, D. M., Steitz, T. A., Goldman, A. (1986). Identifying transbilayer helices in amino acid sequences of membrane proteins. Ann. Rev. Biophys. Biophys. Chem., **15**. pp321-353.

Fernandez-Moreno, M. A., Cabellero, J. L. Hopwood, D. A., Malpartida, F. (1991). The *act* cluster contains regulatory and antibiotic export genes, direct targets for translational control by the *bldA* transfer RNA gene of *Streptomyces*. Cell, **66**. pp769-780.

Fernandez-Moreno, M. A., Martinez, E., Boto, L., Hopwood, D. A. Malpartida, F. (1992a). Nucleotide sequence and deduced functions of a set of cotranscribed genes of *Streptomyces coelicolor* A3(2) including the polyketide synthase for the antibiotic actinorhodin. J. Biol. Chem., **267.** pp19278-19290.

Fernandez-Moreno, M. A., Martinez, E. Triana, A. J., Martinez, E., Niemi. J., Kieser, H. M., Hopwood, D. A., Malpartida, F. (1992b). *abaA*, a new pleiotropic locus for antibiotic production in *Streptomyces ceolicolor*. J. Bacteriol, **174**. pp2958-2967.

Flaschel, E. & Friehs, K. (1993). Improvement of downstream processing of recombinant proteins by means of genetic engineering methods. Biotech. Adv. **11**. pp31-78.

Friend, E. J., Hopwood, D. A. (1971). The linkage map of *Streptomyces rimosus*. Jnl. Gen. Microbiol., **68**. pp187-197.

Fujii, T., Gramajo, H. C., Takano, E., Bibb, M.J. (1996). *redD* and *actII-Orf4*,

pathway specific regulatory genes for antibiotic production in *Streptomyces coelicolor* A3(2), are transcribed *in vitro* by an RNA-polymerase holoenzyme containing sigma (HrdD). Jnl Bacteriol. **178**. pp3402-3405.

Fujihara, E., Kimura, T., Shiina, Y., Yamaguchi, A. (1996). Transmembrane glutamic acid residues play essential roles in the metal-tetracycline/$H^+$ antiporter of *Staphylococcus aureus*. FEBS Lett., **391**. pp243-246.

Gallant, J.A. (1979). Stringent control in *Escherischia coli.* Ann. Rev. Genet., **13.** pp393-415.

Garven, S. (1994) Ph.D. Thesis, University of Glasgow.

Ginn. S. L., Brown, M. H., Skurray, R. A. (1997). Membrane topology of the metal tetracycline/$H^+$ antiporter TetA(K) from *Staphylococus aureus*. J. Bacteriol., **179.** pp3786-3789.

Gorman, M., Chamberlin, J.W., Hammil, R. L. (1968). Monesin, a new biologically active compound. V. Compounds related to monesisn. Antimicrob. Agents. Chemother. **11** pp363-368.

Guo, D., Liu, J., Motlagh, A., Jewell, J., Miller, K. W. (1996). Efficient insertion of odd-numbered transmembrane segments of the tetracycline resistance protein requires even-numbered insertion. J. Biol. Chem., **271**. pp30829-30834.

Griffith, J. K., Baker, M. E., Rouch, D. A., Page, M. G. P., Skurray, R. A., Paulsen, I. T., Chater, K. F., Balwin, S. A., Henderson, P. J .F. (1992). Membrane transport proteins, implications of sequence comparisons. Curr. Opin. Cell. Biol., 4. pp684-695.

Grimm, A., Madduri, K., Ali, A., Hutchinson, C. R. (1994). Characterisation of the *Streptomyces peuticus* ATCC-29050 gene encoding doxorubicin polyketide synthase.

Gene, **151**, pp1-10.

Guay, G. G., Khan, S. A., Rothstein, D. M. (1993a). The *tetK* gene of plasmid pT181 from *Staphylococcos aureus* encodes an efflux protein that contains 14 transmembrane helices. Plasmid, **30**. pp163-166.

Guay, G. G., Tuckman, M., McNicholas, P., Rothstein, D. M. (1993b). The *tetK* gene from *Staphylococcus aureus* mdiates transport of potassium in *Escherichia coli*. Jnl. Bacteriol. 175. pp4927-4929.

Guay, G. G., Tuckman, M., Rothstein, D. M. (1994). Mutations in the *tetA(B)* gene that cause a change in the substrate specificity of the tetracycline efflux pump. Antimicrob. Agents. Chemother., **38**. pp857-860.

Guffanti, A. A., Krulwich, T. A. (1995). Tetracycline/$H^+$ antiport and $Na^+/H^+$ antiport catalysed by the *Bacillus subtilis* TetA(L) transporter expressed in *Escherichia coli*. J. Bacteriol., **177**. pp4557-4561.

Guilfoile, P. G. Hutchinson, C. R. (1992). Sequence and transcriptional analysis of the *Streptomyces glaucesens tcmAR* tetracenomycin resistance and repressor gene loci. J. Bacteriol., **174.** pp3651-3658.

Guthrie, E. P., Chater, K. F. (1990). The level of a transcript required for the production of a *Streptomyces coelicolor* antibiotic is conditionally dependent on a tRNA gene. J. Bacteriol., **172**. pp6189-6193.

Guthrie, E. P., Flaxman, C. S., White, J., Hodgson, D. A., Bibb, M. J., Chater, K. F. (1998). A response-regulator-like activator of antibiotic synthesis from *Streptomyces coelicolor* A3(2) with an amino-terminal domain that lacks a phosphorylation pocket. Microbiology, **144**. pp727-738.

Hahlbrock, K. (1981). Flavenoids. In, "*The Biochemistry of Plants*". Conn, E.E.

(ed.). Academic Press, Inc., New York. pp425-456.

Hanahan, D. (1983). Studies on the transformation of *Escherichia coli* with plasmids. Jnl. Genl. Microbiol., **166**. pp557.

Hansen, L. M., McMurray, L. M., Levy, S. B., Hirsh, D. C. (1993). A new tetracycline determinant, *tetH* from *Pastuerella maltocida* specifying active efflux of tetracycline. Antimicrob. Agents. Chemother., **37**. pp2699-2705.

Hara, O. & Beppu, T. (1982) Induction of Streptomycin inactivating enzyme by A-factor in *Streptomyces griseus*. Jnl. Antibiot, **35**. pp1208-1215.

Hara, O., Horinouchi, S., Uozumi, T., Beppu, T. (1983). Genetic analysis of A-factor synthesis in *Streptomyces coelicolor* A3(2) and *Streptomyces griseus*. Jnl. Gen Microbiol., **129**. pp2939-2944.

Helmann, J. D. Marquez, L. M., Chamberlin, M. J. (1988). Cloning, sequencing and disruption of the *Bacillus subtilis* $\sigma^{28}$ gene. J. Bacteriol. **170**. pp1568-1574.

Henderson, P. J. F. (1990). Proton linked sugar transport systems in bacteria. Jnl. Bioenerget. Biomemb., **22**. pp525-569.

Henderson, P. J. F., MacPherson, A. J. S. (1986). Assay, genetics, proteins and reconstitution of proton linked galactose, arabinose and xyose transport systems of *Escherichia coli*. Methods. Enzymol., **125**. pp387-429.

Henderson, P. J. F., Maiden, M. C. J. (1990). Homologous sugar transport proteins in *Escherichia coli* and their relatives in both prokaryotes and eukaryotes. Phil. Transac. Royal. Soc. Lond. (B)., **326**. pp391-410.

Heukshoven, J. & Dernick, R. (1988). Improved silver staining procedure for fasr staining in phastsystem development unit .1. Staining of Sodium dodecyl sulphate

gels. Electrophoresis. **9**. pp28-32.

Hochstein, F. A., Schach von Wittenau, M., Tanner, F. W. Jr., Muraki, K. (1960). 2 Acetyl-2-decarboxamidooxytetracycline. J.Am. Chem. Soc., **82.** pp5934.

Hochuli, E. (1988). Large-scale chromatography of recombinant proteins. J. Chromatogr., **444**. pp293-302.

Hollingshead, S., Vapnek, D. (1985). Nucleotide sequence analysis of a gene encoding a streptomycin, spectinomycin adenyltransferase. Plasmid, **13**. pp17-30.

Hong, S-K., Kito, M., Beppu, T., Horinouchi, S. (1991.) Phosphorylation of the *afsR* product, a global regulatory protein for secondary metabolite production in *Streptomyces coelicolor* A3(2). J. Bacteriol. **173**. pp2311-2318.

Hopp, T. P. Prickett, K. S., Price, V. L. Libby, R. T. March, C. J., Ceretti, D. P., Urdal, D. L., Conlon, P. J. (1989). A short polypeptide marker sequence useful for recombinant protein identification and purification. Bio/Technology. **6**. pp1204-1210.

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., Schrempf, H. (1985). Genetic manipulation of *Streptomyces*; a laboratory manual. The John Innes Foundation, Norwich, UK.

Hopwood, D. A. & Khosla, C. (1992) Genes for polyketide secondary metabolic pathways in microorganisms and plants. In " *Secondary metabolites, their function and evolution* " Ciba Foundation Symposium 171. pp88-112.

Hopwood, D. A., Sherman, D.H. (1990). Molecular genetics of polyketide synthesis and it's comparison to fatty acid biosynthesis. Ann. Rev. Genet., **24**. pp37-66

Hopwood, D. A., Wildermuth, H. Palmer, H. M. (1970.) Mutants of *Streptomyces*

*ceolicolor* defective in sporulation. Jnl. Gen. Microbiol., **61**. pp397-408.

Hoshino, T., Ikeda, T., Tomizuka, N., Furukawa, K. (1985.) Nucleotide sequence of the tetracycline resistance gene of pTHT15, a thermophillic Bacillus plasmid; comparison with staphylococcal tetracycline resistant controls. Gene, **37**. pp131-138.

Horinouchi, S. & Beppu, T. (1983). Cloning of a pleiotropic gene that positively contols biosynthesis of A-factor, actinorhodin and prodigiosin in *Streptomyces coelicolor* A3(2) and *Streptomyces lividans*. J. Bacteriol. **155**. pp1238-1248.

Horinouchi, S. & Beppu, T. (1994). A-factor as a microbial hormone that controls cellular differentiation and secondary metabolism in *Streptomyces griseus*. Mol. Microbiol. **12**. pp859-864.

Horinouchi, S., Kito, Nishiyama, N., Furuya, K., Hong, S-K., Miyake, K., Beppu, T. (1990). Primary structure of AfsR, a global regulatory protein for secondary metabolite formation in *Streptomyces coelicolor* A3(2). Gene, **95**. pp49-56.

Horinouchi, S., Kumada, Y., Beppu, T. (1984). Unstable genetic determinant of A-factor biosynthesis in streptomycin producing organisms: cloning and characterisation. J. Bacteriol. **158**. pp481-487.

Hoshino, T. T., Ikeda, N., Tomizuka, N., Furukawa. (1985). Nucleotide sequence of the tetracycline resistance gene of pTHT15, a thermophiliic *Bacillus* plasmid; comparison with staphylococcal $Tc^R$ controls. Gene, **37**. pp131-138.

Hunter I.S. & Hill, R. A. (1997). Tetracyclines; Chemistry and molecular genetics of their formation. In "*Biotechnology of Industrial antibiotics*." Strohl, W.D. (ed.) Marcel Decker. pp211-231.

Hunter, I. S. (1985). Gene cloning in *Streptomyces*. In *DNA cloning II; a practical approach*. Glover, D. M. (ed.) IRL Press. pp19-44.

Inouye, M., Arnheim, N., Sternglanz, R. (1973). Bacteriophage T7 lysozyme is an acetylmuranyl L-alanine amidase. J. Biol. Chem. **248**. pp7247-7252.

Ishizuka, H., Horinouchi, S., Kieser, H. M., Hopwood, D. A., Beppu, T. (1992). A putative two component regulatory system involved in secondary metabolism in *Streptomyces* species. Jnl Bacteriol. **174**. pp7585-7594.

Ives, C. L., Bott, K. F. (1990). Cloned *Bacillus subtilis* chromosomal DNA mediates tetracycline resistance when present in multiple copies. J. Bacteriol., **171**. pp1801-1810.

Jessen-Marshall, A. E., Paul, N. J., Brooker, R. J. (1995). The conserved motif, *GXXX(DE)(R/K)XG[X](R/K)(R/K)* , in hydrophillic loop2/3 of the lactose permease. J. Biol. Chem., **270**. pp16251-16257.

Ji, G. Y., Beavis, R. C., Novick, R. P. (1995). Cell density control of staphylococcal virulence mediated by an octapeptide pheremone. Proc. Natl. Acad. Sci. USA. **92**. pp12055-12059

Kanazawa, S. Driscoll, M., Struhl, K. (1988). ATR-1: *Saccharomyces cerevisiae* gene encoding a transmembrane required for aminotiazole resistance. Mol. Cell. Biol., **8**. pp664-673.

Kaneko, M., Yamaguchi, A., Sawai, T. (1985). Energetics of tetracycline efflux encoded by Tn*10* in *Escherichia coli*. FEBS. Lett., **193**. pp194-198.

Katz, E. & Demain, A.L. (1977). The peptide antibiotics of *Bacillus*; chemistry, biogenesis and possible functions. Bacteriol. Rev., **41**. pp449-474.

Katz, L. & Donadio, S. (1993). Polyketide synthesis- Prospects for hybrid antibiotics. Ann. Rev. Microbiol. **47**. pp875-912.

Katz, E., Thopson, C. J., Hopwood. D. A. (1983). Cloning and expression of the tyrosinase gene from *Streptomyces antibioticus* in *Streptomyces lividans*. J. Gen. Microbiol. **129**. pp2703-2714.

Keleman,G. H., Brown, G. L., Kormanec, J., Potuckova, L., Chater, K. F. Buttner, M. J. (1996). The positions of the sigma factor genes *whiG* and *SigF* in the hierarchy controlling the development of spore chains in the aerial hyphae of *Streptomyces coelicolor* A3(2). Mol. Microbiol., **21**. pp595-603.

Keleman, G. H., Plaskitt, K. A., Lewis, C. G., Findlay, K., Buttner, M. J. (1995). Deletion of DNA lying close to the *glkA* locus induces ectopic sporulation in *Streptomyces coelicolor* A3(2). Mol. Microbiol., **17**. pp221-230.

Kendrick, E., Ensign, J. C. (1983). Sporulation of *Streptomyces griseus* in submerged cluture. J. Bacteriol., **155**. pp357-366

Khokhlov, A. S. (1982). Low molecular weight microbial bioregulators of secondary metabolism. In *Over production of Microbial product,* Krumphanzi, V., Sikyta, B & vanek, Z. (eds) London Academic, UK. pp97-109.

Khokhlov, A. S., Tovarova, I. I., Borisova, L. N., Pliner, S. A., Shevshenko, L. A. (1967). A factor responsible for the biosynthesis of streptomycin by a mutant strain of *Actinomyces streptomycinii*. Dokl. Acad. Nauk. SSSR., **177**. pp232-235.

Khokhlov, A. S., Tovarova, I. I., Kleiner, E. Y., Kovalenko, I. V. (1973). Effect of A-factor on the growth of asporogeneous mutants of *Streptomyces griseus* not producing this factor. Z. Allg. Mikrobiol, **13**. pp647-655.

Khosla, C., Daniel, R., Ebert-Khosla, S., Torres, R., Sherman, D. H. Bibb, M. J. Hopwood, D. A. (1993). Genetic construction and functional analysis of hybrid polyketide synthases containing heterologous acyl carrier proteins. J. Bacteriol., **175**.

pp2197-2204.

Khosla, C., Ebert-Khosla, S., Hopwood, D. A. (1992). Targeted gene replacements in a *Streptomyces* polyketide synthase gene cluster: role for the acyl carrier protein. Mol. Microbiol., **6**. pp3237-3249.

Kitamura, I., Tobe, H., Yashimoto, A., Oki, T., Nagahawa, H., Takeuchi, T., Umezawa, H. (1981). Biosynthesis of aklavinone and aclacinomycins. J. Antibiot., **34**. pp959-964.

Kormanec, J., Farkovski, M. (1983). Differential expression of principal sigma factor homologues of *Streptomyces aurefaciens* correlates with the developmental stage. Nucleic Acids. Res., **21**. pp3647-3652.

Kormanec, J., Potuckova, R., Rezuchkova, B. (1994). The *Streptomyces aureofaciens* homologue of the *whiG* gene encoding a putative sigma factor essential for sporulation. Gene, **143**. pp101-103.

Kramer, P. J.& Khosla, C. (1998). Enginering of novel polyketides - Progress and prospects. Annal. N.Y. Acad. Sci., **799**. pp32-34.

Kuhn, A. (1988). Alterations in the extracellular domain of M13 procoat protein make its membrane insertion dependent on *secA* and *secY*. Eur J Biochem. **177**. pp267-271.

Kyte, J. & Doolittle, R. F. (1982). A simple method for displaying the hydrophathic character of a protein. J. Mol. Biol.,**157**. pp105-132.

Lawlor, E. J., Baylis, H. A., Chater, K. F. (1987.) Pleiotropic morphological and antibiotic deficiencies result from mutations in a gene encoding a tRNA-like product of *Streptomyces coelicolor* A3(2). Genes. Dev., **2**. pp1305-1310.

Le Blanc, D. J., Lee., L. N., Titmas, B. M., Smith, C. J., Tenover, F. C. (1988). Nucleotide sequence analysis of tetracycline resistance gene *tetO* from *Streptomyces mutans* DL5. J. Bacteriol., **170**. pp3618-3626.

Leboul, J., Davies, J. (1982). Enzymatic modification of hygromycin B in *Streptomyces hygroscopicus*. Jnl. Antibiot., **35**. pp527-528.

Leskiw, B. K., Bibb, M. J. Chater, K. F. (1991). The use of a rare codon specifically during development? Mol Microbiol., **5**. p2861-2867.

Levy, S. B., McMurray, L. M., Burdett, V., Courvalin, P. Hillen, W., Roberts, M. C., Taylor, D. E. (1989). Nonclamenture for tetracycline resistance determinants. Antimicro. Agents. Chemother. **33**. pp1373-1374.

Lewis, K. (1994). Multidrug resistance pumps in bacteria: variations on a theme. Trends. Biochem. Sci., **19**. pp119-123.

Lineal, M. (1993) Vesicular transporters join the major facilitator superfamily (MFS). Trends. Biochem. Sci., **18**. pp248-249.

MacNeil, D. J. (1988). Characterisation of a unique methyl-specific restriction system in *Streptomyces avermitilis*. Jnl. Bacteriol. 170. pp5607-5612.

Maiden, M. C. J. (1987) Ph.D. Thesis, University of Cambridge.

Maiden, M. C. J., Jones-Mortimer, M. C., Henderson, P. J. F. (1988). The cloning, DNA sequence and over expression of the gene AraE coding for arabinose-proton symport in *Escherichia coli* K12. J. Biol. Chem., **263**. pp8003-8010.

Malpartida, F. M., Hopwood,D. A. (1984). Molecular cloning of the whole biosynthesis pathway of a *Streptomyces* antibiotic and it's expression in a heterologous host. Nature, **309**. pp462-464.

Malpartida, F. M., Hopwood,D. A. (1986). Physical and genetic characterisation of the gene cluster for the antibiotic actinorhodin in *Streptomyces coelicolor* A3(2). Mol. Gen. Genet., **205**. pp66-73.

Manniatis, T., Fritsch, E. F., Sambrook, J. (1982). Molecular cloning: A laboratory manual. Cold Spring Harbour, NY.

Marger, M. D., Saier, M. F. (1993). A major superfamily of transmembrane facilitators that catalyse uniport, symport and antiport. Trends. Biochem. Sci., **18**. pp13-20.

Martinez-Costa, O. H., Arias, P., Romero, N. M., Parro, V., Mellado, R. P., Malpartida, F. (1996). A relA/spoT homologous gene from *Streptomyces coelicolor* A3(2) controls antibiotic synthesis genes. J. Biol. Chem., **271**. pp10627-10634.

Martin, P., Tieu-Cuot, P., Courvalin, P. (1986). Nucleotide sequence of the *tetM* resistance determinant of the streptococcal conjugative shuttle transposon Tn*1545*. Nuc. Acids. Res., **14**. pp7047-7058.

Matsumoto, A., Hong, S-K., Ishizuka, H., Horinouchi, S, Beppu, T. (1994). Phosphorylation of the AfsR protein involved in secondary metabolism in *Streptomyces* species is by a eukaryotic-type kinase. Gene, **146**. pp47-56.

McCormick, J. R. D. , Sjolander, N. O., Hirsch, U., Jenssen, E. R., Doeschuk, A. P. (1957) A new family of antibiotics, the demethyltetracyclines. J. Am. Chem. Soc., **79**. pp4561.

McCormick, J. R. D. (1966). Biosynthesis of the tetracyclines, an integrated biosynthetic scheme. In *Antibiotics, advances in research, production and clinical use*. Herold, M., Gabriel, Z. (ed.'s) Czechoslovak medical press. pp556-573.

McCormick, J. R. D. (1968).Biosynthesis of tetracyclines. X. Protetrone. J. Am. Chem. Soc., **90**. pp7126-7127.

Chem. Soc., **90**. pp7126-7127.

McDaniel, R., Ebert-Khosla, S. Hopwood, D. A., Khosla, C. (1993). Engineered biosynthesis of novel polyketides. Science, **262**. pp1546-1550.

McDaniel, R., Kao, C. M., Hwang, S. J., Khosla, C. (1997). Engineered intermodular and intramodular polyketide synthase fusions. Chemistry & Biology. **4**. pp667-674.

McDowall, K. J., Doyle, D., Butler, M. J., Binnie, C., Warren, M., Hunter, I. S. (1991). Molecular genetics of oxytetracycline production by *Streptomyces rimosus.* In *Genetics and Product formation in Streptomyces.* Baumberg, S. (ed) Plenum Press, NY. pp105-116.

MacGregor-Pryde, S. E. (1995). Ph.D. Thesis. University of Glasgow.

McMurray, L., Levy, S. B. (1995). Decreased function of the class B tetracycline efflux protein TetA, with mutations at asparagine 15, a putative intramembrane residue. J. Bacteriol. **174**. pp6294-6297.

McMurray, L., Petrucci, R. E., Levy, S. B. (1980). Active efflux of tetracycline encoded by four genetically different tetracycline resistance determinants in *Escherichia coli*. Proc. Nat. Acad. Sci. USA., **77.** pp3974-3977.

McNicholas, P., Chopra, I., Rothstein, D. .M. (1992). Genetic analysis of the *tetA(C)* gene on plasmid pBR322. J. Bacteriol., **174**. pp7926-7933.

McNicholas, P., McGlynn, M., Guay, G. G., Rothstein, D. M. (1995). Genetic analysis suggests functional interactions between the N and C terminal domains of the TetA(C) efflux pump encoded by pBR322. J. Bacteriol., **177**. pp5355-5357

Mendez, C., Chater, K. F. (1987). Cloning of *whiG*, a gene critical for sporulation of *Streptomyces coelicolor* A3(2). Jnl. Bacteriol, **169**. pp5715-5720.

Merrick, M. J.(1976). A morphological and genetic mapping study of bald colony mutants of *Streptomyces coelicolor*. Jnl. Gen. Microbiol., **96**. pp299-315.

Metzger, P.A., Schreiber, G., Aizenman, E., Cashel, M., Glaser, G. (1989). Characterisation of the *relA*1 mutation and a comparison of *relA* with new *relA* null alleles in *Escherichia coli*. Jnl Biol. Chem. 264. pp21146-21152.

Miller, P. A., Sjolander, N. O., Nalesnyk, S., Arnold, N., Johnson, S., Doerschuk, A. D., McCormick, J. R. D. (1960). Cosynthetic factor 1, a factor in hydrogen transfer in *Streptomyces aureofaciens*. J. Am. Chem. Soc., **82**. pp5002-5003.

Milton, D. L., Harman, A., Camara, M., Chhabra, S. R., Bycroft, B. W., Stewart, G. S. A. B., Williams, P. (1997). Quorum sensing in *Vibrio anguillarum*; Characterisation of the *vanI/vanR* locus and identification of the autoinducer *N*-(3-Oxodecanoyl)-L-Homoserine lactone. J. Bacteriol. **179**. pp3004-3012.

Miyake, K., Kuzuyama, T., Horinouchi, S., Beppu, T. (1990). The A-factor binding protein of *Streptomyces griseus* negatively controls streptomycin production and sporulation. J. Bacteriol. **172**. pp3003-3008.

Moffatt, B. A. & Studier, F. W. (1987) . T7 lysoszyme inhibits transcription by T7 polymerase. Cell **49**. pp221-227

Muiry, J. A. R. (1989). Ph.D. Thesis. University of Cambridge.

Nakano, H., Matsuhashi, Y., Takeuchi, T., Umezawa, H. (1977). Distribution of chloramphenicol acetyltransferase and chloramphenicol-3-acetate esterase among *Streptomyces* and Corynebacterium. Jnl. Antibiot., **30**. pp76-82.

Neal, R. J., Chater, K. F. (1987). Nucleotide sequence analysis reveals similarities between proteins determining methylenomycin-A resistance in *Streptomyces* and

tetracycline resistance in eubacteria. Gene, **58**. pp229-241.

Neal, R. J. & Chater, K. F. (1991). Bidirectional promoter and terminator regions bracket *mmr*, a resistance gene embedded in the *Streptomyces coelicolor* A3(2) gene cluster encoding methylenomycin production. Gene **100**. pp75-83.

Needleman, S. B. & Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequences of two proteins. J. Mol. Biol. **48**. pp443-453.

Nguyen, T. T., Postle, K., Bertrand, K. P. (1983). Sequence homology between the tetracycline resistance determinants of Tn*10* and pBR322. Gene, **25**. pp83-92.

Nikolich, M. P., Shoemaker, N. B., Salyers, A. A. (1992). A *Bacteroides* tetracycline resistance gene represents a new class of ribosome protection tetracycline resistance. Antimicro. Agents. Chemother., **36**. pp1005-1020.

Noguchi, N., Aoki, M., Sasatsu, M., Kono, M., Shishido, K., Ando, T. (1986). Determination of the complete nucleotide sequence of pNS1, a staphylococcal tetracycline resistance plasmid propagated in *Bacillus subtilis*. FEMS Lett., **37**. pp283-288.

Ochi, K. (1985). Sporulation and antibiotic production by *Bacillus subtilis* mutants deficient in intracellular proteases. Agric. Biol. Chem., **49**. pp905-907.

Ochi, K. (1986a). A decrease in GTP content is associated with aerial mycelium formation in *Streptomyces* MA406-A-1. Jnl. Gen Microbiol., **132**. pp299-305.

Ochi, K. (1986b). Occurrence of the stringent response in *Streptomyces* spp. and it's significance for the initiation of morphological and physiological differentiation. Jnl. Gen. Microbiol., **132**. pp2621-2631.

Ochi, K, (1987a). Changes in nucleotide pools during sporulation of *Streptomyces griseus* in submerged culture. Jnl. Gen. Micro., **133.** pp2787-2795.

Ochi, K. (1987b). A *rel* mutation abolishes an enzyme induction needed for actinorhodin synthesis by *Streptomyces* antibioticus. Agric. Biol. Chem., **51.** pp829-835.

Ochi, K. (1990). *Streptomyces relC* mutants with an altered ribosomal protein St-L11 and genentic analysis of a *Streptomyces griseus relC* mutant. J. Bacteriol., **172.** pp4008-4016.

Ochi, K., Oshawa, S. (1984). Initiation of antibiotic production of the stringent response of *Bacillus subtilis* Marburg. J.Gen. Microbiol., **130**. pp2473-2482.

Ohnuki, T., Katoh, T., Imanka, T., Aiba, S. (1985). Molecular cloning of the tetracycline resistance gene from *Streptomyces rimosus* in *Streptomyces griseus* and characterisation of the cloned genes. J. Bacteriol., **161**. pp1010-1016.

Oliva, B., Chopra, I. (1992). Tetracycline determinants provide poor protection against some tetracyclines; furhter evidence for division of tetracyclines into two classes. Anti. Micro. Agents. Chemo. **36** pp876-878.

Oliva, B., Gordon, G G., McNicholas, P., Ellestad, G., Chopra, I. (1992). Evidence that tetracycline analogues whose targets are not the bacterial ribosome cause lysis of *Escherichia coli.* Antimicro. Agents. Chemother. **36**. pp913-919.

Omura, S. (1986). Philosophy of new drug discovery. Microbiol. Rev., **50**. pp259-279.

Omura, S., Tanaka, Y. (1984). Biochemistry, regulation and genetics of macrolide production. In *Macrolide antibiotics* Omura, S. (ed) Chemistry, Biology and practice. Academic Press, Orlando. Fl. US. pp199-229.

Park, B. H., Hendricks, M., Malamy, M. H., Tally, F. P., Levy, S. B. (1987). Cryptic tetracycline determinant (class F) from *Bacteroides fragalis* mediates resistance in *Escherichia coli* by actively reducing tetracycline accumulation. Anti. Microb. Agents. Chemother., **31**. pp1739-1743.

Park, B. H., Levy, S. B. (1988). The cryptic tetracycline resistance determinant on Tn*4400* mediates tetracycline degradation as well as tetracycline efflux. Anti. Microb. Agents Chemother., **32**. pp1797-1800.

Parker, J., Watson, R. J. Friesen, J. D. (1976). A relaxed mutant with an altered ribosomal protein L11. Mol. Gen. Genet., **144**. pp111-114.

Paulsen, I. T., Brown, M. H., Skurray, R. A. (1996). Proton dependent multidrug efflux systems. Microb. Rev., **60.** pp575-608.

Paulsen, I. T., Skurray, A. R. (1993). Topology, evolution and structure of two families of proteins involved in antibiotic and antiseptic resistance in eukaryotes and prokaryotes: an analysis. Gene, **124**. pp761-768.

Peden, K. W. (1983). Revised sequence of the tetracycline-resistance gene of pBR322. Gene, **22**. pp277-280.

Pigac, J., Alacevic, M. (1979). Mapping of oxytetracycline genes in *Streptomyces rimosus*. Period. Biol., **81**. pp575-582.

Potuckova, L., Keleman, G. H., Lonetto, M.A., Buttner, M. J., Kormanec, J. (1995). A new RNA polymerase sigma factor, sigma (F) is required for the late stages of morphological differentiation in *Streptomyces* spp.
Mol. Mocrobiol., **17**. pp37-48.

Revill, W. P., Bibb, M. J., Hopwood, D. A. (1996). Relationships between fatty acid

and polyketide synthases from *Streptomyces coelicolor* A3(2) : characterisation of the fatty acid synthase acyl carier protein. J. Bacteriol. **178**. pp5560-5567

Reynes, J. P., Calmels, T., Drocourt, D., Tiraby, G. (1988). Cloning, expression in *Escherichia coli* and nucleotide sequence of a tetracycline resistance gene from *Streptomyces rimosus*. Jnl. Gen. Microbiol., **134.** pp585-598.

Rhodes, P. M., Hunter, I. S., Friend, E. J. Warren, M. (1984). Recombinant DNA methods for the oxytetracycline producer *Streptomyces rimosus*. Biochem. Soc. Transac., **12**. pp586-587.

Rhodes, P. M., Winskill, N., Friend, E. J., Warren, M. (1981). Biochemical and genetic characterisation of *Streptomyces rimosus* mutants impaired in oxytetracycline biosynthesis. Jnl. Gen. Microbiol., **124**. pp329-338.

Roberts, M. C. (1994). Epidemiology of tetracycline resistance determinants. Trends Microbiol., **2.** pp353-357.

Rogalski, W. (1985). Chemical modification of the tetracyclines. In *Handbook of experimental pharmacology*. Hvlaka, J.J., Booth, J.H. (ed.'s) vol. **78.** Springer, Berlin. pp179-316.

Rouch, D. A., Cram, D. S., DeBerardino, D., Littlejohn, T. G., Skurray, R. A. (1990). Efflux mediated antiseptic resistance gene *qacA* from *Staphylococcus aureus*; common ancestry with tetracycline and sugar transport proteins. Mol. Microbiol., **4.** pp2051-2062.

Rubin, R. A., Levy, S. B., Heinrikson, R. L., Kezdy, F. J. (1990). Gene duplication in the evolution of the two complementing domains of Gram-negative bacterial tetracycline efflux proteins. Gene, **87**. pp7-13.

Rudd, B. A. M., Hopwood, D. A. (1979). Genetics of actinorhodin biosynthesis by

*Streptomyces coelicolor* A3(2). Jnl. Gen. Microbiol. **114**. pp35-43.

Ryan, M. J., Lotvin, J.A., Strathy, N., Fantini, S. E. (1997). Cloning of the biosynthetic pathway for chlortetracycline and tetracycline formation and cosmids useful therein. U.S. Patent **5589385-A** 31-DEC-1996.

Saier, M. H. Jr. (1990). Evolution of permease diversity and energy coupling mechanisms - an introduction. Res. Microbiol., **141**. pp281-286.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., Erlich, H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable polymerase. Science, **239**. pp487-491.

Salyers, A. A., Speer, B. S., Shoemaker, N. B. (1990). New perspectives in tetracycline resistance. Mol. Microbiol., **4**. pp151-156.

Satoh, A., Ogawa, H., Saomura, Y. (1975). Role and regulation mechanism of kanamycin acteyltransferase in kanamycin biosynthesis. Agric. Biol. Chem., **39**. pp2331-2336.

Sarrubi, E., Rudd, K. C., Xiao, H., Ikehara, K., Kalman, M., Cashel, M. (1989. The characterisation of the *spoT* gene of *Escherichia coli* J. Biol. Chem., **264**. pp15074-15082.

Schatz, P. J. & Beckwith, J. (1990). Genetic analysis of protein export in *Escherichia coli*. Annu. Rev. Genet. 24. pp215-248.

Schnappinger, D. Hillen, W. (1996). Tetracyclines, antibiotic action, uptake and resistance mechanisms. Arch. Microbiol., **165.** pp359-369.

Schuldiner, S. (1994). A molecular glimpse of vesicular monoamine transporters. Jnl. Neurochem., **62**. pp2067-2078.

Schuldiner, S., Shirvan, A., Lineal, M. (1995. Vesicular neurotransmitters: from bacteria to humans. Physiol. Rev., **75**. pp369-392.

Sheridan, R. P., Chopra, I. (1991. Original tetracycline efflux proteins; Conclusions from nucleotide sequence analyses. Mol. Microbiol., **5**. pp895-900.

Sherman, D. H., Malpartida, F., Bibb, M. J., Kieser, H. M., Hopwood, D. A. (1989). Structure and deduced function of the granaticin producing polyketide synthase gene cluster of *Streptomyces violaceoruber* Tu22. EMBO. Jnl., **8**. pp2717-2725.

Sloan, J., McMurray, L. M., Lyras, D., Levy, S. B., Rood, I. J. (1994). The *Clostridium perfringens tetP* determinant compromises two overlapping genes, *tetA(P)* which mediates active tetracycline efflux and *tetB(P)* which is related to the ribosomal protection family of resistance determinants. Mol. Microbiol., **11**. pp403-415.

Someya, Y., Niwa, A., Sawai, T., Yamaguchi, A. (1995). Site-specificity of the second-site suppressor mutation of the Asp-285-Asn mutant of metal-tetracycline/ $H^+$ antiporter of *Escherichia coli* and the effects of amino acid substitutions at the first and second sites. Biochemistry. **34**. pp7-12.

Speer, B. S., Bedzyk, L., Salyers, A. A. (1991). Evidence that a novel tetracycline resistance gene found on two *Bacteroides* transposons encodes an NADP-requiring oxidoreductase. J. Bacteriol., **173**. pp176-183.

Speer, B. S., Salyers, A. A. (1990). A tetracycline efflux gene on *Bacteroides* transposon Tn*4400* does not contribute to tetracycline resistance. J. Bacteriol., **172**. pp292-298.

Speer, B. S., Shoemaker, N. B., Salyers, A. A. (1992). Bacterial resistance to tetracycline; mechanisms, transfer and clinical significance. Clin. Microbiol. Rev., **5**. pp387-399.

Stone, M. J. & Williams, D. H. (1992). On the evolution of functional secondary metabolites, natural products. Mol. Microbiol., **6**, pp29-34.

Strauch, E., Takano, E., Bayliss, H. A Bibb, M. J. (1991). The stringent response in *Streptomyces coelicolor* A3(2). Mol. Mirobiol., **5.** pp289-298.

Studier, F. W. (1991). Use of bacteriophage T7 lysozyme to improve and inducible T7 expression system. J. Mol. Biol. **219**. pp37-44.

Studier, F. W. & Moffatt, B.A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high level expression of cloned genes. J. Mol. Biol. **189**. pp113-130.

Sugiyama, M., Paik, S-Y., Nomi, R. (1985.) Mechanism of self protection in a puromycin producing microorganism. Jnl. Gen. Microbiol., **131**. pp1999-2005.

Summers, R. G., Ali, A., Shen, B., Wessel, W. A., Hutchinson, C. R. (1995). Malonyl Coenzyme-A acyl carrier protein acyltransferase of *Streptomyces glaucesens*, a possible link between fatty acid and polyketide biosynthesis. Biochemistry, **34**. pp9389-9402.

Swift, S., Throup, J. P., Williams, P., Salmond, G. P., Stewart, G. S. A. B. (1996). Quorum sensing; a population density component in the determination of bacterial phenotype. TIBS **21**. pp214-219.

Tabor. S. & Richardson, C. C. (1985). A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc. Nat. Acad. Sci (USA). **82**. pp1074-1078.

Takano, E., Gramajo, H. C., Strauch, E., Andres, N., White. J., Bibb, M. J. (1992). Transcriptional regulation of the *redD* transcriptional activator gene accounts for growth phase dependent production of the antibiotic undecylprodigiosin in

*Streptomyces coelicolor* A3(2). Mol. Microbiol., **6.** pp2797-2804.

Takano, E., White, J., Thompson, C. J., Bibb, M. J. (1995) Construction of thiostrepton-inducible, high copy-number expression vectors for use in *Streptomyces* spp. Gene. **166**. pp133-137.

Tanaka, K., Shiina, T. Takahashi, H. (1988). Multiple principal sigma factor homologues in eubacteria: Identifcation of the "*rpoD* box". Science, **242**. pp1040-1042.

Tanaka, K., Shiina, T. Takahashi, H. (1991). Nucleotide sequence of genes *hrdA*, *hrdC* and *hrdD* from *Streptomyces coelicolor* A3(2) having similarity to *rpoD* genes. Mol. Gen. Genet., **229**. pp334-340.

Taylor, D.E. (1986). Plasmid mediated tetracyline resistance in *Campylobacter jejuni*- Expression in *Escherichia coli* and identification of homology with Streptococcal TetM determinant. Jnl Bacteriol. **165**. pp1037-1039.

Taylor, R. D.(1992). Ph.D Thesis, University of Glasgow

Tercero JA, Lacalle RA, Jimenez A. (1993). The *pur8* gene from the pur cluster of *Streptomyces alboniger* encodes a highly hydrophobic polypeptide which confers resistance to puromycin. Eur. J. Biochem. **218**. pp963-971.

Thamchaipenet, A. (1994). Ph.D. Thesis. University of Glasgow.

Thanassi, D. G., Suh, G. S. B., Nikaido, H. (1995. Role of outer membrane barrier in efflux mediated tetracycline resistance of *Escherichia coli*. J. Bacteriol., **177.** pp998-1007.

Thomas, R., Williams, D. J. (1983). Oxytetracyline biosynthesis: origin of the carboxamide substituent. J. Chem. Soc. Chem. Commun., **12**. pp677-679.

Thompson, J., Schmidt, F., Cundliffe, E. (1982). Site of action of a ribosomal RNA methylase conferring resistance to thiostrepton. J. Biol. Chem., **257**. pp7915-7917.

Towbin, H., Staehlin, T., Godon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications.Proc. Nat. Acad. Sci., **76**. pp4350-4354.

Varela M. F., Sansom, C. E., Griffith J. K. (1995). Mutational analysis and molecular modelling of an amino acid sequence motif conserved in antiporters but not symporters in a transporter superfamily. Mol. Membr. Biol., **12**. pp:313-319.

Voet, D Voet JG (1990). Chapter 23, Lipid Metabolism in "*Biochemistry*" John Wiley & sons (pub.)

Vogtli, M., Chang, P. C., Cohen, S. N. (1994). AfsR2: a previously undetected gene encoding a 63 amino acid protein that stimulates antibiotic production in *Streptomyces lividans*. Mol. Microbiol., **14**. pp643-653.

Von Heijne, G. (1986). The distribution of positively charged residues in bacterial inner membrane proteins correlates with the transmembrane topology. EMBO Jnl., **5**. pp3021-3027.

Vujaklija, A. D., Horinouchi, S., Beppu, T. (1993). Detection of an A-factor responsive protein that binds to the upstream activation sequences of StrR, a regulatory gene for streptomycin synthesis in *Streptomyces griseus*. J. Bacteriol. **175**. pp2652-2661.

Vujaklija, A. D., Ueda,K., Hong, S-K., Beppu, T., Horinouchi, S. (1991). Identification of an A-factor dependent promoter in the streptomycin biosynthetic gene cluster of *Streptomyces griseus*. Mol. Gen. Genet., **229**. pp119-128.

Walker, J. B. & Skorvaga, M. (1973a). Phosphorylation of streptomycin and dihydrostreptomycin by *Streptomyces*. Enzymatic synthesis of different diphosphorylated derivatives. J. Biol Chem. **248**. pp2435-2440.

Walker, J. B. & Skorvaga, M. (1973b). Streptomycin biosynthesis and metabolism. Phosphate transfer from dihydrostreptomycin 6-phosphate to inosamines, streptamine, and 2-deoxystreptamine. J. Biol. Chem. **248**. pp2441-2446.

Wang, I. K., Reeves, C., Gaucher, G. M. (1990). Isolation and sequencing of a genomic DNA clone containig the 3' terminus of the 6 methyl salicylic acid polyketide synthetase gene of *Penicillium urtiacae*. Can. Jnl. Micro., **133**. pp86-95.

Waters, S. H., Rogowsky, P. Grinstead, J., Altenbuchner, J., Schmitt, R. (1983). The tetracycline resistance determinants of RP1 and Tn*1721*; nucleotide sequence analysis. Nucleic Acids Res., **11**. pp6089-6105.

White, J. & Bibb, M. J. (1997). *bldA* dependence of undecylprodigiosin production in *Streptomyces coelicolor* A3(2) involves a pathway-specific regulatory cascade. J. Bacteriol. **179**. pp627-633.

Wirth, R., Muscholi, A., Wanner, G. (1996). The role of pheremones in bacterial interactions. Trends Microbiol. **4**. pp96-102.

Xiao, H., Kalman, M., Ikehara, K., Zemel, S. Glaser, G., Cashel, M. (1991). Residual guanosine 3' 5' bispyrophosphate activity of *relA* null mutants can be eliminated by *spoT* mutations. J. Biol. Chem., **266**. pp5980-5990.

Yamaguchi, A., Inagaki, Y., Sawai, T. (1995). Second-site supressor mutations for the Asp-66-Cys mutant of the transposon Tn*10* encoded metal-tetracycline / $H^+$ antiporter of *Escherichia coli.* Biochemistry, **34**. pp11800-11806

Yamaguchi, A., Kimura, T., Sawai, T. (1994). Hot spots for sulfhydryl inactivation of

cys mutants in the widely conserved sequence motifs of the metal-tetracycline $H^+$ antiporter of *Escherichia coli*. Jnl. Biochem. 115. pp958-964

Yamaguchi, A., Kimura, T., Someya, T., Sawai, T. (1992b). Metal tetracycline/$H^+$ antiporter of *Escherichia coli* encoded by transposon Tn*10* : The structural resemblence and functional difference in the role of the duplicated sequence motif between hydrophobic segment 2 and segment 3 and segment 8 and segment 9. J. Biol. Chem., **267**. pp19155-19162.

Yamaguchi, A., Ono, N., Akasaka, T., Noumi, T., Sawai, T. (1990). Metal tetracycline/$H^+$ antiporter of *Escherichia coli* encoded by transposon, Tn*10* : The role of the conserved dipeptide, $Ser^{65}$-$Asp^{66}$, in tetracycline transport. J. Biol. Chem., **265**. pp15525-15530.

Yamaguchi, A., Ono, N., Akasaka, T., Sawai, T. (1992c) Serine residues responsible for tetracycline transport are on a vertical stripe including Asp-84 on one side of transmembrane helix 3 in transposon Tn*10*-encoded tetracycline/ $H^+$ antiporter of *Escherichia coli*. FEBS letters, **307**. pp229-232.

Yamaguchi, A., Someya, Y, Sawai, T. (1992a). Metal tetracycline/$H^+$ antiporter of *Escherichia coli* encoded by transposon Tn*10*. The role of a conserved sequence motif, GXXXXRXGRR, in a putative cytoplasmic loop between helices 2 and 3. J. Biol. Chem., **267**. pp19155-19162.

Yamaguchi, A., Someya, Y, Sawai, T. (1993). The *in vivo* assembly and function of the N- and C-terminal halves of the Tn*10* encoded TetA protein in *Escherichia coli*. FEBS letters, **324**. pp131-135.

Yanish-Peron, C., Viera, J., Messing, J. (1985). Improved M13 phage cloning vectors and host strains: Nucleotide sequence of the M13mp18 and pUC19 vectors. Gene **33**. pp103-119.

Zen, K.H., McKenna, E. Bibbi, E., Hardy, D., Kaback, H. R. (1994). Expression of lactose permease in contiguous fragments as a probe for membrane spanning domains. Biochemistry. **33**. pp8198-8206.

Zhao, J., Aoki, T. (1992). Nuceotide sequence analysis of the class G tetracycline resistance determinant from *Vibrio anguillarum*. Micrbiol. Immunol., **36**. pp1051-1060.

GLASGOW
UNIVERSITY
LIBRARY