Construction, Characterisation and Utilisation of a Novel E. coli Expression Vector

by

STEPHEN P CHAMBERS

This thesis is presented for the degree of Doctor of Philosophy in the Department of Biological Sciences, University of Warwick.

This work was carried out between 1983-1987, while employed at the Biotechnology Division, Centre for Applied Microbiology & Research, Porton Down, Salisbury.

JANUARY 1989
Dedicated to my Mother and Father
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Contents</td>
<td>1</td>
</tr>
<tr>
<td>List of Figures</td>
<td>viii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>xii</td>
</tr>
<tr>
<td>Abbreviations</td>
<td>xiii</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>xv</td>
</tr>
<tr>
<td>Summary</td>
<td>xvi</td>
</tr>
</tbody>
</table>

## CHAPTER 1

1 Introduction

1.1 Plasmids and Genetic Manipulation of *E. coli*                   2
1.2 Plasmids as Cloning Vectors in *E. coli*                        3
1.3 Plasmids as Expression Vectors in *E. coli*                     6
1.4 High Expression of Cloned Genes                                 7
1.5 Promoters                                                       8
1.5.1 Promoter Regulation                                           11
1.6 Translation Initiation Sequences                                15
1.7 Plasmid Stability                                               16
1.7.1 Replication Control of the ColEl Plasmids                     16
1.7.2 Partitioning                                                  21
1.8 Plasmid Copy Number                                             25
1.9 Protein Stability                                               26
1.10 Host Physiology                                                29
1.11 Continuous Cultivation                                         30
1.11.1 Theory of Continuous Culture                                 30
1.12 Aims of Thesis                                                 34
## CHAPTER 2

### 2.1 Materials

#### 2.1.1 Chemicals

#### 2.1.2 Enzymes

#### 2.1.3 Media

#### 2.1.4 Buffers

#### 2.1.5 Bacterial Strains and Plasmids

### 2.2 Methods

#### 2.2.1 Storage and Growth of Bacterial Strains

#### 2.2.2 Beta-lactamase Plate Assay

#### 2.2.3 Carboxypeptidase Plate Assay

#### 2.2.4 Screening for Recombinants using X-Gal Plates

#### 2.2.5 Transformation of E. coli Competent Cells

#### 2.2.6 Small Scale Plasmid Preparations

#### 2.2.7 Large Scale Plasmid Preparations

#### 2.2.8 Isolation of Plasmid DNA by Centrifugation on a CsCl Gradient

#### 2.2.9 Plasmid Purification

##### 2.2.9.1 Phenol Extraction of DNA

##### 2.2.9.2 Precipitation of DNA

#### 2.2.10 Treatment of DNA with Enzymes

##### 2.2.10.1 Restriction Endonuclease

##### 2.2.10.2 T4 Polynucleotide Ligase

##### 2.2.10.3 DNA Polymerase, Large Fragment (Klenow)

##### 2.2.10.4 Alkaline Phosphatase

#### 2.2.11 Agarose Gel Electrophoresis
2.2.11 Vertical Gels
2.2.11.1 Horizontal Gels
2.2.11.2 Mini Gels
2.2.11.3 Sizing DNA Fragments
2.2.11.4 Tracking Dye
2.2.11.5 Recovery of DNA from Agarose Gels
2.2.11.6 Quantitation of DNA
2.2.12 Sequencing with Chain Terminating Dideoxynucleotides
2.2.12.1 M13 Sequencing Vectors
2.2.12.2 Large Scale Preparation of Bacteriophage
2.2.12.3 Transfection
2.2.12.4 Preparation of Templates for Sequencing
2.2.12.5 Sequence Reactions
2.2.12.6 Acrylamide Gel Fraction
2.2.12.7 Running Conditions
2.2.12.8 Fixing and Autoradiography
2.2.13 Colony Hybridization
2.2.14 Carboxypeptidase Assay
2.2.15 Chloramphenicol Acetyl Transferase Assay
2.2.16 Beta-lactamase Assay
2.2.17 Protein Assay
2.2.18 Polyacrylamide Gel Electrophoresis
2.2.19 Linear Gel
2.2.20 Gradient Gel
2.2.20.4 Running Conditions 73
2.2.20.5 Staining and Destaining 73
2.2.20.6 Quantitation of Protein Components in Gel Electrophoresis 74

2.2.21 Estimation of Copy Number 74
2.2.21.1 Measurement of Gene Dosage 74
2.2.21.2 Measurement of Plasmid/Chromosome Ratio 75
2.2.21.3 Definition of Plasmid Copy Number 75

2.2.22 Continuous Culture 77
2.2.22.1 Apparatus and Culture Conditions 77

CHAPTER 3 80
3. Effect of par Locus on Plasmid Stability, Copy Number and Host Fitness 80

Introduction 81

Results and Discussion 82
3.1 Plasmid Stability under Continuous Cultivation 82
3.1.1 Stability of pAT153 82
3.1.2 Stability of pNM21 87
3.1.3 Stability of pAT153 par 91
3.1.4 Stability of pNM21 par 91
3.1.5 Stability of pJCB 91
3.1.6 Stability of pUC8 par 98

3.2 Can the Calculated Copy Numbers Account for the Varying Instabilities of the Plasmids Examined? 98
3.3 Plasmid Copy Number Changes During Cultivation 106
3.4 Effect of par on Copy Number 108
CHAPTER 4

4. Construction and Characterisation of a Series of Novel Cloning Vectors

Introduction

Results and Discussion

4.1 Construction of the pMTL Backbone

4.2 Derivation of Novel Linker Regions

4.2.1 pMTL 20/21

4.2.2 pMTL 22/23

4.2.3 pMTL 24/25

4.3 Mobilisation of pMTL and pUC Vectors

4.4 Copy Number of pMTL and pUC Vectors

4.5 Sequence Variation Between pBR322 and pUC19 RNAI/RNAII Regions

4.6 Construction of pMTL par

4.6.1 Stability of pMTL 20 par

Concluding Remarks

CHAPTER 5

5. Examination of Pl Promoter Systems

Introduction

Results and Discussion

5.1 Expression Vector Constructions Examined
5.2 Host Strains for Induction 153

5.3 P\_ Directed Expression of CPG\_ in the Host 154

K\_ AH\_Atro

5.3.1 Effect of Copy Number on Expression 158

5.4 P\_ Directed Expression of CPG\_ in the Host W5445 163

5.4.1 What Effect Does Cloning the cl Gene into pNM502 Have on Induced Expression? 165

5.5 P\_ Directed Expression of Chloramphenicol Acetyl-transferase (CAT) in the Host W5445 166

5.6 Construction of pCM10 170

5.7 Effect of Induction on Plasmid Stability 173

Concluding Remarks 177

CHAPTER 6 180

6. Scale-up of Production of CPG\_ Utilising pCM10 180

Introduction 181

Results and Discussion 181

6.1.1 Fermentation Medium 181

6.1.2 Seed Preparation 182

6.1.3 Fermentation 182

6.1.4 Harvesting 183

6.1.5 Monitoring of Culture 184

6.1.5.1 Culture Conditions 184

6.1.5.2 Biochemical Analysis 185

6.1.5.3 Genetic Analysis 185

6.2 151 Small Scale Culture 185

6.3 Large Scale Production Culture 190
Concluding Remarks

CHAPTER 7

7. General Discussion

References

Appendix

Page

197

198

199

203

232
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Fig</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>The Origin of pBR322</td>
<td>5</td>
</tr>
<tr>
<td>1.2</td>
<td>A Schematic Representation of the Transcription and Translation Processes</td>
<td>10</td>
</tr>
<tr>
<td>1.3</td>
<td>The Consensus Sequence and the -10 and -35 Regions of lac, trp, PL, tacI and tacII Promoters</td>
<td>12</td>
</tr>
<tr>
<td>1.4</td>
<td>Replication Control Region of CoIE1</td>
<td>19</td>
</tr>
<tr>
<td>1.5</td>
<td>The pSC101 par Fragment</td>
<td>23</td>
</tr>
<tr>
<td>1.6</td>
<td>Nucleotide Sequence of cpr Gene</td>
<td>36</td>
</tr>
<tr>
<td>2.1</td>
<td>Standard Curve Used for the Determination of DNA Concentrations</td>
<td>65</td>
</tr>
<tr>
<td>2.2</td>
<td>Estimation of Copy Number using the Plasmid/Chromosome Ratio</td>
<td>76</td>
</tr>
<tr>
<td>2.3</td>
<td>Flow Diagram of a Chemostat</td>
<td>78</td>
</tr>
<tr>
<td>3.1</td>
<td>Plasmid pBR322</td>
<td>83</td>
</tr>
<tr>
<td>3.2</td>
<td>Segregational Stability of pAT153 in the Host W5445</td>
<td>85</td>
</tr>
<tr>
<td>3.3</td>
<td>Plasmid pNM21</td>
<td>88</td>
</tr>
<tr>
<td>3.4</td>
<td>Segregational Stability of pNM21 in the Host W5445</td>
<td>89</td>
</tr>
<tr>
<td>3.5</td>
<td>Variation of CPG2 Expression and Segregational Stability of pNM21 in the Host W5445 During Large Scale Fermentation</td>
<td>90</td>
</tr>
<tr>
<td>3.6a</td>
<td>Plasmid pAT153 par</td>
<td>92</td>
</tr>
<tr>
<td>3.6b</td>
<td>Determination of the Orientation of the par Locus within pAT153 par</td>
<td>93</td>
</tr>
<tr>
<td>3.7</td>
<td>Segregational Stability of pAT153 par in the Host W5445</td>
<td>94</td>
</tr>
<tr>
<td>3.8</td>
<td>Plasmid pNM21 par</td>
<td>95</td>
</tr>
</tbody>
</table>
3.9 Segregational Stability of pNM21 par in the Host W5445
3.10 Plasmid pUC8
3.11 Segregational Stability of pUC8 in the Host W5445
3.12 Plasmid pUC8 par
3.13 Segregational Stability of pUC8 par in the Host W5445
3.14 Copy Number and Segregation Frequency of Randomly Partitioned Plasmids
3.15 Reduction in Copy Number of pAT153 During Continuous Culture
3.16 Variation in Maintenance of pAT153 par and pAT153 in a Population of W5445 at Steady State Under Phosphate Limitation Conditions
4.1 The pUC Series of Plasmids
4.2 The Relationship Between pBR322, pAT153 and the pUC and pMTL Backbones
4.3 Construction of the pMTL Backbone
4.4 The Multiple Cloning Sites of the pMTL Vectors and their Complementary M13 Vectors
4.5 The Two Complementary Oligonucleotides used in the Construction of the pMTL20 and pMTL21 Linkers
4.6 Restriction Digests of pMTL20
4.7 The Two Pairs of Complementary Oligonucleotides used in the Construction of the pMTL22 and pMTL23 Linkers
<table>
<thead>
<tr>
<th>Fig.</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.8</td>
<td>The Construction of pMTL24 and pMTL25</td>
<td>130</td>
</tr>
<tr>
<td>4.9</td>
<td>Whole Cell Lysates Run on an Agarose Gel</td>
<td>135</td>
</tr>
<tr>
<td>4.10</td>
<td>DNA Sequencing Gel of the 3' Coding Sequence of the RNAI/RNAII Region in pBR322 and pMTL/pUC</td>
<td>137</td>
</tr>
<tr>
<td>4.11</td>
<td>Segregational Stability of pMTL20par, pMTL20 and pUC8 in the Host JM83</td>
<td>140</td>
</tr>
<tr>
<td>4.12</td>
<td>Nucleotide Sequence of pMTL20par</td>
<td>145</td>
</tr>
<tr>
<td>5.1</td>
<td>Expression of CPG2 by pNM830, when Regulated by the Compatible plasmid pNM52 in the Host JM83.</td>
<td>149</td>
</tr>
<tr>
<td>5.2</td>
<td>Schematic Diagram of (a) pNM232 and (b) pNM502</td>
<td>152</td>
</tr>
<tr>
<td>5.3</td>
<td>Expression of CPG2 by pNM232 and pNM502 in the Host K12AH1Atrp Carried Out in 1:1 Batch Cultures</td>
<td>156</td>
</tr>
<tr>
<td>5.4</td>
<td>Expression of CPG2 by pNM232 and pNM502 in the Host K12AH1Atrp Carried Out in 5:1 Batch Cultures</td>
<td>157</td>
</tr>
<tr>
<td>5.5</td>
<td>Copy Number of pNM232 and pNM502 Pre- and Post-Induction in the Host K12AH1Atrp</td>
<td>159</td>
</tr>
<tr>
<td>5.6</td>
<td>Expression of CPG2 by pNM502 in the Host W5445 Carried out in 5:1 Batch Cultures</td>
<td>164</td>
</tr>
<tr>
<td>5.7</td>
<td>Expression of CPG2 by pNM502 cl in the Host W5445 Carried out in 5:1 Batch Cultures</td>
<td>167</td>
</tr>
<tr>
<td>5.8</td>
<td>Construction of pCM10</td>
<td>171</td>
</tr>
<tr>
<td>5.9</td>
<td>Expression of CPG2 by pCM10 in the Host W5445 Carried out in 5:1 Batch Cultures</td>
<td>172</td>
</tr>
<tr>
<td>5.10</td>
<td>Segregational Stability of pNM232 and pNM502 when Induced in the Host K12AH1Atrp</td>
<td>174</td>
</tr>
<tr>
<td>5.11</td>
<td>Segregational Stability of pNM232 and pNM502 with pcl857 and pNM502 cl and pCM10 in the Host W5445</td>
<td>176</td>
</tr>
<tr>
<td>6.1</td>
<td>Fermentation Parameters of a 15:1 Small-Scale Culture</td>
<td>187</td>
</tr>
<tr>
<td>Fig.</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>------</td>
<td>------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>6.2</td>
<td>Levels of CPG2 Expression During Fermentation of 15 l Small Scale Culture</td>
<td>188</td>
</tr>
<tr>
<td>6.3</td>
<td>Segregational Stability of pCM10 in W5445 During 15 l Small Scale Culture</td>
<td>189</td>
</tr>
<tr>
<td>6.4</td>
<td>Fermentation Parameters of Seed Culture</td>
<td>191</td>
</tr>
<tr>
<td>6.5</td>
<td>Fermentation Parameters of Production Culture</td>
<td>192</td>
</tr>
<tr>
<td>6.6a</td>
<td>Levels of CPG2 Expression During Fermentation of Production Culture</td>
<td>193</td>
</tr>
<tr>
<td>6.6b</td>
<td>Levels of CPG2 Expression During Fermentation of Production Culture as Observed on an SDS PAGE</td>
<td>194</td>
</tr>
<tr>
<td>6.7</td>
<td>Segregational Stability of pCM10 in W5445 During Production Culture</td>
<td>195</td>
</tr>
<tr>
<td>6.8</td>
<td>Variation of Copy Number During Fermentation</td>
<td>196</td>
</tr>
<tr>
<td>Table</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>-------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>1.1</td>
<td>Factors Affecting the Expression of Cloned Genes</td>
<td>9</td>
</tr>
<tr>
<td>2.1</td>
<td>Bacterial Strains used in this Study</td>
<td>51</td>
</tr>
<tr>
<td>2.2</td>
<td>Plasmids used in this Study</td>
<td>52</td>
</tr>
<tr>
<td>3.1</td>
<td>Calculated Copy Numbers from Plasmid/Chromosome Ratios Obtained from Densitometry of EtBr-Stained Electrophoresis Gels</td>
<td>86</td>
</tr>
<tr>
<td>3.2</td>
<td>Calculated Copy Number per Dividing Cell</td>
<td>105</td>
</tr>
<tr>
<td>3.3</td>
<td>Summary of Competition Events Carried out in Continuous Culture</td>
<td>111</td>
</tr>
<tr>
<td>3.4</td>
<td>Growth Rate Variation Between Organisms Containing pAT153, pAT153 par and Plasmid Free Cells</td>
<td>114</td>
</tr>
<tr>
<td>4.1</td>
<td>Mobilisation Frequencies of the Cloning Vectors</td>
<td>132</td>
</tr>
<tr>
<td>4.2</td>
<td>Calculated Copy Numbers from Plasmid/Chromosome Ratios, Obtained from Densitometry of EtBr-Stained Electrophoretic Gels</td>
<td>134</td>
</tr>
<tr>
<td>5.1</td>
<td>Variation in Growth Rates Pre- and Post-Induction of pNM232 and pNM502 in the Host K12AH1Apr</td>
<td>161</td>
</tr>
<tr>
<td>5.2</td>
<td>Variation in Expression and Relative Promoter Strengths of the Pl Promoter in the Constructions and Hosts under Examination</td>
<td>162</td>
</tr>
<tr>
<td>5.3</td>
<td>A Comparison of the Codon Usage of the cpg Gene and the cat Gene.</td>
<td>168</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>ccc</td>
<td>covalently closed circular</td>
<td></td>
</tr>
<tr>
<td>kb</td>
<td>kilo base</td>
<td></td>
</tr>
<tr>
<td>bp</td>
<td>base pair</td>
<td></td>
</tr>
<tr>
<td>kD</td>
<td>kilo dalton</td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>resistant</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>sensitive</td>
<td></td>
</tr>
<tr>
<td>Ap</td>
<td>ampicillin</td>
<td></td>
</tr>
<tr>
<td>Tc</td>
<td>tetracycline</td>
<td></td>
</tr>
<tr>
<td>Km</td>
<td>kanamycin</td>
<td></td>
</tr>
<tr>
<td>Co</td>
<td>chloramphenicol</td>
<td></td>
</tr>
<tr>
<td>ORI</td>
<td>origin of replication</td>
<td></td>
</tr>
<tr>
<td>rBS</td>
<td>ribosome binding site</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>Shine-Dalgarno</td>
<td></td>
</tr>
<tr>
<td>ROP</td>
<td>repressor of primer</td>
<td></td>
</tr>
<tr>
<td>mPF</td>
<td>partition function</td>
<td></td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
<td></td>
</tr>
<tr>
<td>X-Cna</td>
<td>5-bromo-4-chloro-3-indoly-Beta-D-galactopyranoside</td>
<td></td>
</tr>
<tr>
<td>IPTG</td>
<td>isopropyl-Beta-D-thiogalactoside</td>
<td></td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
<td></td>
</tr>
<tr>
<td>DTNB</td>
<td>5,5'-dithio-bis-2-nitrobenzoic acid</td>
<td></td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulphate</td>
<td></td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
<td></td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylene diamine tetra acetic acid</td>
<td></td>
</tr>
<tr>
<td>EtBr</td>
<td>ethidium bromide</td>
<td></td>
</tr>
<tr>
<td>CAT</td>
<td>chloramphenicol acetyl transferase</td>
<td></td>
</tr>
<tr>
<td>CPC2</td>
<td>carboxypeptidase G2</td>
<td></td>
</tr>
<tr>
<td>CPD</td>
<td>carboxypeptidase gene</td>
<td></td>
</tr>
</tbody>
</table>
bla  beta-lactamase gene
tet  tetracycline resistance gene
lacZ beta-galactosidase gene
cat chloramphenicol acetyl transferase gene
PAGE polyacrylamide gel electrophoresis
APS ammonium persulphate
ATP adenosine triphosphate
dNTP deoxynucleotide triphosphate
ddNTP dideoxynucleotide triphosphate
CIP calf intestinal phosphatase
PEG poly ethylene glycol
TEMED tetramethyl ethylene diamine
OD optical density
UV ultra violet
V volts
mA milli Amps
a centripetal force equal to gravitational acceleration
nm nano meters
klm volumetric oxygen transfer coefficient
DOT dissolved oxygen tension
rpm revolutions per min
LDW Lab. dry weight
psi pounds per square inch
vvm volume of air/volume of culture/min
w/v weight to volume ratio
v/v volume to volume ratio
ACGM Advisory Committee on Genetic Manipulation
ACKNOWLEDGEMENTS

I should like to thank my supervisors R Sharp, N Mann and N Minton who have provided help and advice during the course of this work. I am also particularly indebted to Professor T Atkinson for his constant encouragement.

I would also like to thank Val Lancaster for typing this thesis.
SUMMARY

At the onset of this thesis the aim was to design and develop an expression system capable of producing high levels of expression of carboxypeptidase (CPG2). The previous plasmid employed (pNM21) delivered relatively low and variable levels of CPG2.

Plasmid segregational instability had been thought to be a potential cause of low yields of CPG2. The effect of the par locus on the stability of a series of related plasmids was examined. The utilisation of par within a vector did enhance its stability and fitness when examined under continuous cultivation. The par locus was therefore considered to be of significant value to an expression vector.

The par locus subsequently became an integral feature incorporated into any vector to be constructed. The copy numbers of the series of plasmids examined for stability revealed that pUC8 exhibited a higher copy number than pAT153. To investigate more fully this anomaly, a vector was constructed which provided the backbone for a series of cloning vectors (the pMTL series) and allowed the copy number variation to be examined. Sequencing pAT153 and pUC8 revealed a single base change, which could account for the copy number variation. The increase in copy number of pUC8 over pAT153 was incorporated in the pMTL series of vectors.

The induced expression of CPG2 was investigated using the Pl promoter on a high copy number vector, exhibiting the aforementioned base change. Incomplete repression of the Pl promoter resulted when the cl gene was located in the chromosome or on a co-resident low copy number vector. The high copy number of the vector and the Pl operator it carried, titrates out the effectiveness of the cl repressor molecule. This problem was overcome by incorporation of the cl gene into the Pl expression vector, elevating the gene dosage of the cl gene, sufficient to repress the Pl promoter.

The features investigated were combined in an expression vector pCM10 and used to express CPG2 in a series of small scale fermentations. The process was then scaled up to a 400 L fermentation, and a suitable process for large scale production developed.
CHAPTER 1

INTRODUCTION
INTRODUCTION

1.1 Plasmids and Genetic Manipulation in E. coli

Plasmids are autonomous, extrachromosomal genetic elements found in a wide variety of Gram-positive and Gram-negative bacteria. Plasmids usually exist within cells as double stranded, covalently closed circular (ccc) DNA molecules (Clewell and Helinski, 1969). They range in size from 1 kb to greater than 200 kb. Plasmids often confer a wide range of biological functions on the host cell, which can be used as phenotypic markers when selecting cells containing the plasmid. Plasmids may also specify the ability to transfer themselves from the host cell (donor) to a recipient cell by conjugation, which involves direct cell to cell contact. A method known as transformation, has been developed whereby naked plasmid DNA can be introduced into E. coli cells (Mandel and Higa, 1970; Cohen et al., 1972). The new phenotype conferred on the recipient by the plasmid (e.g., resistance to an antibiotic) allows a simple method of selection of cells which have been successfully transformed.

Along with the ability to isolate and transform plasmid DNA from one bacteria strain to another, the discovery of site specific restriction endonucleases (Roberts, 1981), particularly type II enzymes (Smith and Wilcox, 1970), made it possible to manipulate DNA molecules in vitro. Type II enzymes cleave duplex DNA to produce fully double-stranded (flush or blunt-ended) termini or fragments with single-stranded, self complementary (cohesive or sticky-ended) termini. The restriction endonuclease EcoRI was shown to produce DNA molecules with cohesive termini after cleavage (Mertz and Davis, 1972) which could be reannealed by the use of E. coli polynucleotide ligase (Dugaiczyk et
al., 1975) and transformed into E. coli cells. Blunt-ended DNA fragments can also be joined by the use of T4 polynucleotidase ligase (Heynaker et al., 1976). Further methods for joining together DNA molecules have been developed. The enzyme terminal-deoxynucleotidyl transferase (Chang and Bollum, 1974) can be used to add homopolynucleotide tails to DNA fragments to be joined. After mixing and annealing such DNA molecules can then be used directly to transform E. coli where repair of single-stranded gaps occur in vivo (Hutchinson and Halvorson, 1980). Synthetic DNA fragments (linkers) can also be used for linking together DNA molecules (Maniatis et al., 1978).

1.2 Plasmids as Cloning Vectors in E. coli

To be useful as a cloning vector, a plasmid should possess several properties. It should have a low molecular weight, making the plasmid more resistant to damage by shearing and therefore easier to handle. Also a large number of plasmid copies should be present in the cell. High copy number plasmids, facilitate isolation and manipulation of the plasmid DNA. It should also carry one or more selectable markers to allow identification of transformants and to maintain the plasmid in the bacterial population. Another requirement of a cloning vector is that it contains a single recognition site for one or more restriction enzymes in regions of the plasmid that are not essential for replication. Preferably, these restriction sites, into which foreign DNA can be inserted, should be located within the genes coding for selectable markers so that insertion of a foreign DNA fragment inactivates the gene. Finally, the cloning vector should exhibit a degree of stability, to overcome plasmid loss from a bacterial population in the absence of any selective pressure.
Although some 'natural' plasmids have been used as cloning vectors most cloning experiments now make use of 'purpose-built' vectors which are themselves the product of in vitro gene manipulations. ColE1 is a naturally occurring plasmid which is the progenitor of many of the cloning vectors used. pBR322 is one of the more widely known cloning vectors, utilising the ColE1 replicon while substituting colicin E1 immunity and production (which are not convenient genetic markers) with the beta-lactamase (bla) and tetracycline resistance (tet) genes of RSF2124 and pSC101 respectively (Fig. 1.1) (Bolivar et al., 1977a, b). pBR322 exhibits many of the qualities required by a successful cloning vector, a low molecular weight (2.6 MD), with a relatively high copy number of 50 copies per cell, containing bla and tet genes as genetic markers and a number of convenient restriction sites within these marker genes.

A whole generation of cloning vectors have developed from pBR322, improving on the basic requirements for a cloning vector achieved by pBR322. (for review see Balbas et al., 1986). pAT153 was constructed by the removal of the Haell B and G fragments from pBR322, resulting in a 3-fold increase in its copy number (Twigg and Sherratt, 1980). This additional advantage of increased levels of plasmid content results in an increase in gene dosage of the plasmid specified gene products. A further extension to the usefulness of cloning vectors has been made by inserting polylinkers, containing closely spaced sites for several restriction enzymes, into non-essential regions of the vector (Davison et al., 1984; Heustersprente and Davison, 1984; Heustersprente et al., 1985). Another development of the cloning vector was made by Vieria and Messing (1982) who constructed the pUC plasmid series. A 2.30 kb PvuII, EcoRI fragment
The Origin of pBR322

The plasmid pBR322 was constructed from the his and tet genes of the transposon Tn1 from RSF2124 and pSC101 respectively, combined with replication elements of a ColE1-like plasmid.
from pBR322, provides the cloning vector with its replicon and the ampicillin resistance (Ap\(^R\)) antibiotic marker. In addition the pUC plasmids contain multiple restriction site polylinkers within the coding region of the alpha-peptide of the E. coli lacZ gene product (Langley et al., 1975). Cloning into the polylinker region results in inactivation of the lacZ', so that the plasmid in a suitable E. coli host produces white colonies on medium containing 5-bromo-4-chloro-3-indoyl-beta-D-galacto-pyranoside (X-Gal), in contrast to the parental blue colour (Messing, 1983).

As with pBR322, the pUC series of plasmids have spawned a whole new generation of cloning vectors, each one slightly improving on the basic requirements of a cloning vector. An ever increasing number of unique restriction sites within the alpha-peptide coding sequence of lacZ' (Norrander et al., 1983) and various copy number derivatives have also been developed (Stewart et al., 1986a,b).

1.3 **Plasmids as Expression Vectors in E. coli**

Cloning vectors such as pBR322 and pAT153 can be employed as expression vectors, utilising the bla or tet gene promoters (Minton et al., 1983a). Unfortunately these promoters are partially constitutive and therefore provide unregulated expression which can be lethal to the cell. The pUC series of plasmids can also be used to obtain expression of cloned DNA as lacZ fusions when joined in-phase with the reading frame of the alpha-peptide. The regulation of transcription can be achieved in strains that overproduce the lacI encoded repressor (lacI\(^Q\)) by the addition of an inducer such as lactose, or its synthetic analogue, isopropyl-beta-D-thiogalactoside (IPTG).

The first experiments designed to overproduce foreign gene
products, the synthetic genes coding for the hormones somatostatin (Itakura et al., 1977) and the insulin A and B chains (Goeddel et al., 1979b), utilised cloning vectors. Once the foreign gene was in the cloning vector, it was put under transcriptional control by cloning a DNA cartridge containing the regulatory regions of lac upstream of the synthetic genes. The strategy for expression of heterologous or foreign protein has subsequently become more sophisticated and now a wide variety of expression vectors are available, into which a gene can be cloned and expression regulated.

1.4 High Expression of Cloned Genes

Much of the interest surrounding in vitro gene manipulation involves the commercial exploitation of bacteria to produce foreign proteins such as those from human, animal or viral sources. Detectable expression is no longer sufficient to satisfy the demands of production, high expression of the gene product is required.

Nature has provided us with many examples of the overproduction of gene products by bacteria. In E. coli the ribosomal proteins are one classic example; the related protein elongation factor Tu, for instance, can accumulate to 5% of the soluble cell protein. At the extreme, infection of E. coli with bacteriophages T4 or T7, leads to total remobilisation of the host replication, transcription and translation machinery to produce almost exclusively the viral-encoded products. Additionally, there are naturally occurring plasmids ColE2 and ColE3 which upon induction produces colicin. This then becomes the major synthesised protein in the cell (Tyler and Sherratt, 1975). Synthesis of a functional protein depends upon transcription of the appropriate gene, efficient translation of the mRNA and in many cases,
post-translational processing (Table 1.1). A failure to perform correctly any one of these processes can result in the failure of a given gene to be expressed. It is the role of an expression vector to maximise expression; this can be done at specific stages of the process and determines the components essential for an expression vector.

1.5 Promoters

To obtain high expression of a cloned gene transcription should be maximised using a strong promoter. A promoter is a DNA sequence that directs DNA dependent RNA polymerase to bind to DNA and to initiate RNA synthesis. Strong promoters cause mRNAs to be initiated at high frequency; while weak promoters direct the synthesis of rarer transcripts. A large number of promoters for E. coli RNA polymerase have been analysed (Hawley and McClure, 1983); comparison of the sequences reveals two highly conserved regions, one located about 10 bp (-10 region or Pribnow box; Pribnow, 1975) and the other about 35 bp (-35 region) upstream from the point at which transcription starts (+1) (Rosenberg and Court, 1979) (Fig. 1.2). These two regions are thought to be important in determining promoter strength because mutations that decrease the frequency of transcription usually decrease the amount of homology with these conserved sequences. Other more moderately conserved regions of promoters may also contribute to promoter strength. Also, the number of nucleotides that separate the conserved sequences is important for efficient promoter function. Mutations altering the spacing between the -10 region and -35 region in a lac promoter (Stefano and Gralla, 1982), and in the beta-lactamase promoter (Jaurin et al., 1981), change the strength of the promoter. Hawley and McClure (1983) found that in all cases the promoter strength
Factors Affecting the Expression of Cloned Genes

<table>
<thead>
<tr>
<th>Promoter strength</th>
<th>Translation initiation sequences</th>
<th>Codon choice</th>
<th>Transcriptional termination</th>
<th>Plasmid copy number</th>
<th>Plasmid stability</th>
<th>Protein stability</th>
<th>Host cell physiology</th>
</tr>
</thead>
</table>

Table 1.1
Fig. 1.2
A Schematic Representation of the Transcription and Translation Processes

Identifying the -10, -35 consensus promoter regions. Also the Shine-Dalgarno site and the AUG start codon.
was stronger if the spacing was moved closer to 17 bp and weaker if moved further away from 17 bp. Of the three *E. coli* promoters, lac UV5, trp and the P\textsubscript{L} promoter from coliphage lambda, most widely used in expression vectors, none shows absolute identity with the consensus sequence (Fig. 1.3). In terms of transcriptional strength yet more efficient promoters have been made which are hybrids of the trp and lac UV5 promoters. The two synthetic hybrid promoters tac I and tacII were eleven and eight times stronger than the lac promoter (deBoer et al., 1983).

Many *E. coli* genes are controlled by relatively weak promoters, therefore the expression of such genes can be increased by placing them downstream from efficient promoters. Eukaryotic promoters function extremely poorly, if at all, in *E. coli*, and efficient expression of eukaryotic proteins has been achieved only when the coding sequence is placed under the control of a strong *E. coli* promoter.

The effect of introducing a strong promoter into a vector must also be taken into account. Readthrough beyond the gene of interest can be an unnecessary energy drain on the cell, producing transcripts coding for non-essential protein. Also transcription from the promoter may interfere with an essential or regulatory gene (Stueber and Bujard, 1982). Transcription terminators at the end of the cloned gene will prevent the deleterious effects of readthrough from a strong promoter (Gentz et al., 1981).

### 1.5.1 Promoter Regulation

The most useful promoters for expressing heterologous genes in *E. coli* are those that are both strong and also regulated. Although
# Fig. 1.3

The Consensus Sequence, and the -10 and -35 Regions of lac, trp, Fl, tacI and tacII Promoters

<table>
<thead>
<tr>
<th></th>
<th>-35 REGION</th>
<th>-10 REGION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consensus</td>
<td>T T G A C A</td>
<td>T A T A A T</td>
</tr>
<tr>
<td>lac</td>
<td>T T T A C A</td>
<td>T A T A T T</td>
</tr>
<tr>
<td>trp</td>
<td>T T G A C A</td>
<td>T T A A C T</td>
</tr>
<tr>
<td>Fl</td>
<td>T T G A C A</td>
<td>G A T A C T</td>
</tr>
<tr>
<td>tacI</td>
<td>T T G A C A</td>
<td>T A T A A T</td>
</tr>
<tr>
<td>tacII</td>
<td>T T G A C A</td>
<td>T T T A A T</td>
</tr>
</tbody>
</table>
constitutive expression was capable of producing human growth hormone (Goeddel et al., 1979a) and calf prochymosin (Emtage et al., 1983) as 5% of the cells soluble protein, various problems emerged. It was found that in the absence of any selection for plasmids expressing the desired gene product, plasmid-free cells frequently appeared (Imanaka et al., 1980; Steuber and Bujard, 1982; Skogman et al., 1983). Freed from the plasmid burden, these cells proved to have a growth advantage over those carrying the expressing plasmid (Caulcott et al., 1985) and could therefore outgrow plasmid-bearing cells during a fermentation. As a result it is quite possible, during a production process involving several inoculum stages, for the proportion of cellular population carrying the plasmid to fall considerably. The discovery of these problems led to a general appreciation of the fact that expression of recombinant protein was a burden on the host cell (Carrier et al., 1983).

An alternative solution to this problem would be to operate a biological or metabolic switch that would enable product expression to be turned on and off, dividing cell growth and cell biosynthesis into two distinct phases. The switching on of a promoter at a stage in growth when cell densities are at their highest will result in the maximum protein synthesising capacity of the culture being achieved.

Two basic promoter systems are available for the development of controllable expression vectors: either a temperature sensitive repressor/promoter combination or one derived from a metabolically controlled gene could be used. The latter type of control system could be affected by the addition of an inducer to the growth medium. The lac promoter being regulated by the lac repressor can be induced by the addition of an inducer lactose or its analogue IPTG to the medium.
(Miller and Reznikoff, 1978). The trp promoter, regulated by the trp repressor, can be induced by the addition of 3-indolylacetic acid or 3-indoleacrylic acid to the medium (Morse et al., 1970). This promoter can also be regulated by the removal of a nutrient e.g., tryptophan depletion; the trp promoter being induced by tryptophan starvation (Miller and Reznikoff, 1978).

Unfortunately, metabolic promoters on high copy number plasmids have not proved very successful. In general, control of such systems is by a repressor protein synthesised from a single gene copy on the host chromosome. For example, in the absence of lactose, the lac repressor protein binds to the promoter region and switches off expression of any genes under its control. In the presence of lactose, it binds to the repressor protein, preventing it from binding to the lac promoter/operator. It has been shown that on high copy number plasmids the lac operator cannot be solely controlled by the chromosomal repressor gene (O'Farrell et al., 1978; Lee and Bailey, 1984) since there are insufficient repressor molecules. A similar situation occurs with the trp promoter system. Here the repressor protein is activated by tryptophan, and is then able to switch off the trp promoter. However, if present on a high copy number plasmid, the number of trp promoters in a cell would almost certainly titrate out the amount of trp repressor protein synthesized from the chromosomal gene (Hallewell and Emtage, 1980).

Metabolically-controlled promoters have in general been superceded by temperature controlled systems. P_L, the leftward promoter of bacteriophage lambda, provides a strong promoter system controllable by the temperature-sensitive lambda repressor gene product, cI857 (Isacoff et al., 1965). This gene product is denatured
at temperatures above 37°C, allowing initiation of transcription at the PL promoter when the culture is raised above 37°C. This promoter has proved to be the most successful for expression vector systems (Remaut et al., 1981) and their commercial exploitation (Simons et al., 1984; Waldman et al., 1983; Liu and Milman, 1983).

1.6 Translation Initiation Sequences

The efficiency of translation can also be influenced by manipulating the secondary structure of the ribosome-binding sites (RBS) (Gren, 1984). Efficient translation depends on ribosomes recognizing the initiation codon (AUG) at the start of a gene sequence. With most E. coli genes translation is initiated after a ribosome binds to the mRNA at the RBS. This is a region which is 3–9 nucleotides long, located 3–11 nucleotides upstream from the initiation codon (Shine and Dalgarno, 1975; Steitz, 1979) known as the Shine-Dalgarno (SD) sequence. It contains part of the sequence 5'-AGGAGGUG-3' which is complementary to the 3' terminus of the 16S ribosomal RNA (-CACCUCCUAOH). Binding of the ribosome to mRNA is thought to be promoted by base pairing between the SD sequence in the mRNA and the sequence at the 3' end of the 16S ribosomal RNA. The degree of complementation between these two sequences will therefore have a marked effect on the efficiency of translation of the mRNA. When expressing heterologous genes in E. coli it is essential that a RBS precedes the coding sequence. The precise location of the RBS site relative to the initiation codon and the composition of the intervening mRNA can affect dramatically the level of expression. A 100-fold increase in the level of beta-interferon expression was obtained by reducing from 11 nucleotides to 6 the distance between the RBS site and
the initiation codon of the beta-interferon gene (Shepard et al., 1982). The composition of the triplet immediately preceding the AUG start codon also affects the efficiency of translation. For translation of beta-galactosidase mRNA the most favourable combinations of bases in this triplet are UAU and CUU. If UUC, UCA or AGG replace UAU or CUU the level of expression was 20-fold less (Hui et al., 1984). Consensus RBS sequences have been made from hybrid and synthetic constructions but have met with limited success (Crowl et al., 1985).

1.7 Plasmid Stability

Although aggregational stability of plasmids is not considered to be a problem at the laboratory scale, in large scale production, however, plasmid instability can pose a very serious problem. The use of antibiotics at this scale is not desirable because of effectiveness, cost and waste disposal considerations.

In order that plasmids are maintained within bacteria cells, they must replicate at least once before cell division and be accurately partitioned to ensure that each daughter cell receives at least one plasmid copy at cell division.

1.7.1 Replication Control of the ColEl Plasmide

Plasmid ColEl is a small (6.6 kb) multicopy plasmid whose replication has been extensively studied both in vitro and in vivo (Staudenbauer, 1978a). Since the ColEl replicon is the replicon used in the vectors studied in this thesis, its replication properties are discussed (for reviews see Scott, 1984; Cesareni and Banner, 1985).
ColEl has been shown to replicate unidirectionally from a fixed origin (Lovett et al., 1974) by Cairns-type replication (Inselburg and Fuke, 1971) and more recent studies have demonstrated that initiation of ColEl DNA synthesis occurs at either one of 3 consecutive bases in the replication origin (Tomizawa et al., 1977).

Electron microscopic analysis of intermediates of ColEl replication in the presence and absence of various DNA gene products led to the proposal of a two-stage model for ColEl replication (Staudenbauer, 1978b). DNA Polymerase I extends the DNA chain from the RNA primer for 500 nucleotides, whereupon DNA Polymerase III continues to synthesise DNA on both strands. This explains why ColEl, unlike many other plasmids has an absolute dependence on PolI for replication (Kingsbury and Helinski, 1973). The observation that ColEl plasmids could replicate for 10 – 15 hours after the addition of chloramphenicol, an inhibitor of protein synthesis, demonstrated that ColEl does not require any plasmid encoded proteins for replication (Clewell, 1972). Moreover, since bacterial replication cannot continue, the host proteins necessary for ColEl replication must be stable.

ColEl replication is, however, sensitive to rifampicin, an RNA polymerase inhibitor, which first indicated that an RNA primer may be involved in the replication of this plasmid (Clewell, 1972). The first direct evidence that ColEl specified an RNA primer came from Itoh and Tomizawa (1980) who demonstrated in vitro that an RNA transcript initiating 553 bp upstream of the defined origin, was cleaved by RNAase H indicating that an RNA-DNA hybrid was formed. After cleavage, DNA polymerase I added deoxyribonucleotides to the primer RNA (RNA II). Formation of the pre-primer RNA-DNA hybrid is negatively controlled by
a small RNA molecule RNAI (108 nucleotides), which is encoded within the region specifying the pre-primer RNAII (Tomizawa et al., 1981). In order for inhibition by the anti-sense RNA molecule to be effective, RNAI binding to the pre-primer RNAII must occur prior to the extension of RNAII to the replication origin (Tomizawa and Itoh, 1981). Both RNAI and RNAII have the capacity to form three stem-loop structures (Morita and Oka, 1979) and the importance of secondary structure to RNAII processing has been demonstrated by replacing 'G' residues with inosine which weakly pairs with 'C' residues, disrupting secondary structure and inhibiting maturation of the primer.

Many in vitro experiments have suggested a mechanism for the binding of RNAI to the RNAII pre-primer (Tomizawa and Som, 1984) (Fig. 1.4). Inc mutations within any of the three single stranded loop regions, which do not alter the secondary structure of RNAI but affect the binding rate of RNAI to RNAII, demonstrated the importance of the loop regions to the binding process. Furthermore, the 5' end of RNAI is involved, since its removal greatly reduced the binding activity of RNAI to RNAII although once again the secondary structure of RNAI was not altered.

From these observations the "stepwise pairing model" was proposed. Initially, RNAI and RNAII form stem-loop structures; the two molecules make transient contact, termed "kissing", at the single stranded loops. This contact facilitates pairing at the 5' end of RNAI, which then propagates along the entire length of RNAI/RNAII, forming stable binding. In addition to RNAI, CoIE1 specifies a second product called ROP (repressor of primer) which is involved in replication control. The construction of a CoIE1 deletion derivative exhibiting a higher copy number which could be reduced in trans by
Interaction between the complementary RNAI and RNAlII molecules prevents processing of the pre-primer RNA (RNAlII) by RNAase H to form the active replication primer. The 63 residue protein, ROP, catalyses the interaction between RNAI and RNAlII, enhancing the negative replication control mechanism. This schematic representation of the replication region was described by Cenareti and Bannar (1985).
ColEl, first suggested that another region of ColEl was involved in copy number control (Twigg and Sherratt, 1980). Later, Cesareni et al. (1982) suggested that the rop gene located between the ColEl replication origin and mobility genes specified a 63 amino acid polypeptide (6.3 KD). This protein has subsequently been purified (Som and Tomizawa, 1983) and shown to inhibit transcription of the galK gene when it was fused to the primer promoter. Inhibition was strong when 135 bases of the primer region downstream of the promoter was fused to the galK gene but weak when only 52 bases were present, indicating that a region downstream of the first 52 bases was involved in the action of the ROP protein.

RNAI has been shown to be an absolute requirement for ROP mediated inhibitory action both in vitro (Lacatena et al., 1984) and in vivo (Cesareni et al., 1984). In vitro studies on pMB1, a derivative of ColEl, have shown that ROP affects primer formation in two ways, firstly by causing transcription termination to occur at nucleotide 220 in 10% - 20% of the primer transcripts and secondly, by increasing the ability of RNAI to inhibit RNAase H processing of the primer.

Experiments carried out in vitro have demonstrated that ROP enhances the binding rate of RNAI to RNAII under physiological conditions (Tomizawa and Som, 1984) by catalysing a stage which precedes the propagation of stable pairing between the two molecules. Also, there is an inverse correlation between copy number and the rate of binding of RNAI to RNAII. The highest binding rate was observed for the wild type RNAI/RNAII species from ColEl, this was enhanced 2-fold by ROP. These results were substantiated by in vivo analysis of plasmid copy numbers. ROP+ plasmids showed approximately a 3-fold increase in copy number.
Thus, ROP enhances the inhibitory activity of RNA I, however, the exact mechanism of replication inhibition is unknown. It is generally accepted that binding of RNAI to RNAII prevents processing of the pre-primer by RNAase H, but inhibition of expression of fusion proteins between galK or lacZ and part of the pre-primer argues to the contrary. It has been suggested that binding of RNAI makes RNAII more sensitive to degradation by ribonucleases, supporting the finding in another system involving binding of an RNA molecule to the complementary region of the 5' part of an mRNA (Coleman et al., 1984).

Recently, the structure of the ROP protein has been determined and shown to be a small rigid dimer of exactly two-fold symmetry (Cesareni and Banner, 1985). They have proposed that each sub-unit binds to either RNAI or RNAII and positions them correctly for the transient initial loop-loop interaction.

1.7.2 Partitioning

In order that plasmids are stably maintained, in addition to controlled replication, they must be partitioned accurately at cell division so that each daughter cell receives at least one plasmid copy. The hereditary stability of low copy number plasmids such as F, pSC101 and R1 can only be explained by accurate partitioning occurring at cell division. Indeed, several low copy number plasmids have been shown to encode their own partition function (par) (Nordstrom et al., 1980; Ogura and Hiraga, 1983). For ColEl and related plasmids however, there have been no substantiated reports of par-like functions and it has been suggested that active partitioning may not be critical for multicopy plasmids and random partitioning may suffice (Summers and Sherratt, 1984). Unfortunately many highly engineered plasmids based
on ColEI seem to be much less stable in comparison with their high copy number derivatives. Meacock and Cohen (1980) identified a 270 bp region of pSC101 which is responsible for efficient segregation and partitioning of replicated plasmid molecules to daughter cells at division. The par function is not essential for replication, it is cis acting and lacks coding potential, indicating that it is more than likely a site, which accomplishes active partitioning and may be analogous to the centromeres of eukaryotic chromosomes. In addition to stabilizing pSC101 par derivatives, it fully stabilizes an unrelated multicopy plasmid pACYC184 indicating that par is not replicon specific.

Fine deletion mapping analysis has delimited par to a 130 bp region within the conveniently clonable 372 bp EcoR1 - AvaI fragment of pSC101 (Miller et al., 1983; Kalla and Gustafsson, 1984) (Fig. 1.5). Three discrete segments designated a, b and a' within this region are directly involved in the activity of par and therefore have been termed "partition related". Segments a and a' are direct repeats and 14/16 bp are identical, while b is an inverted repeat of both, with 8/9 and 9/12 base matches respectively. The position and dyad symmetry of these segments are such that hairpin-loops could potentially form between the middle segment b and either a or a'.

Deletion analysis revealed that removal of any 2 of the 3 partition related segments resulted in a par- phenotype, while removal of all 3 resulted in a "super-par-" phenotype, plasmids in this latter category being more unstable than the former category. Deletion of one segment produced a Cmp- phenotype although plasmids of this type were still stably maintained. The Cmp phenotype relates to the ability of a
Fig. 1.5

The pSC101 par Fragment

The sequence shows the EcoRI-AvaI fragment, (Miller et al., 1983) with the EcoRI linker addition. The partition related segments a, b and a' shown to be involved in par function are underlined.
The pSC101 par Fragment

The sequence shows the EcoRI-AvaI fragment, (Miller et al., 1983) with the EcoRI linker addition. The partition related segments a, b and a' shown to be involved in par function are underlined.
plasmid to transform a host strain already carrying an incompatible, wild-type pSC101 plasmid. These plasmids have a reduced transformation frequency which greatly exceeds that seen due to normal incompatibility, however, two Cmp⁻ plasmids can stably co-exist and show normal incompatibility properties. The Cmp phenotype was demonstrated to be associated with the replication system of the plasmid by the stable co-existence of a Cmp⁻ pSC101 derivative with a par pACYC184 plasmid (Tucker et al., 1984). It is possible, that the Cmp phenotype represents an intermediate stage in par function. Deletion of one segment may reduce the affinity of par for its binding site. If the replication enzymes are adjacent to this site, par⁺ plasmids may be preferentially replicated, rather than Cmp⁻ plasmids. The stable co-existence of two Cmp⁻ plasmid indicates that both have an equal probability of being replicated, while the stable co-existence of a Cmp⁻ pSC101 plasmid with a par⁺ pACYC184 implies that there may be different replication sites for unrelated plasmids. Furthermore, "Super-par" and par⁻ pSC101 derivatives with copy number equivalents to the wild-type, exhibit segregation frequencies which are several fold lower and slightly lower, respectively, than the calculated copy numbers of these plasmids. Also, temperature-sensitive pSC101 replicons at the non-permissive temperature for replication, are lost from cells at a rate consistent with the calculated copy number when they are par⁺. But when they are "Super-par", half of the cell population loses the plasmid after only one cell doubling, indicating that this latter group segregate as a single unit (Tucker et al., 1984). The previous data implies that pSC101 plasmids possessing an intact par region can be counted as individual molecules and segregated
evenly to daughter cells. Moreover, the Cmp- phenotype suggests that the plasmids containing par regions deleted for one or more partition-related segments a, b or a' are at replicative disadvantage in the presence of par+ plasmids.

It has been suggested that par functions by binding to some host encoded component of the chromosome partitioning system. Evidence for this theory is supported by Gustafsson et al. (1983) who has demonstrated an association between the pSC101 par region and the outer membrane of E. coli. Recent work by Wahle and Kornberg (1988) has shown that the par sequence is a specific binding site for DNA gyrase. The tight binding of DNA gyrase to one particular site (alone or in conjunction with other proteins) might serve in the intracellular organisation of the DNA at cell division.

1.8 Plasmid Copy Number

The simplest way of achieving high expression is by cloning the gene of interest on a high copy number plasmid. Using a high copy plasmid increases the gene dosage providing more mRNA transcripts, thereby overcoming a major rate limiting step in protein synthesis. A range of high copy number vectors have been derived from ColE1, all of them utilising the RNAI/II interaction as the basis of achieving high copy numbers (Twigg and Sherratt, 1980; Boros et al., 1984; Gayle et al., 1986). Potentially more interesting are the temperature inducible runaway replication vectors developed by Uhlin et al. (1979). These vectors, based on the low copy number plasmid R1 (10 - 25 copies per chromosome), have lost the ability to control its copy number at increased growth temperatures. At temperatures above 35°C control of
replication is lost, resulting in an increase in copy number to more
than 2,000 copies per chromosome. ColEl derived vectors have also been
employed in this way. Wright et al. (1986) have constructed a dual
origin vector, comprising an inducible ColEl-derived origin controlled
by the lambda Pr promoter, the c1857 temperature-sensitive repressor
gene and the pSC101 origin with the par locus. Frey and Timmis (1985)
obtained an elevated copy number, by employing ColD based vectors which
auto-amplify their own copy-number when the cultures enter the
stationary phase of growth. All these expression vectors overcome the
problems of metabolic burden and instability associated with
maintaining a high copy number, by maintaining a low copy number during
the growth phase and then inducing a higher copy number with expression
at the end of the growth prior to harvesting. However, few have been
used in high expression work since the energy of the cell is directed
towards DNA rather than protein synthesis.

1.9 Protein Stability

A number of heterologous proteins expressed in E. coli have
been reported as being unstable e.g., somatostatin (Itakura et al.,
1977), human pre-interferon and interferon (Taniguchi et al., 1980).
Not all recombinant proteins are unstable in E. coli. The variation in
half life of 'normal' resident proteins is large; 7% of all proteins
have a half life of less than 15 min, another 20-30% are not broken
down unless under conditions of starvation, while the remainder are not
turned over at all (Nath and Koch, 1971). The structural features
which determines the half life of proteins are not known, therefore it
is not possible to predict which proteins will be unstable.

Instability of heterologous protein in E. coli can be
circumvented or reduced by changing the host strain or the cellular location. *E. coli* mutant strains which have a reduced complement of intracellular proteases have been employed to stabilise heterologous proteins. *E. coli* appears to have two proteolytic systems, one involved in the slow turn over of normal cellular proteins and one which rapidly degrades a labile fraction including abnormal proteins resulting from mutation, errors in translation, denaturation and may include heterologous proteins (Swarms and Goldberg, 1982). An *E. coli* mutant has been isolated which lacks the *lon* protease, a DNA binding protein with ATP dependent proteolytic activity, thought to have a role in the degradation of abnormal proteins (Buell et al., 1985). An alternative approach to the problem of protease activity is to clone the antiprotease gene of phage T₄ into the expression vector. This phage gene product reduces proteolysis. Simon et al. (1983) found that fibroblast interferon, a labile eukaryotic protein cloned in *E. coli*, was stabilized in cells where the T₄ antiprotease gene was expressed.

The classic solution to protein instability has been to produce fusion polypeptides. This involves the protection of the recombinant protein, particularly small polypeptides which are susceptible to proteolysis, by the whole or part of a host protein molecule fused to the N-terminus of the recombinant protein. Itakura et al. (1977) used beta-galactosidase to afford considerable protection for the low molecular weight polypeptide somatostatin, whose synthesis is undetectable without the fusion to beta-galactosidase.

By the addition of a leader sequence to a cloned gene the location of proteins produced within the cell can be directed into the periplasm. One of the major advantages of producing the products in the periplasmic space is the reduction in proteolysis. Swarms and
Goldberg (1982) showed that of the 8 proteases identified the distribution within \textit{E. coli} is 5 in the cytoplasm, 2 in the periplasm and 1 distributed equally between the two. The distribution of these proteases is such that fusion of a signal peptide or leader sequence to a cloned gene product ensuring its secretion to the periplasm increases its stability to proteolytic activity. Talmadge and Gilbert (1982) demonstrated that preproinsulin molecules were 10 times more stable in the periplasm than in the cytoplasm. A range of bacterial, eukaryotic or hybrid signal sequences have been used to direct heterologous protein out of the cytoplasm (Talmadge \textit{et al.}, 1980; Gray \textit{et al.}, 1985; Ohuye \textit{et al.}, 1983; Nagahari \textit{et al.}, 1985; Abrahamsen \textit{et al.}, 1986) with varying degrees of success.

Cytoplasmic localisation of heterologous protein can also induce aggregation of the expressed protein into inclusion bodies. The 'packaging' of insulin into dense cytoplasmic inclusion bodies within \textit{E. coli} was originally observed by Williams \textit{et al.} (1982) and Paul \textit{et al.} (1983), although the accumulation of 'abnormal' proteins in \textit{E. coli} into intracellular granules had been previously described (Prouty and Goldberg, 1972; Prouty \textit{et al.}, 1975). The formation of such aggregates provides a means of stabilisation for foreign proteins over-expressed in \textit{E. coli} was demonstrated with an 'abnormal' beta-galactosidase construct (Cheng \textit{et al.}, 1981; Cheng, 1983). This phenomenon has since been observed in the expression of several eukaryotic proteins in \textit{E. coli} including interleukin - 2 (Devos \textit{et al.}, 1983), prochymosin (Etage \textit{et al.}, 1983; Kawaguchi \textit{et al.}, 1984) gamma-interferon (Simons \textit{et al.}, 1984) and bovine growth hormone (Schoner \textit{et al.}, 1985). The common feature of all the systems is that gene expression and product formation is high (4 to 30\% of soluble protein of the cell) and, in the
case of all eukaryotic proteins studied to date, involves proteins normally secreted from their host eukaryotic cell but 'constructed' to be cytoplasmic products in E. coli. Such normally secreted proteins all have tertiary structure dependent on the formation of disulphide bridges. It appears that the redox system of E. coli is incapable of fully oxidising such proteins since a large percentage of the sulphydryl groups in these proteins are found in the reduced rather than the oxidised form (Marston et al., 1984). Although inclusions probably offer protection against proteases, they do present problems of extraction and purification. In most instances denaturants (SDS or urea) have to be used to extract the protein. For proteins of pharmaceutical interest, the use of detergents is undesirable since it is difficult to remove them completely. The formation of inclusion granules can provide a rapid means of purifying foreign proteins. Such granules have been isolated by a single step centrifugation in the case of prochymosin (Marston et al., 1984). Regardless of the extraction method used, the protein will almost certainly have to be renatured, which may prove difficult, if not impossible.

1.10 Host Physiology

In spite of the many problems of expressing heterologous protein in E. coli, it is still the paradigm against which all other systems are measured. While E. coli will not carry out post-translational modifications, such as glycosylation of eukaryotic proteins, it will express larger quantities of heterologous protein compared to other systems available.

The choice of which E. coli host to be used to express the heterologous protein can have a marked effect on its expression
Less tangible is the physiology of the host cell. The choice of nutrients, the way in which they are supplied to the culture (Cheng, 1983) and environmental parameters such as temperature (Mizukami et al., 1986) and dissolved oxygen tension are all of particular importance to the expression of heterologous protein in *E. coli*.

1.11 Continuous Cultivation

Continuous cultivation was used extensively in this study to examine the segregative stability of plasmids. The advantages of using continuous culture over batch culture to study the kinetics of plasmid loss are numerous. In batch culture the cells alternate between conditions of nutrient excess and nutrient starvation and, as a result, the degree of competition between plasmid containing and plasmid free cells may vary. In such circumstances kinetic analysis of the data is impossible. Once cells reach the stationary phase in batch cultures the death rate of plasmid free and plasmid containing cells would not be the same. Batch cultures would therefore have to undergo repeated subculturing prior to stationary phase. Also the determination of the number of generations which the cells have undergone is not accurate.

Continuous culture does not suffer any of these disadvantages. The steady state environment achieved in continuous culture provides a means of quantitatively studying segregational plasmid stability.

1.11.1 Theory of Continuous Culture

The continuous culture of microorganisms has come a long way since its introduction (Novick and Szilard, 1950a, b; Monod, 1949).
Here, no attempt will be made to cover all the theoretical and applied aspects of this technique, instead the general theory, with reference to its role in the study of segregational stability, will be developed (for review see Tempest, 1969). The simplest and most straightforward theory of continuous culture, based on Monod kinetics as applied to a culture at equilibrium, was utilized in our study of segregation plasmid stability. Any cell in a medium suitable for growth will multiply at rates that are, overall, constant. Therefore, even though there may be a spread of individual generation times, the population density as a whole increases exponentially with time.

If the original population is \( x_0 \) cells, after \( n \) generations the population is \( x \) cells.

Then, \( x = x_0 \cdot 2^n \)

Also, if the mean generation time is \( g \) and the overall population is achieved after time \( t \).

Here, \( n \) represents the number of doublings that have occurred after some time interval \( t \).

Thus \( n = \frac{t}{g} \).

It follows therefore, that the number of organisms present in a culture after \( t \) hours of incubation will be related to the initial population by the equation

\[
\begin{align*}
x &= x_0 \cdot 2^n \\
x &= x_0 \cdot 2^{t/g} \\
x &= 2^{t/g} x_0
\end{align*}
\]

and taking logarithms,

\[
\ln \left( \frac{x}{x_0} \right) = (\ln 2) \cdot \frac{t}{g}.
\]
\( \frac{(\ln x - \ln x_0)}{t} = \frac{0.693}{g} \)

\[ \mu = \frac{1}{x} \cdot \frac{dx}{dt} = \ln 2 = \frac{0.693}{g} \]

\( \mu \) is called the specific growth rate constant. It is the rate of increase in cell numbers per unit of cell numbers.

The steady state environment necessary to study segregational plasmid stability within cells was achieved by the fact that the rate of change in each and every environmental parameter is linked directly to the rate of change of biomass concentration (that is, to the growth rate of the organism). Therefore, all one needs to do is add fresh medium to the culture at a rate sufficient to maintain the culture population density at some prescribed submaximal value, thereby replenishing nutrients that are being consumed and, simultaneously, diluting out end-products that are accumulating at rates exactly sufficient to ensure that the environment no longer varies with time. The culture was run as an open system, employing an overflow tube (or weir) into the culture so that excess culture can flow from the growth vessel at the same rate that the fresh medium was added, thereby maintaining the culture volume constant.

The construction of a fully defined medium whose composition allows one nutrient to be present at a concentration such that in a batch culture, it would become depleted before any other nutrient was fully used up, has been achieved. Thereby the rate of supply of that nutrient, added along with the other essential nutrients in the fresh medium, would prescribe the rate of growth of the organism in the culture.
Within such nutrient-limited cultures it would be the concentration of nutrients in the culture extracellular fluids that is actually limiting the rate of cell synthesis, since the rate at which such a substrate is taken into the cell will be a function of its concentration.

If the uptake process involves some enzyme-catalysed reaction, then one might expect that the relationship between uptake rate (and growth rate) and concentration of limiting substrate would be of the form of a Michaelis-Menton equation.

That is

\[ V = \frac{V_{\text{max}} S}{K_S + S} \]

or

\[ \mu = \frac{\mu_{\text{max}} S}{K_S + S} \]

where

- \( V \) = rate of penetration of substrate,
- \( V_{\text{max}} \) = maximum rate of substrate uptake
- \( S \) = concentration of substrate
- \( K_S \) = saturation constant

- \( \mu \) = specific growth rate
- \( \mu_{\text{max}} \) = maximum growth rate
- \( K_S \) = saturation constant

Monod (1949) found experimentally that the above relationship did apply to many cultures, and though it may not be valid under all circumstances, it forms the cornerstone of much of the theory of
microbial growth in continuous culture, as applied to a chemostat.

The change in concentration of the organism with time will depend on the balance between the growth rate $\mu$ and the dilution rate $D$.

\[
\text{Change} = \text{Growth} - \text{Washout} \\
\frac{dx}{dt} = x - Dx
\]

Therefore at equilibrium when $\frac{dx}{dt} = 0$

\[
\mu = D
\]

steady state conditions are ultimately established. Where the specific growth rate ($\mu$) and the dilution rate ($D$) are equal, with the dilution rate set below a critical value at which the organism express their maximum growth rate.

Increasing the dilution rate will increase the specific growth rate of the organism. A point will be reached at which the growth rate will no longer be able to keep pace with the dilution rate at $\mu_{max}$, this is termed the maximum specific growth rate and the point at which washout occurs.

1.12 Aims of Thesis

The main aim of this thesis was to obtain high expression of the Pseudomonas derived enzyme carboxypeptidase G2 (CPG2) in E. coli and scale up its expression for production.

CPG2 belongs to a class of enzymes which hydrolyse the C-terminal glutamate moiety from folic acid and its analogs (Sherwood et al., 1985). The key role of reduced folates as co-enzymes in many
biochemical pathways has led to the folic acid molecule becoming a prime target in cancer chemotherapy (Bertino et al., 1971). Folate depletion has most commonly been achieved using the folic acid antagonist methotrexate to inhibit dihydrofolate reductase (Blayer, 1978), but may also be achieved directly using the carboxypeptidase G class of enzyme (Kalghatgly and Bertino, 1981). In addition, these enzymes may be used in rescue therapy following high dosage methotrexate regimes (Chabner et al., 1972). More recently an antibody – CPG2 conjugate has been shown to be cytotoxic to a choriocarcinoma cell line in vitro (Searle et al., 1986), and to be selectively targeted to tumour tissue in a choriocarcinoma xenograft in nude mice (Melton et al., 1986). The potential of CPG2 in cancer chemotherapy has brought forward the requirement for large quantities of purified protein for experimental and clinical testing. To alleviate this situation the gene encoding CPG2 (cpg) was cloned in E. coli and its entire nucleotide sequence determined (Fig.1.6) (Minton et al., 1983a; Minton et al., 1984). Expression was obtained in E. coli from a 2.03 kb BamH1 fragment of the cpg gene subcloned into the BamH1 site of pAT153. This construction in the host W5445 directed expression of CPG2 to between 1% and 2.5% of the cells soluble protein into the periplasmic space. The relatively high level of expression, compared to the native promoter (0.2%) was due to transcriptional read through from the promoter of the upstream tet gene. The expression vector pNM21 in the host W5445, had exhibited plasmid segregational instability during cultivation. This was therefore the starting point to improve CPG2 expression. Other factors affecting expression were also investigated; copy number and a temperature inducible promoter system.
Fig. 1.6

Nucleotide Sequence of the cog gene.

Dashed overlining represents an amino terminal sequence determination in an automated spining cup sequencer. Solid overlining represents elastase peptides which have been completely sequenced and fitted to the translation DNA sequence. Solid-dotted overlining represents peptides which have been fitted on the basis of the partial sequencing (solid) and amino acid analysis (dotted). S.D. refers to the putative Shine-Dalgarno sequence (or ribosome-binding site). Possible promoter sequences (i.e., -35 and -10 regions) are indicated by the solid lines above and below the relevant sequence. The signal peptide is indicated by the heavy arrows, as subdivided into its hydrophilic and hydrophobic.
CHAPTER 2

MATERIALS AND METHODS
2.1 MATERIALS

2.1.1 Chemicals

All the materials used in this study were of the highest grade commercially available (Anala R) and unless otherwise stated in the text were obtained from the following sources:

Microbiological media - Oxoid Ltd or Difco Laboratories
Chemicals - BDH Ltd or Sigma (UK) Ltd

2.1.2 Enzymes

All restriction endonuclease, T4 polynucleotide ligase, E. coli DNA polymerase I; large fragment (Klenow), lysozyme (Grade 1), Proteinase K and Bovine serum albumin (BSA) fraction V were purchased from Bethesda Research Laboratories, Gibco Ltd., Paisley, Scotland.

2.1.3 Media

The following media were prepared with double distilled, deionised water and autoclaved at 15 psi (121°C) for 15 min unless otherwise stated. Post sterilisation additions were made aseptically by filtration of solutions through a 0.22 μm filter.

Luria - Bertani (LB) Medium

Used routinely for growth of E. coli strains in both solid and liquid form.
Contains per litre -

Tryptone 10 g
Yeast extract 5 g
NaCl 10 g

pH adjusted to 7.5 with NaOH.

For solid media (L-agar) 1.5% (w/v) purified agar was added.

**Nutrient Broth**

Used routinely for growth of *E. coli* strains.

Contains per litre -

Nutrient broth 16 g
NaCl 5 g

Natural pH of 7.5.

For solid media 1.5% (w/v) purified agar was added.

**YT Broth**

Used to grow *E. coli* strains prior to making competent cells ready for transformation and transfection. (Also used at double strength concentration).

Contains per litre -

Tryptone 8 g
Yeast extract 5 g
NaCl 5 g

For solid media, 1.5% (w/v) purified agar was added.

**M-9 Medium**

Minimal media used for the growth of *E. coli* strains.
**M-9 Buffer (x 10)**

Contains per litre -

- $\text{Na}_2\text{HPO}_4$ 60 g
- $\text{KH}_2\text{PO}_4$ 30 g
- NaCl 5 g
- $\text{NH}_4\text{Cl}$ 10 g

**Mg/Ca solution (x 100)**

- 0.01 M CaCl$_2$
- 0.1 M MgCl$_2$

**M-9 media contains per 100 ml**

- M-9 Buffer 10 ml
- Mg/Ca solution 1 ml
- 20% (w/v) glucose 1 ml

M-9 buffer was made up to 100 ml with H$_2$O and its pH adjusted to 7.4 prior to being autoclaved. Mg/Ca solution, glucose and amino acid requirements were sterilised separately by filtration before being added to the sterile M-9 buffer.

**Iso-sensitest Agar**

Used for the identification of beta-lactamase producing colonies.

Contains per litre -

- Iso-sensitest agar 31.4 g

made up to 1 litre with H$_2$O.

**H-top**

Used to an overlay for plating out bacteriophage. For convenience was
made up in 200 ml volumes.

Contains per litre -

- Bacto Tryptone 10 g
- Difco minimal agar 8 g
- NaCl 8 g

Fermentation Media

Semi-defined media was used for scale-up of fermentations while fully defined media was required for continuous cultivation. The bulk of the medium was prepared with demineralised water and autoclaved at 20 psi (125°C) for 30 min. Heat labile component(s) were either sterilised separately and then added to the bulk sterile medium or filter sterilised and then added.

Medium A

Contains per litre -

- K$_2$HPO$_4$ 4.0 g
- KH$_2$PO$_4$ 1.0 g
- NH$_4$Cl 1.0 g
- CaCl$_2$·6H$_2$O 0.01 g
- K$_2$SO$_4$ 2.6 g
- Casamino acids 20.0 g
- Yeast extract 3.0 g
- Glycerol 46.0 g
- 1M MgCl$_2$ 1 ml
- Trace elements 10 ml
Trace elements stock solution:

Contains per litre -

- EDTA.Na$_2$ 5.0 g
- FeCl$_3$.6H$_2$O 0.5 g
- ZnO 0.05 g
- CuCl$_2$.2H$_2$O 0.01 g
- Co(NO$_3$)$_3$.6H$_2$O 0.01 g
- NH$_4$Mo$_7$O$_2$$_4$ 0.01 g

Trace elements stock solution were made up by dissolving 5 g EDTA (disodium salt) in approximately 800 ml H$_2$O and then separately adding the salts while maintaining the pH between 7 and 8. The trace elements and MgCl$_2$ can either be autoclaved separately and then aseptically added to the bulk medium or filter sterilised along with any amino acids, vitamins and antibiotics required.

Continuous Culture Media

Developed by Tempest et al. (1969) for the continuous cultivation of E. coli.

Stock solutions for simple salts media:

<table>
<thead>
<tr>
<th>Solution</th>
<th>Concentration</th>
<th>Concentration g l$^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phosphorus source</td>
<td>2M Na$_2$PO$_4$.2H$_2$O</td>
<td>312</td>
</tr>
<tr>
<td>Nitrogen source</td>
<td>4M NH$_4$Cl</td>
<td>214</td>
</tr>
<tr>
<td>Source</td>
<td>Concentration</td>
<td>Value</td>
</tr>
<tr>
<td>----------------------------</td>
<td>---------------</td>
<td>-------</td>
</tr>
<tr>
<td>Potassium source</td>
<td>2M KCl</td>
<td>149</td>
</tr>
<tr>
<td>Sulphur source</td>
<td>1M Na₂SO₄.10H₂O</td>
<td>322</td>
</tr>
<tr>
<td>Chelating source</td>
<td>1M Citric acid</td>
<td>210</td>
</tr>
<tr>
<td>Magnesium source</td>
<td>0.25M MgCl₂</td>
<td>23.8</td>
</tr>
<tr>
<td>Calcium source</td>
<td>0.02M CaCl₂</td>
<td>2.2</td>
</tr>
<tr>
<td>Molybdenum source</td>
<td>0.001M Na₂MoO₄</td>
<td>0.2</td>
</tr>
</tbody>
</table>

**Trace elements:**

Contains per litre:

- FeCl₃.6H₂O 5.4 g
- MnCl₂.4H₂O 2.0 g
- CuCl₂.2H₂O 0.17 g
- CaCl₂.6H₂O 0.48 g
- H₃BO₄ 0.06 g

**Composition of simple salts media used in nutrient limitation studies:**

<table>
<thead>
<tr>
<th>Solution</th>
<th>Concentration of limiting solution ml⁻¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phosphate</td>
<td>C: 5.0, P: 1.0</td>
</tr>
<tr>
<td>Nitrogen</td>
<td>C: 25.0, P: 25.0</td>
</tr>
<tr>
<td>Potassium</td>
<td>C: 5.0, P: 5.0</td>
</tr>
<tr>
<td>Sulphate</td>
<td>C: 2.0, P: 2.0</td>
</tr>
<tr>
<td>Chelate</td>
<td>C: 2.0, P: 2.0</td>
</tr>
<tr>
<td>Magnesium</td>
<td>C: 5.0, P: 5.0</td>
</tr>
<tr>
<td>Calcium</td>
<td>C: 1.0, P: 1.0</td>
</tr>
<tr>
<td>Trace elements</td>
<td>C: 5.0, P: 5.0</td>
</tr>
<tr>
<td>Molybdate</td>
<td>C: 0.1, P: 0.1</td>
</tr>
<tr>
<td>Glucose or other</td>
<td>C: 10, P: 30</td>
</tr>
</tbody>
</table>

**Carbon source (g):**

43
2.1.4 Buffers

Distilled and deionised water was used in the preparation of all buffers. Most of the buffers were prepared as stock solutions in concentrated form.

**T.E. (x 100)**
Stock solution contains per litre:
- Tris 121 g
- EDTA.Na$_2$ 3.72 g
Natural pH of 8.0 – 8.5.

**TM (x10)**
- Tris-HCl 100 mM (pH 8.5)
- MgCl$_2$ 50 mM

**T.A.E. (x 20)**
Stock solution contains per litre:
- Tris 96.80 g
- Sodium acetate 32.81 g
- EDTA.Na$_2$ 7.44 g
pH adjusted to 8.0 with glacial acetic acid.

**T.B.E. (x 10)**
Stock solution contains per litre:
- Tris 108 g
- Boric acid 55 g
- EDTA.Na$_2$ 9.3 g
Natural pH of 8.3.
Glycine Buffer (x 10)

Stock solution contains per litre:

- Glycine: 150 g
- NaOH: 6 g
- EDTA.Na₂: 7.44 g

Natural pH of 9.0.

STET Buffer

- Sucrose: 8% (v/v)
- Triton X-100: 5% (v/v)
- EDTA.Na₂: 50 mM
- Tris-HCl: 50 mM (pH 8.0)

Ligation Buffer (x 10)

- NaCl: 300 mM
- Tris-HCl: 300 mM (pH 7.5)
- MgCl₂: 75 mM
- Spermidine: 10 mM
- DTT: 10 mM
- EDTA.Na₂: 2 mM

Tracking Dye (x 6)

- Glycerol: 50% (v/v)
- Xylene cyanol: 0.02% (w/v)
- Bromophenol Blue: 0.02% (w/v)
- Sucrose: 40% (w/v)
S.S.C. (x 20)
Stock solution contains per litre:

- NaCl 175.3 g
- Na Citrate 88.2 g

Made up to 800 ml originally, pH adjusted to 7.4 with NaOH, then adjust volume to 1 litre.

**Kinase Buffer (x 10)**

- Tris-HCl 500 mM (pH 8.0)
- MgCl$_2$ 100 mM

**Pre-Hybridising Solution**

- 6 x SSC
- 10 x Denhardt's solution
- 0.1% (v/v) SDS
- 100 g ml$^{-1}$ denatured, Salmon sperm DNA
- 50% (v/v) Formamide

**Oligo-Labeling Buffer**

Made up from the following components:

**Solution O**

- Tris-HCl 1.25 M (pH 8.0)
- MgCl$_2$ 0.125 M

Stored at 4°C.

**Solution A**

- Solution O 1 ml
- 2-mercaptoethanol 18 µl
0.1 M dCTP  5 µl  
0.1 M dTTP  5 µl  
0.1 M dGTP  5 µl  

Stored at -20°C.

Solution B

2 M Hepes, titrated to pH 6.6 with 4 M NaOH stored at 4°C.

Solution C

Hexadeoxyribonucleotides, evenly suspended (this does not completely dissolve) in TE at 90 OD units ml⁻¹. Stored at -20°C.

Oligo-labelling buffer consists of solutions A:B:C mixed in the ratio of 100:250:150. Stored at -20°C.

Denhardt's Solution (x 50)

Stock solution contains per litre:

- Ficoll  10 g
- Polyvinylpyrrolidone  10 g
- B.S.A. (Fraction V)  10 g

**CAT Assay Buffer**

1 ml of Tris-HCl (100 mM, pH 7.8) containing 4 mg of DTNB was added to 0.2 ml of Acetyl Co-enzyme A (5 mM) and made up to 10 ml with 8.8 ml of H₂O.

**CPG Assay Buffer**

Consists of Tris-HCl (100 mM, pH 7.3) with 0.2 mM ZnSO₄. The substrate for the reaction being 60 nmol methotrexate (3.4 mg Methotrexate in 20 ml CPG assay buffer).
Restriction Endonuclease Buffers

For convenience, restriction enzymes were divided into 3 groups, in which optimum performances took place in buffer with high ionic strength, medium ionic strength or low ionic strength. In this way only 3 buffers were required for the vast majority of enzymes which can be categorised into these 3 groups. There were still some enzymes which did not fall into any of the 3 groups, a separate buffer was therefore made up, obtained from the product profile provided by the manufacturer.

Low Salt x 10

<table>
<thead>
<tr>
<th>Buffer Component</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tris-HCl</td>
<td>100 mM (pH 7.5)</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>100 mM</td>
</tr>
<tr>
<td>Dithiothreitol (DTT)</td>
<td>10 mM</td>
</tr>
</tbody>
</table>

Medium Salt x 10

<table>
<thead>
<tr>
<th>Buffer Component</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>NaCl</td>
<td>500 mM</td>
</tr>
<tr>
<td>Tris-HCl</td>
<td>100 mM (pH 7.5)</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>100 mM</td>
</tr>
<tr>
<td>Dithiothreitol</td>
<td>10 mM</td>
</tr>
</tbody>
</table>

High Salt x 10

<table>
<thead>
<tr>
<th>Buffer Component</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>NaCl</td>
<td>1 mM</td>
</tr>
<tr>
<td>Tris-HCl</td>
<td>500 mM (pH 7.5)</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>100 mM</td>
</tr>
<tr>
<td>Dithiothreitol</td>
<td>10 mM</td>
</tr>
<tr>
<td><strong>Acrylamide Solution</strong></td>
<td><strong>Acrylamide</strong></td>
</tr>
<tr>
<td>-------------------------</td>
<td>---------------</td>
</tr>
<tr>
<td></td>
<td><strong>Bis-acrylamide</strong></td>
</tr>
</tbody>
</table>

| **2x Running Gel Buffer** | **Tris-HCl (pH 8.8)** | 0.75M |
| (lower Tris) | **SDS** | 0.2% (w/v) |

| **2x Stacking Gel Buffer** | **Tris-HCl (pH 6.8)** | 0.25M |
| (upper Tris) | **SDS** | 0.2% (w/v) |

| **Sample Buffer** | **Tris-HCl (pH 6.8)** | 12.5 ml |
|                  | **Glycerol** | 20 ml |
|                  | **beta-mercaptoethanol** | 10 ml |
|                  | **10% (w/v) SDS** | 40 ml |
|                  | **Bromophenol Blue** | 1.5 mg |

*made up to 100 ml with H₂O*

| **SDS-PAGE Buffer** | **Tris** | 3.03 g |
|                     | **Glycine** | 14.4 g |
|                     | **SDS** | 1.0 g |

*Natural pH of 8.3*

<p>| <strong>Staining Solution</strong> | <strong>Coomassie Blue R250</strong> | 1.25 g |
|                       | <strong>50% Methanol</strong> | 454 ml |
|                       | <strong>Glacial Acetic Acid</strong> | 46 ml |</p>
<table>
<thead>
<tr>
<th>Destaining Solution</th>
<th>Acetic Acid</th>
<th>75 ml</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Methanol</td>
<td>50 ml</td>
</tr>
<tr>
<td></td>
<td>H₂O</td>
<td>875 ml</td>
</tr>
</tbody>
</table>
### 2.1.5 Bacterial Strains and Plasmids

#### Table 2.1

**Bacterial Strains used in this Study**

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Source or Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>W5445</td>
<td><em>pro</em>, <em>leu</em>, <em>thi</em>, <em>thr</em>, <em>supE44</em>, <em>lacY</em>, <em>tonA</em>, <em>hadM</em>, <em>hadR</em>, <em>atR</em></td>
<td>N Mann</td>
</tr>
<tr>
<td>JM83</td>
<td><em>ara</em>, <em>A lac</em>, <em>pro</em>, <em>strA</em>, <em>thi</em>, <em>F8OD</em>, <em>lacZ M15</em></td>
<td>Messing <em>et al.</em> (1981)</td>
</tr>
<tr>
<td>JM101</td>
<td><em>supE</em>, <em>thi</em>, <em>A lac</em>, <em>pro</em>, <em>F</em>, <em>traD36</em>, <em>proAB</em>, <em>lac T9 ZM15</em></td>
<td></td>
</tr>
<tr>
<td>K12Miatrp</td>
<td><em>SmR</em>, <em>lac Z'me</em>, <em>bio-uVR</em>, <em>A trpEA2</em></td>
<td>Bernard <em>et al.</em> (1979)</td>
</tr>
<tr>
<td></td>
<td>(Nam 7-Naa53 c(857H1))</td>
<td></td>
</tr>
<tr>
<td>GM161</td>
<td><em>thr-1</em>, <em>leuB</em>, <em>dam-4</em>, <em>thi-1</em>, <em>hadS1</em>, <em>lacY1</em>, <em>tonA2I</em>, <em>L</em>-* supE44*</td>
<td>Marinus <em>et al.</em> (1983)</td>
</tr>
<tr>
<td>GM271</td>
<td><em>leuB6</em>, <em>dcm-6</em>, <em>hiaG4</em>, <em>thi-1</em>, <em>hadR2</em>, <em>ara-14</em>, <em>lacY1</em>, <em>salK2</em>, <em>salT22</em>, <em>xy1-5</em>, <em>mtl-1</em>, <em>rpsl136</em>, <em>tonA31</em>, <em>tax-78</em>, <em>L</em>-* supE44*</td>
<td></td>
</tr>
<tr>
<td>GM2199</td>
<td><em>thi-1</em>, <em>ara-14</em>, <em>leuB6</em>, <em>tonA31</em>, <em>lacY1</em>, <em>tax-78</em>, <em>supE44</em>, <em>salK2</em>, <em>salT22</em>, <em>L</em>-* hiaG4*, <em>rpsl136</em>, <em>dam-13</em>: <em>Tn9</em>, <em>xy1-5</em>, <em>mtl-1</em>, <em>thi-1</em></td>
<td></td>
</tr>
<tr>
<td>GM2163</td>
<td><em>ara-14</em>, <em>leuB6</em>, <em>tonA13</em>, <em>lacY1</em>, <em>tax-78</em>, <em>supE44</em>, <em>salK2</em>, <em>L</em>-* dcm-6*, <em>hiaG4</em>, <em>rpsl136</em>, <em>dam-13</em>: <em>Tn9</em>, <em>xy1-5</em>, <em>mtl-1</em>, <em>thi-1</em>, <em>hadR2</em></td>
<td></td>
</tr>
<tr>
<td>JC2926</td>
<td><em>thr-1</em>, <em>leu-6</em>, <em>proA2</em>, <em>his-4</em>, <em>asaE3</em>, <em>recA-3</em>, <em>D J Sherratt</em></td>
<td>N Mann</td>
</tr>
</tbody>
</table>

---

51
Table 2.2
Plasmids used in this Study

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description</th>
<th>Source or Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pNM21</td>
<td>Ap^R, Tc^S, cpg</td>
<td>Minton et al. (1983a)</td>
</tr>
<tr>
<td>pAT153 par</td>
<td>Ap^R, Tc^R, par</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>pUC8 par</td>
<td>Ap^R, lacZ', par</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>pNM21 par</td>
<td>Ap^R, Tc^S, cpg, par</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>pMTL</td>
<td>Ap^R, lacZ' nic/bom'</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>pMTL par</td>
<td>Ap^R, lacZ', nic/bom', par</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>pPIc28</td>
<td>Ap^R</td>
<td>Remaut et al. (1981)</td>
</tr>
<tr>
<td>pNM232</td>
<td>Ap^R, cpg</td>
<td>N Minton</td>
</tr>
<tr>
<td>pNM501</td>
<td>Ap^R, lacZ'</td>
<td>N Minton</td>
</tr>
<tr>
<td>pNM502</td>
<td>Ap^R, lacZ', cpg</td>
<td>Chapter 5</td>
</tr>
<tr>
<td>pNM503</td>
<td>Ap^R, lacZ', cat</td>
<td>Chapter 5</td>
</tr>
<tr>
<td>pCI857</td>
<td>Km^R, cI</td>
<td>N Minton</td>
</tr>
<tr>
<td>pCM10</td>
<td>Ap^R, lacZ', par, cpg</td>
<td>Chapter 5</td>
</tr>
<tr>
<td>pLVC9</td>
<td>Cm^R, Km^S</td>
<td>Warren and Clark</td>
</tr>
</tbody>
</table>
2.2 METHODS

2.2.1 Storage and Growth of Bacterial Strains

Dorset egg agar slopes (Oxoid Ltd., England) provided a convenient laboratory storage method for bacterial strains. Bacterial strains maintained at -70°C in a glycerol containing medium provided a long term method of storage. Before storage at -70°C the strains were grown overnight in L-Broth at the appropriate temperature. Glycerol was then added at 10% (v/v) final concentration, 1 ml aliquots were then placed in sterile vials, mixed and then frozen at -70°C. Prior to storage, and when regenerated from storage plasmid encoded markers and auxotrophic markers were checked by plating onto the appropriate media.

2.2.2 Beta-lactamase Plate Assay

This assay was developed by Boyko and Ganschow (1960), based upon the removal of starch iodine complex by penicilloic acid which forms when penicillin is cleaved by beta-lactamase.

A 20 ml solution of Benzyl penicillin (0.03 g 1⁻¹) stored at 4°C was mixed with 3 ml of iodine solution (203 g KI and 10 g I₂ per litre, stored in the dark) and used to flood the test plates. Isosensitext agar was used in this assay, because it contains soluble starch which provides the blue-black colouration on reaction with iodine. After the solution uniformly stained the plate a bluish-black colour (10 - 15 sec), the mixture was discarded and the plates drained. All the colonies were initially dark yellow brown colour, however plasmid containing colonies soon began to turn white on the stained agar surrounding those colonies begins to clear, until a 'halo' around the colony was formed. This halo increased in size with time.
contrast, those colonies which had lost their plasmid remained brown in
colour, where no localised clearing occurred. After 15 min, there was
a general clearing of media. To obtain optimum results with this
technique it was advisable to have less than 200 colonies per plate,
otherwise clearing around one colony affected adjacent colonies.

2.2.3 Carboxypeptidase Plate Assay

Colonies containing carboxypeptidase gene produce concentric
yellow halos of precipitated pteroic acid, the insoluble product of
folate hydrolysis when grown on folate plates.

Folate plates are made up from M9 media containing 0.1% (v/v)
folate solution. The folate solution was made up from 100 g folate
dissolved in 1 litre H2O, its pH was then adjusted to 7.5 with NaOH to
achieve a clear orange solution which was then stored in a light proof
bottle at 4°C.

2.2.4 Screening for Recombinants using X-Gal Plates

M13 viral DNA and the pUC vectors contain within their linker
region, the so-called alpha peptide. This is a portion of the lac gene
of E. coli, including the operator, promoter and the N-terminal of
beta-galactosidase which is functional when in the E. coli host JM101;
Which carries its own defective beta-galactosidase on its partially
deleted chromosomal loci. The two defective beta-galactosidase
polypeptides complement each other and thus produce a functional beta-
galactosidase. This hydrolyses the colourless X-Gal compound and
results in blue colouration of plaques containing M13 and colonies
containing pUC vectors without cloned inserts. Consequently this
provided a convenient screening method for vectors possessing cloned
inserts in the polylinker region because these plaques and colonies will be white when plated out on X-Gal plates.

2.2.5 Transformation of E. coli Competent Cells

This is the procedure used to introduce plasmid DNA into E. coli cells and is a modified version of that described by Cohen et al. (1972). A fresh stationary phase overnight culture was used to inoculate 50 ml of L-broth, and grown in a 500 ml conical flask at 37°C to an OD450 nm of 0.6. The flask was then placed on ice for 20 min, the culture was then centrifuged (1,000 g for 10 min at 4°C). The pellet was resuspended in 25 ml 0.1 M MgCl₂ and repelleted. The cells were then resuspended in 2.5 ml 0.1M CaCl₂ and held on ice for at least 2 hr. The cells were then competent for a period up to 24 hr, after this time the efficiency of transformation decreases rapidly back to its original level.

100 μl of the competent cells were added to the transforming DNA (10 μl or less) and then left on ice for 30 min. The cells were then heat shocked at 42°C for 2 min, and placed on ice for a further 30 min. During this period the DNA was taken up by the cells. The transformed cells were then allowed to come to room temperature. The competent cells were then diluted serially, plated out onto a selective media and incubated overnight at 37°C.

2.2.6 Small Scale Plasmid Preparations

This method derived from Holmes and Quigley (1981) is used routinely for the rapid isolation of plasmid DNA from small cultures and colonies from plates. The yield of DNA from this method is of sufficient quantity and quality for 3 restriction digests (0.75 ng); 1
ml of a 5 ml overnight culture or 1 ml from a 1 ml 5 hr culture, can provide enough cells to give sufficient yield of DNA. The cells were pelleted in a 1 ml Eppendorf tube, the medium removed and the pellet resuspended in 100 µl STET buffer. 8 µl of freshly prepared lysozyme (10 mg ml⁻¹) was added and left on ice for 5 min. The Eppendorf tube was then placed in a boiling water bath for 40 sec. The suspension was then centrifuged for 10 min in a bench top microcentrifuge. This separated the cell debris as a spongy pellet from the supernatant which contains the plasmid DNA, the pellet was removed with a toothpick. To the supernatant an equal amount of propan 2-ol (75 µl) was added, mixed, then kept at -20°C for 60 min. The suspension was then centrifuged and the ethanol removed. Any excess solvent remaining was removed by desiccation under vacuum. The pellet produced was resuspended in TE buffer and the extrachromosomal DNA was ready for analysis.

2.2.7 Large Scale Plasmid Preparations

The purification of plasmid DNA from E. coli strains utilized a modified method from that described by Clewell and Helinski (1969).

A single colony from a fresh L-agar plate was used to inoculate 200 ml of L-broth in a 500 ml conical flask. The cells were grown at 37°C with minimal agitation until OD₆₅₀ nm reached 0.9. Chloramphenicol (170 g ml⁻¹ final concentration) was then added to the culture and left growing overnight. The cells were harvested by centrifugation (15,000 g for 10 min at 4°C) the pellet was then resuspended in 5 ml 25% (w/v) sucrose; 50 mM Tris-HCl (pH 8.0), then 1 ml of lysozyme (10 mg ml⁻¹ in 25 mM Tris-HCl; pH 8.0) was added and left on ice for 2 min. 2 ml of 250 mM EDTA was added and mixed on ice
for 2 min. 8 ml of Brij/DOC solution (1%, w/v, Brij 58; 0.4%, w/v, sodium deoxycholate in 50 mM Tris- HCl; 25 mM EDTA, pH 8.0) was added, mixed and allowed to stand for 5 min. The lysed cells were then transferred to a 30 ml polypropylene centrifuge tube and centrifuged (45,000 g, 30 min at 4°C). The pellet produced contains the cell debris and the majority of the chromosomal DNA. The supernatant containing the plasmid DNA, was further purified on a CsCl gradient.

2.2.8 Isolation of Plasmid DNA by Centrifugation on a CsCl Gradient

Purification of plasmid DNA was carried out by centrifugation to equilibrium in a CsCl-EtBr gradient. For every 1 ml of DNA solution, 1 gm of CsCl and 0.1 ml EtBr solution (10 mg ml⁻¹) was added. The CsCl solution is transferred to 10 ml Beckman Quick-Seal Centrifuge tubes. The centrifugation was carried out at 15°C, 48K (200,000 g) for 48 hr. After centrifugation the plasmid band was visualised with a U.V. transilluminator (TM 36, Ultra-violet Products Inc., California, USA) and removed using a hypodermic needle inserted into the side of the tube and withdrawn with a 2 ml syringe. The ethidium bromide was removed by extracting three times with isoamyl alcohol saturated with CsCl. CsCl was removed by dialysis against three changes of TE buffer at 4°C. The TE buffer was changed after 2, 4 and 12 hr of dialysis.

2.2.9 Plasmid Purification

Estimation of the concentration and purity was carried out by measuring the optical density at 260 nm and 280 nm. An optical density reading at 260 nm allows estimation of the concentration of DNA, an OD of 1 corresponds to a DNA concentration of 50 μg ml⁻¹. The ratio of
the optical density readings at 260 nm and 280 nm provides an estimate of the purity of the DNA. Protein contamination is considered negligible where the ratio is 1.8 or greater, protein contamination will tend to reduce this ratio. Where the protein contamination was too high or the DNA concentration was low the sample was subjected to phenol extraction to remove proteins and ethanol precipitation to concentrate the DNA.

2.2.9.1 Phenol Extraction of DNA

Ultra-pure phenol (BRL) was melted at 55°C and neutralised by equilibrating (99:1) with TE buffer. Phenol extraction was carried out by the addition of equal amounts of the equilibrated phenol to the sample, mixing, allowing to stand for 2 min, then mixing again followed by centrifugation in a microfuge for 1 min. The aqueous layer was then removed and phenol extractions repeated if necessary. Extractions with di-ethyl ether removed traces of phenol in the sample. Ether was removed from the sample by incubation at 50°C for 30 min or by passing nitrogen over the sample.

2.2.9.2 Precipitation of DNA

DNA was precipitated by the addition of 2 volumes ethanol or 1 volume propan-2-ol, in the presence of 0.1 volume 3 M sodium acetate (pH 5.5). Storage at low temperature allowed the DNA precipitate to form, usually -70°C for 20 to 30 min or -20°C for 1 to 2 hr (overnight). Centrifugation for 10 min in a microfuge at 4°C pelleted the DNA, the supernatant was decanted, the pellet dried and resuspended in TE buffer.
2.2.10 Treatment of DNA with Enzymes

2.2.10.1 Restriction Endonuclease

1 to 5 units of restriction enzyme were used to digest 1 ng of DNA in the appropriate digestion buffer and at the recommended temperature. After 1 to 2 hr incubation, the digestion was stopped by heat inactivation by maintaining the mix at 70°C for 10 min or by the addition of 0.5 M EDTA (pH 7.5).

2.2.10.2 T₄ Polynucleotide Ligase

T₄ polynucleotide ligase catalyses the formation of phosphodiester bonds between complementary 3' hydroxyl and 5' phosphate termini ("sticky ends"). It is also possible to join blunt and double stranded DNA using T₄ DNA ligase under high DNA concentrations.

Digestion of the vector DNA and fragment DNA to provide compatible cohesive termini precedes ligation. The concentration of vector to fragment DNA used for ligation is approximately 1 to 3 for sticky ends and 1 to 10 molar excess for blunt end clonings. Sticky end ligations were incubated for 3 to 12 hr at 14°C, while blunt end ligations required 24 hr incubation.

A typical ligation mix of 10 µl total volume contained 1 µl 10 x ligation buffer, 1 µl 10 mM ATP, 1 µl ligase (10 u µl⁻¹) vector and fragment DNA made up to the final volume with water. Following ligation the mixture was ready for transformation with competent cells.

2.2.10.3 DNA Polymerase, Large Fragment (Klenow)

The 5' – 3' polymerase activity of Klenow was employed for
dideoxynucleotide sequencing and oligo labelling. The sequencing protocol is more fully described in later sections. The oligo-labelling protocol was that described by Feinberg and Vogelstein (1983) with slight modifications. The reaction was carried out at room temperature by the addition of the following reagents in the stated order: (i) H₂O to a total volume of 50 µl; (ii) 10 µl oligo labelling buffer; (iii) 2 µl of 10 mg ml⁻¹ BSA; (iv) DNA; (v) 5 µl of 32p ATP; (vi) 2 u of Klenow.

The polymerisation reaction reaches a plateau within 2 to 8 hr, but was normally left incubating overnight. Prior to hybridisation separation of labelled DNA from unincorporated d NTP's was carried out by desalting on a Pharmacia PD-10 (G-25) disposable column.

2.2.10.4 Alkaline Phosphatase

The terminal 5' phosphate groups on linear vector molecules can be removed by treatment with alkaline phosphatase, preventing self-ligation of the linear vector molecules, thus reducing the efficiency of cloning. This is due to the poor ability of linear vector DNA to undergo in vivo circularisation and therefore resulting in very low frequency of transformation of vector molecules alone. The removal of the terminal 5' phosphates of the vector DNA, leaves the 3' hydroxyl ends available for joining to the 5' phosphate terminals of fragment DNA to be cloned. In this way self ligation is minimised since T₄ DNA ligase requires 5' terminal phosphate on one of the precursor molecules.

Calf intestinal phosphatase (CIP) used at 1 u µl⁻¹ and incubated at 37°C for 90 min. After this period the reaction can be terminated by heat denaturation (70°C for 10 min) or removal by phenol
2.2.11 Agarose Gel Electrophoresis

Electrophoresis through agarose gels was the method used to routinely separate, identify and purify DNA fragments. The rate of DNA migration through the gel is dependent on the concentration of agarose. It was therefore possible using gels of different agarose concentration to resolve a range of different sized DNA fragments.

The agarose was dissolved by heating in the appropriate buffer at the required concentration for efficient fragment separation. Following cooling the agarose was poured into the cast (either horizontal or vertical) allowed to cool until set, and was then ready to be set up for electrophoresis.

2.2.11.1 Vertical Gels

Vertical gels are used for running large volumes of digested DNA for fragment isolation. The gels were run overnight at 15 mA or at 50 mA for at least 3 hr in TBE electrophoresis buffer. After the gel has been run it was stained in EtBr (0.5 μg ml⁻¹) for 30 min prior to being photographed. The EtBr stained gels were photographed on a UV transilluminator (Transilluminator model TM36) using a Polaroid MP-4 camera, 667 film and fitted with a Kodak Wratten Filter No. 9.

2.2.11.2 Mini Gels

A horizontal gel system was used for the rapid analysis of digest and for fragment isolation. Gels were run at 100 V 60 mA in TBE electrophoresis buffer for approximately 60 min. Usually the EtBr (0.5 μg ml⁻¹) was incorporated into both the gel and the running buffer prior
to electrophoresis. This enabled the gel to be examined directly on the UV transilluminator at the end of the run.

2.2.11.3 Horizontal Gels

Were used for the analysis of a large number of digests and can be run in a double decker fashion to double the number of samples run on one gel. The gels were run overnight submerged at 100 V or using a wick at 60 V using TBE electrophoresis buffer. Glycine horizontal wick gels were used for more accurate analysis of digests. The electrophoresis buffer used in this system was glycine buffer. The gels were run at 240 V for 2 to 3 hr. Staining and photographing the gels are the same as for vertical gels.

2.2.11.4 Sizing DNA Fragments

Molecules of linear double stranded DNA travel through agarose gels at rates which are inversely proportional to the log 10 of the molecular weight. Therefore the distance migrated by standards of known molecular weights can be used to calculate the molecular weights of unknown fragments in the gel run alongside the standards. The DNA Star Inc. computer program DIGIGEL was used to measure restriction mapping gels. The migrations of standard fragments were used to plot a mathematical approximation to the migration curve. Unknown fragment migrations were measured and the fragment sizes were interpolated from the migration curve.

2.2.11.5 Tracking Dye

Prior to loading on a gel, samples were mixed with a high density solution (tracking dye) to aid loading DNA solution into wells.
The tracking dyes were also used to indicate the migration front of the electrophoresis buffer and the position of fast moving fragments in the gel.

2.2.12 Recovery of DNA from Agarose Gels

After the agarose gel had been run and stained, the DNA band of interest was located using a long wave transilluminator (to minimize UV damage to the DNA). The agarose containing the band was excised and placed in a dialysis tube filled with 0.5 x TBE, completely covering the gel. The sealed tube was then placed in an electroelution tray parallel to the electrodes in 0.5 x TBE. The electroelution conditions were 200 V for 2 hr, during this period the DNA was electroeluted out of the gel and onto the inner wall of the dialysis tube. Removal of the DNA from the wall of the dialysis tubing into the electroelution buffer was done by diffusion, i.e. leaving the tubing in the buffer without a potential difference across it or reversal of the polarity of the current 3 times for 5 min. The tube was then opened and the buffer surrounding the gel slice removed. The DNA in the buffer was then extracted with isoamyl alcohol removing the ethidium bromide present. Further purification and concentration follows before DNA can be used (2.2.9).

2.2.13 Quantitation of DNA

2.2.13.1 Spectrophotometric Determination of DNA

The optical density at 260 nm was used to estimate the concentration of DNA. An optical density of 1 corresponds to approximately 50 µg ml⁻¹ for double stranded DNA, 40 µg ml⁻¹ for single
stranded DNA, and 20 μg ml⁻¹ for oligonucleotides. Accurate quantitation of DNA is only possible where the sample is relatively pure.

2.2.13.2 Diphenylamine Assay

This was used to estimate the concentration of low levels of DNA. The procedure was as described by Burton (1956) and subsequently simplified by Giles and Myers (1965). DNA concentration was estimated by comparison against a standard curve obtained from calf thymus DNA. Standard DNA stock solution, 6 mg ml⁻¹ was degassed and stored at 4°C. Diphenylamine reagent (0.4 g diphenylamine dissolved in 40 ml glacial acetic acid with 1.1 ml concentrated sulphuric acid), was used on day of preparation. The stock solution of DNA was diluted with TE buffer to cover the required range, usually between 1 μg - 100 μg ml⁻¹.

2 ml of diphenylamine reagent was then added to the 1 ml DNA solution and boiled for 10 min in a fume cupboard. The solutions were then allowed to cool, mixed and then the optical density at 600 nm was recorded. The DNA concentrations were estimated by reference to the standard curve (Fig. 2.1).

2.2.14 Sequencing with Chain Terminating Dideoxynucleotides

The method used is essentially that described by Sanger et al. (1977) with alterations described in personal communication with N Minton.

2.2.14.1 M13 Sequencing Vectors

Vectors used are a series of M13 mp bacteriophages developed
Fig. 2.1

Standard Curve used for the Determination of DNA Concentrations

![Graph showing DNA concentration vs. OD at 600nm](image-url)
by Messing (1983) specifically for DNA sequencing with universal M13 primers.

2.2.14.2 Large Scale Preparation of Bacteriophage

10 ml of the supernatant of a plaque infected JM101 culture was used to inoculate a 1 l culture of JM101 in L-broth when at an optical density at 600 nm of 0.6. The culture was harvested 4 hr later and a conventional large scale plasmid prep was carried out to isolate the replicating form (RF) of the viral DNA.

2.2.14.3 Transfection

The method used to introduce the M13 vector into E. coli cells was identical to that described by Cohen et al. (1972) for transformation. The transformed cells were mixed with exponential and spread onto agar with H-top to obtain a lawn. Vectors containing inserts could be identified using a beta-galactosidase complementation assay using X-Gal and IPTG in the overlay.

2.2.14.4 Preparation of Templates for Sequencing

The plaques providing the DNA to be sequenced were used to infect 2 ml of 2 x YT containing 1/100 fold dilution of an overnight JM101 culture. This culture was then grown at 37°C for 5 hr. 1.5 ml of this culture was then pelleted, the supernatant was then decanted and added to 250 µl of PEG (2.5 M NaCl, 20% PEG 6000). After 15 min at room temperature the mixture was centrifuged for 5 min to obtain a pellet, any traces of the PEG solution were removed at this stage and 100 µl of TE was added to the pellet which was then extracted with phenol. Traces of phenol were removed with several extractions of
ether. 10 µl of Na Acetate (3 M pH 5.5) and 250 µl ethanol were added and frozen at -70°C for 60 min. Following the ethanol precipitation, the DNA was pelleted by centrifugation and desiccated to remove traces of EtOH. The DNA pellet was then resuspended in 25 µl TE and stored frozen in Eppendorfs until used.

2.2.14.5 Sequence Reactions

The single stranded template and primer in equimolar quantities were annealed in TM buffer, annealing takes place in an Eppendorf tube maintained at 80°C in a water bath for 3 min and allowed to cool to 30°C. Primed synthesis was then affected by the addition of DNA polymerase (Klenow) fragment, which lacks 5' exonuclease activity, and the 4 precursor nucleoside triphosphates (dNTPs) one of which is radioactively labelled in the alpha position so that the reaction products can be visualised by subsequent autoradiography. Specific termination at each of the 4 different nucleotides is brought about by addition, at the same time chain terminating dideoxy nucleotide triphosphates (ddNTPs). These are incorporated randomly at which point the chain cannot be further extended because the ddNTP lack the 3' hydroxyl group. Four separate reactions are carried out each with a different ddNTP present yielding 4 separate sets of reaction products ending in a specific ddNTP. Any chains not ending in ddNTPs are then chased out by flooding the reaction mixture with a solution of all 4 dNTPs. These chains will stay at the top of the gel in the subsequent fractionation procedure because of their high molecular weight. Prior to loading the reaction was stopped by the addition of 4 µl of formamide dye and boiling for 3 mins in a water bath.
2.2.14.6 Acrylamide Gel Fraction

By creating a gradient of potential difference down the gel, it is possible to increase the size range of fragments that can be separated on an acrylamide gel. This was achieved by utilising the differing resistances of different ionic strength buffers. A potential difference gradient was generated by a gradient of TBE gel buffer concentration with a high voltage drop across a low concentration of TBE at the top of the gel and a low voltage drop across a high concentration of TBE at the bottom. The overall effect reduces the band spacing over the lower portion of the gradient and to increase the band spacing over the upper part of the gel. A crude gradient was obtained by drawing up 6 ml of a 40 ml stock of 0.5 TBE gel mix polymerised with 80 µl TEMED and 80 µl 25% (w/v) APS (150 ml 40% Acrylamide, 50 ml 10 x TBE, 460 g Urea made up to litre with H$_2$O) followed by 7 ml of 5.0 TBE gel mix polymerised with 14 µl TEMED and 14 µl 25% (w/v) APS (150 ml 40% Acrylamide, 500 ml 10 x TBE, 460 g Urea, 50 mg Bromophenol Blue, made up to litre with H$_2$O) and introducing 2-4 air bubbles.

The gradient can then be poured slowly into the gel template in the usual manner. As the last of the gradient gel solution enters the gel template it should be lowered to horizontal to slow down the rate of entry until the template can be carefully filled with the remaining 0.5 TBE stock.

2.2.14.7 Running Conditions

The gel is best run at constant power of around 40 Watts as this maintains a hot, denaturing environment in the gel, thereby reducing the artefacts produced by secondary structures. At this
setting the voltage varies between 1.2 - 1.5 KV and the current between 26-33 mA. The minimum running time depends on the vector, insert and cloning sites used.

2.2.14.8 Fixing and Autoradiography

After the gel had been run it was fixed with 10% v/v Acetic acid and 10% v/v Methanol for 15 min. The gel was then dried, placed in a film cassette with X-Ray film in direct contact with the gel and left exposing overnight.

2.2.15 Colony Hybridization

Using a method adapted from Grunstein and Hogness (1975) it was possible to screen a large number of clones with a radioactive probe. The colonies to be screened were picked onto a nitrocellulose filter, grown, lysed and then the liberated DNA fixed to the filter by baking. After hybridization to a $^{32}$P labelled probe, the filter was monitored by autoradiography. A colony whose DNA gives a positive autoradiographic result can then be recovered from the master plate. This method can also be adapted for bacteriophage plaques.

Lysis of the colonies was carried out on several layers of Whatman 3 MM paper saturated with different solutions. Lysis was initiated by treatment with 0.5N NaOH for 10 min, followed by 3 treatments with 1.0 M Tris-HCl (pH 7.4) of 10 min duration. The filters were then treated with a neutralizing solution 1.5M NaCl, 0.5M Tris-HCl (pH 7.4) for 10 min.

The filters are then immersed for 1 hour in 1 x SSC containing 1 mg ml$^{-1}$ Proteinase K. The filters were then washed twice with 2 x SSC removing all cell debris from the filters. The filters
are then allowed to dry at room temperature for 30 min before being baked for 2 hr at 80°C in a vacuum oven. The filters were then prehybridized for 2 hr at 37°C in prehybridization solution, the probe was then added and left to incubate overnight at 37°C. The filters are then washed several times with 2 x SSC before being dried and autoradiographed.

2.2.16 Carboxypeptidase (CPG2) Assay

0.9 ml of CPG2 buffer (0.8 ml if levels of CPG2 are high) were mixed with 0.1 ml methotrexate (5.4 mg in 20 ml CPG2 buffer) and incubated in a 1 ml cuvette at 37°C.

10 µl to 50 µl of the sample to be assayed was added to the incubated mixture and the rate of change of optical density at 320 nm was measured. Complete hydrolysis gives an absorbance change of approximately 0.5 at 320 nm. The molar extinction coefficient at 320 nm was 8,300 at pH 7.3 and 37°C.

2.2.17 Chloromphenicol Acetyl Transferase (CAT) Assay

A coupled spectrophotometric assay which measured the cleavage of acetyl-coenzyme A was used. The reduction at 37°C of 5,5-dithio-bis-2-nitrobenzoic acid (Ellman's reagent or DTNB) by reduced CoA was followed at 412 nm. 0.95 ml of CAT buffer was incubated in a 1 ml cuvette at 37°C. 5 to 50 µl of sample to be assayed was added depending on enzyme concentration. The addition of 20 µl of chloromphenicol (1.6 mg ml⁻¹) initiated the reaction. The initial rate of reaction was followed at 412 nm.

The molar extinction coefficient at 412 nm was 13,600 at pH 7.8.
2.2.18 Beta-lactamase Assay

Nitrocefin is a cephalosporin used to detect beta-lactamase activity. In aqueous solution at pH 7.0, with the lactam ring intact, nitrocefin has a main absorption peak at 390 nm. Cleavage of the lactam ring by the action of the lactamase results in a shifting of the absorption peak to 486 nm.

A stock solution of Nitrocefin was prepared by dissolving 5 mg of Nitrocefin in 0.5 ml dimethyl sulphoxide and then subsequently made up to 10 ml with 0.1 M potassium phosphate buffer at pH 7.0 containing 1 mM EDTA.

1 ml of the Nitrocefin solution was incubated in a 1 ml cuvette at 30°C. 5 to 50 µl of the sample to be assayed was added. The initial rate of reaction was followed at 480 nm. The molar extinction coefficient at 486 nm was 18,350 at pH 7.0.

2.2.19 Protein Assay

Protein assays, based on Coomassie blue protein binding were carried out using Pierce protein assay reagent. BSA stock solution was diluted to produce a range of protein concentrations, around the unknown concentration. 50 µl of test sample was added to 1 ml of water, a similar volume of water was added to standards. To this was then added 1 ml of Protein assay reagent, and after 5 min the optical density at 595 nm was recorded against a blank. A calibration curve of known concentration of BSA against absorbance at 595 nm was plotted. From this calibration curve the concentration of the test sample can be determined.
2.2.20 Polyacrylamide Gel Electrophoresis (PAGE)

Bio-Rad's 'Protean' double slab electrophoresis cell was used for all polyacrylamide gels. The gel slabs had the dimensions 15 cm x 14 cm x 0.2 cm.

2.2.20.1 Linear Gel

Essentially that described by Laemmli (1970). A running gel consisting of 15 ml Acrylamide solution was added to 22.5 ml of 2 x running gel buffer and made up to 45 ml with water. This mixture was then degassed and polymerised by the addition of 0.5 ml 10% (w/v) APS and 20 μl TEMED. This was then poured into the cast and allowed to set before casting the stacking gel. The stacking gel consists of 2 ml Acrylamide solution, 6 ml 2 x stacking gel buffer made up to 12 ml with water. This mixture was also degassed and polymerised by adding 0.1 ml 10% (w/v) APS and 10 μl TEMED. Placing the comb in the gel displaces any bubbles and any excess mixture. After 20 min polymerisation was complete and the comb was carefully removed. The wells were then flushed with electrophoresis buffer.

2.2.20.2 Gradient Gel

Gradient gels were utilised to obtain greater resolution of low molecular weight proteins. The gradient required for the gel is produced by the mixing of two solutions (light and dense) in a gradient mixer. The dense solution consists of 18 ml 40% (w/v) acrylamide solution, 3.8 ml 3% (w/v) bis-acrylamide solution and 4.5 ml lower Tris. The light solution contains 5.4 ml 40% (w/v) acrylamide, 2.3 ml 3% (w/v) bis-acrylamide and 4.5 ml lower Tris made up to 26 ml with water. A gradient maker was used to mix 15 ml of the light and dense
solutions once they are 0.18 ml 10% (w/v) APS 14 µl TEMED. Once the
gels have been poured the gel can be overlayed with 30% methanol and
25% lower Tris. Following polymerisation the overlay solution can be
removed and the top of the gel rinsed with water. The stacking gel
consists of 1.1 ml Acrylamide solution, 2.5 ml upper Tris and made up
to 10 ml with water 50 µl 10% (w/v) APS and 85 µl TEMED will polymerise
the stacking gel.

2.2.20.3 Sample Preparation

A sample from a fermentation culture was prepared for running
on a polyacrylamide gel by diluting the sample to an optical density of
0.2 at 600 nm. 0.5 ml of this diluted sample was then added to 100 µl of
sample buffer, sonicated for 5 sec and then boiled for 5 min. 25 µl of
the treated sample was then loaded on the gel.

For a sample of cell paste, 10 mg of cell paste was added to
400 µl of sample buffer and then treated as before. 20 µl of sample
was loaded on the gel.

2.2.20.4 Running Conditions

Linear gels were run at 30 mA for approximately 5 hr.
Gradient gels run at 40 mA required 8 hr. The electrophoresis buffer
used was SDS-PAGE buffer.

2.2.20.6 Staining and Destaining

Gels were staining with a Coomassie blue solution for
approximately 2 hr. While destaining was normally overnight with
slight agitation.
2.2.20.6 **Quantitation of Protein Components in Gel Electrophoresis**

Destained gels were shrunk and dried between two transparent membranes. The dehydrated transparent gel was then scanned using a Joyce-Loebl Chromoscan 3. The areas under the traced peaks corresponding to the protein bands are given as integral values which can then be used to express an individual protein component as a % total amount of the soluble protein.

2.2.21 **Estimation of Copy Number**

There are several methods available to estimate the plasmid copy number. They can be classified into two groups: Methods which employ gene dosage and those methods which measure the ratio of plasmid to chromosomal DNA.

2.2.21.1 **Measurement of Gene Dosage**

Depends upon the gene dosage dependent production of antibiotic inactivating enzymes, as demonstrated by Uhlin and Nordstrom (1978). This method is therefore limited to plasmids which encode antibiotic inactivating enzymes such as beta-lactamase.

The beta-lactamase activity/cell of bacteria containing an unknown number of plasmids which code for enzyme activity is measured, and related to the beta-lactamase activity/cell of bacteria containing a known number of plasmids. It is therefore possible to estimate the number of plasmids per cell with reference to a standard cell/plasmid combination, usually pBR322/HB101, which has a quoted copy number of 50.
2.2.21.2 Measurement of Plasmid/Chromosome Ratio

This can either be done by hybridisation analysis of total DNA or by physical separation of plasmid and chromosome species. The latter method was employed in our study of copy number. This technique involves screening whole cell lysates, utilising the electrophoretic separation of plasmid and chromosome DNA species on agarose, and then quantifying the amount of DNA in the bands. The screening method was a modified version of that described by Eckhardt (1978) and further adapted for copy number estimation by Projan et al. (1983). 1 ml of culture diluted to an optical density of 2.0 at 600 nm was centrifuged in an eppendorf centrifuge for 5 min, the pellet was then resuspended in a lysing solution (20 mM Tris-Cl, pH 8.0, 10 mM EDTA, 100 mM NaCl, 20% w/v sucrose, 25 \( \mu \)g ml\(^{-1} \) lysozyme, and 2 units ml\(^{-1} \) pancreatic RNase). This suspension was then incubated at 37\(^\circ\)C for 30 min. Lysis was effected by addition of 100 \( \mu \)l 2% SDS while vortexing the cells. The lysate was vortexed at top speed for an additional minute, freeze-thawed (-70-37\(^\circ\)C) twice and treated with Proteinase K at 10 \( \mu \)g ml\(^{-1} \) for 30 min at 37\(^\circ\)C. The sample was prepared for electrophoresis by addition of 50 \( \mu \)l of tracking dye.

The gels obtained from the electrophoresis of the whole cell lysates are stained and photographed in the usual manner. A negative of the photographed gel was scanned using Joyce Loebl Chromoscan-3. The scanner gives an integral value of the area under the peaks which correspond to chromosome and plasmid on the gel. The ratio of plasmid to chromosome was then used to calculate the copy number (Fig. 2.2).

2.2.21.3 Definition of Plasmid Copy Number

Copy number was defined throughout this thesis as the number
Estimation of Copy Number using the Plasmid/Chromosome Ratio

Electrophoretic separation of plasmid and chromosome species on an agarose gel was quantitatively estimated by scanning a photographic negative of the gel. The integral values obtained corresponding to the plasmid and chromosome species were then used to calculate the number of plasmids per chromosome equivalents in the following equation.

\[
\text{Copy number} = \frac{\text{Plasmid integral} \times \text{Mwt of chromosome}}{\text{Chromosomal integral} \times \text{Mwt of plasmid}}
\]
of plasmids per chromosome equivalent. Defining copy number as copies per chromosome, rather than copies per cell relies on the ability to assume that each chromosome was associated at all times with a specific number of plasmid copies.

It has been shown from a variety of studies (Rownd, 1969; Bazaral and Helinski, 1970; Gustafsson and Nordstrom, 1975; Gustafsson et al., 1978) that plasmid copy number is controlled and varies independently of chromosomal replication; both plasmid and chromosomal content vary under different growth conditions and they do so independently. Therefore, the definition of copies per chromosome can only be used with confidence when applied to the relative copy number of different plasmids in the same host when all measurements were done under constant growth conditions.

2.2.22 Continuous Culture

2.2.22.1 Apparatus and Culture Conditions

Continuous culture was carried out in a LH500 series fermenter and control package. A flow diagram of the continuous culture apparatus (chemostat) are described in Fig. 2.3. A 1 l continuous culture vessel was used with its overflow height set to a working volume of 600 ml. Cultures were maintained unless otherwise stated in the text at 37°C ± 1°C and at a pH of 7.0 ± 0.1 with an aeration rate of 1 vvm. Following inoculation cultures were allowed to grow batchwise for 4 to 5 hr before the flow of fresh medium was initiated. The dilution rate (D) was set to 0.1 hr⁻¹, which was equivalent to a mean generation time of 6.93 hr. Samples were removed periodically and serial diluted onto isosensitest agar from which the %
Fig. 2.3

Flow Diagram of a Chemostat
Ap<sup>R</sup> cells in the population could be calculated. Viable counts and dissolved oxygen tension (DOT) gave a guide to the steady state conditions existing in the culture.
CHAPTER 3

EFFECT OF par LOCUS ON PLASMID STABILITY,
COPY NUMBER AND HOST FITNESS
INTRODUCTION

A cis acting DNA region of pSC101 designated par was shown to confer segregational stability not only to pSC101 but also to the unrelated plasmid pACYC184 (Meacock & Cohen, 1980). par is thought to cause the accurate partitioning of plasmids at cell division by binding to specific sites on the outer membrane fraction of the host cell (Gustafsson et al., 1983). The par function has previously been utilised in expression vectors (Skogman et al., 1983; Zurita et al., 1984), although complete stability was not obtained, an increase in stability over the par- vector was achieved.

Since a basic requirement of our vectors would be segregational stability, it was of interest to investigate the stability of our commonly used vectors and the effect par has on them. There are several reports in the literature of certain DNA elements e.g., ISSO, Lambda, Mu, P1 and P2 (Hartl et al., 1983; Edlin et al., 1975; Edlin et al., 1977; Lin et al., 1977; Bier and Hartl, 1983) conferring a growth advantage on host cells containing them. The possibility that par enhances the ability of a host cell to compete favourably with any plasmid-free cells which may arise, was therefore also examined.

Continuous cultivation in a chemostat was chosen as an effective method to examine the stability of plasmids, since continuous cultivation at low growth rates with phosphate or carbon nutrient limitation have been shown as particularly likely to emphasise any segregational instability a plasmid may exhibit (Jones et al., 1980; Wouters et al., 1980; Moek et al., 1981). The study of competition between plasmid-free cells and plasmid-bearing cells is difficult to
perform in batch culture due to fluctuations in both nutrient availability and physiological state of the cells. These problems were overcome with the use of continuous culture techniques.

RESULTS AND DISCUSSION

3.1 Plasmid Stability under Continuous Cultivation

The stability of pAT153, pNM21, pUC8 and their par derivatives in the E. coli host W5445 were examined by continuous culture. Cultures were subjected to either glucose or phosphate nutrient limitation in non-selective medium at a dilution rate of 0.1 hr⁻¹.

3.1.1 Stability of pAT153

pBR322 in the host W5445 when grown in a chemostat under phosphate limitation produced plasmid free segregants after 50 generations, resulting in complete plasmid loss from the culture (Jones et al., 1980). A similar loss was observed by Wouters et al. (1980) when pBR322 was in the host PC221. The pBR322 species used by Wouters et al. (1980) and Jones et al. (1980) is regarded as being genetically different from the pBR322 species used by Noack et al. (1981) which did not exhibit instability in the hosts GY2354 and GM31. Although the genotypes of the host strain and experimental differences in medium might also be reasons for the contrary results concerning pBR322, pAT153, the high copy number derivative of pBR322 (Fig. 3.1) (Twigg and Sherratt, 1980), also provides contradictions. Jones and Melling (1984) explained the stability of pAT153 in HB101, in contrast to the instability observed in W5445 observed by Primrose et al.
Fig. 3.1

Plasmid pBR322

Schematic diagram of pBR322, showing the location and direction of transcription of the bla and tet genes. The origin of replication, showing the direction of the RNA II transcript is also identified. The two contiguous HaeII B and C fragments which on deletion produce pAT153 are also identified.
(1984), as being due to host genetic background variation. Another possible explanation is that insufficient generations were allowed to elapse before plasmid-free segregants manifested themselves. pAT153 stability was only monitored over 100 generations, while pAT153 instability was observed after 150 generations. In an attempt to resolve this paradox, an 'in-house' construction of pAT153 in the strain W5445 was monitored under the carbon and phosphate limiting conditions of a chemostat. pAT153 in W5445 proved to be stable for over 100 generations (Fig. 3.2). The stability of pAT153 provides a reference point to which all subsequent constructions can be compared.

pAT153 is maintained in the population over a longer period of time than pBR322, showing greater stability. The persistence of pAT153 compared to pBR322 could be explained by the elevated copy number of pAT153 compared to pBR322 (Table 3.1). pBR322 has a copy number of 50 (Timmis, 1981) while pAT153 has a copy number of 150 due to the ROP deletion (Cesareni et al., 1984). It can be postulated that pAT153 in the host W5445 will after prolonged cultivation have produced plasmid free segregants, and it is only due to the raised copy number of pAT153 in contrast to pBR322 that it does not result in plasmid free segregants sooner.

Twigg and Sherratt (1980) described pAT153 as a 'potentially useful cloning vector' because of its higher copy number and its low frequency of mobilisation; considered an advantage in terms of biological containment for in vitro genetic manipulation. It was for these reasons that pAT153 has become ubiquitous in genetic engineering, only recently superseded by the pUC series of cloning vectors developed by Vieira and Messing (1982).
Sequenational Stability of pAT153 in the Host W3443

Grown under carbon (●) and phosphate (○) limiting conditions. pAT153 shown to be stable over a 100 generation under both forms of nutrient limiting conditions.
### Table 3.1
Calculated Copy Numbers from Plasmid/Chromosome Ratios Obtained from Densitometry of EtBr-Stained Electrophoretic Gels

<table>
<thead>
<tr>
<th>Host</th>
<th>Plasmid</th>
<th>Plasmid size (kb)</th>
<th>Copy Number (a)</th>
<th>% Standard deviation (b)</th>
<th>(c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>W5445</td>
<td>pBR322</td>
<td>4.36</td>
<td>50</td>
<td>+4.3</td>
<td>3</td>
</tr>
<tr>
<td>W5445</td>
<td>pAT153</td>
<td>3.68</td>
<td>147</td>
<td>+4.0</td>
<td>3</td>
</tr>
<tr>
<td>W5445</td>
<td>pAT153 par</td>
<td>4.10</td>
<td>195</td>
<td>+2.7</td>
<td>3</td>
</tr>
<tr>
<td>W5445</td>
<td>pNM21</td>
<td>8.20</td>
<td>55</td>
<td>+3.7</td>
<td>3</td>
</tr>
<tr>
<td>W5445</td>
<td>pNM21 par</td>
<td>6.10</td>
<td>44</td>
<td>+4.2</td>
<td>3</td>
</tr>
<tr>
<td>W5445</td>
<td>pUC8</td>
<td>2.72</td>
<td>495</td>
<td>+4.3</td>
<td>3</td>
</tr>
<tr>
<td>W5445</td>
<td>pUC8 par</td>
<td>3.20</td>
<td>460</td>
<td>+1.8</td>
<td>3</td>
</tr>
</tbody>
</table>

Plasmid copy numbers were calculated per chromosome. Assuming one plasmid per chromosome equivalent exists and based on a figure of $3.8 \times 10^3$ kb for the size of a non-replicating *E. coli* chromosome (Shepard and Polisky, 1979). The calculated copy numbers per chromosome equivalent are given in column (a), the standard deviations in column (b) and the number of independent trials in column (c). The copy number calculation was as follows:

$$\text{Copy number} = \frac{\text{plasmid}}{\text{chromosome}} \times \frac{\text{Size of } E. \text{ coli chromosome (kb)}}{\text{Size of plasmid (kb)}}$$
3.1.2 Stability of pNM21

pNM21 was derived from the cloning of the *Pseudomonas* gene encoding CPG₂ (*cpp*) as a *Bal*II fragment into the *Bam*HI site of pAT153 (Fig. 3.3). Expression was from its own promoter and read through from the *tet* promoter of pAT153 (Minton et al., 1983a).

Having established previously (3.1) that pAT153 is stably maintained for 100 generations under both carbon and phosphate limitation, the expectation was that pNM21, employed as an expression vector, would exhibit the stability of pAT153. However, pNM21 was lost from the host W5445 after 10 generations under phosphate limitation and after 30 generations under carbon limitation (Fig. 3.4). This stability is more characteristic of pBR322 than pAT153; on examination the copy number of pNM21 was found to be 55 (Table 3.1). The cloning of the *cpp* gene into pAT153 has effectively reduced its copy number to a level comparable with pBR322 and in so doing has increased its segregational instability.

pNM21 in W5445 was the plasmid host combination used in the batch fermentation for the large scale production of CPG₂. A recurring problem in the batch growth of W5445 containing pNM21 was plasmid instability. When monitoring the fermentation, segregational instability invariably took place half way through a typical 12 hour fermentation (Fig. 3.5). This plasmid loss has a correspondingly negative effect on the expression of CPG₂. Therefore any element that could maintain the plasmid in the population until the end of the fermentation would increase expression and the final yield obtained.
Schematic diagram of pNM21, showing the position and orientation of the cpg gene within the vector.
Fig. 3.4

Segregational Stability of pNM21 in the Host W5445

Grown under carbon (●) and phosphate (○) limiting conditions. pNM21 being lost from the population after 30 generations under carbon limitation, while being lost after 10 generations under phosphate limitation.
Variation of CPG2 Expression and Segregational Stability of pNM221 in the Host W5445 During Large-Scale Fermentation

Fig. 3.5

[Graph showing variation of CPG2 expression and segregational stability over time.]
3.1.3 **Stability of pAT153 par**

pAT153par contains the complete pAT153 replicon with the pSC101 region containing the par locus cloned via EcoRI linkers into the EcoRI site of pAT153. The subsequent vector produced was designated pAT153 par (Fig. 3.6A). The AvaI digest of pAT153 par shown in Fig. 3.6B demonstrates the orientation of the par fragment. Analysis of pAT153 par stability in the host W5445 under carbon and phosphate nutrient limitation of a chemostat revealed no instability (Fig. 3.7). pAT153 par, like pAT153, appeared to be fully stable.

3.1.4 **Stability of pNM21 par**

The plasmid pNM21 par was obtained by cloning the cpo gene on a BamHI fragment into the BamHI site of pAT153 par (Fig. 3.8). The resulting plasmid, pNM21 par, when examined under carbon and phosphate limitation in a chemostat, although having a similar copy number to pNM21 reversed the previous instability observed in pNM21. pNM21 par was stably maintained in the host W5445 for 100 generations under both carbon and phosphate limitations (Fig. 3.9) compared to pNM21 which was lost from the population after 30 and 10 generations for carbon and phosphate limitation respectively.

3.1.5 **Stability of pUC8**

pUC8 one of the pUC series of cloning vectors developed by Vieira and Messing (1982), (Fig. 3.10) exhibits an elevated copy number compared to pAT153 (Table 3.1). It was therefore of interest to examine its stability. pUC8 was lost from the host W5445 under the phosphate limitation in a chemostat after 15 generations, while no segregation of instability was observed under carbon limitation (Fig. 3.10).
**Fig. 3.6A**

Plasmid pAT153 par

Schematic diagram of pAT153 par, showing the position and orientation of the par locus within the vector. The orientation of the par locus (determined in Fig. 3.6B) is shown by the asymmetrically located Aval site within the cloned fragment.
Determination of the Orientation of the par Locus within pAT153 par

AvaI digestion resulted in the production of two fragments compatible only with the orientation shown in Fig. 3.6A, the smaller fragment being approximately 1.43 kb instead of 1.796 kb.

Lane (1) pAT153 par AvaI digest (1.43, 2.67 kb). (2) pAT153 AvaI digest (3.68 kb). (3) Lambda AvaI digest. (4) pAT153 par uncut. (5) pAT153 uncut
Segregational Stability of pAT153 par in the Host US445

Grown under carbon (●) and phosphate (○) limiting conditions. pAT153 par shown to be stable over 100 generations under both forms of nutrient limiting conditions.
**Fig. 3.8**

**Plasmid pNM21 par**

Schematic diagram of pNM21 par, showing the location and orientation of the par locus and the cpa gene within the vector.
Grown under carbon (•) and phosphate (○) limiting conditions. pNM21 par shown to be stable over 100 generations under both forms of nutrient limiting conditions; in contrast to the instability of pNM21.
Fig. 3.10

Plasmid pUC8

Schematic diagram of pUC8, showing the location and orientation of the alpha-peptide containing the polylinker region.
3.11). This result, obtained in combination with the segregational instability of pNM21 observed under phosphate limitation, supports the general finding that plasmid instability is increased by the more stringent conditions for DNA replication under phosphate limitation (Godwin and Slater, 1979).

3.1.6 Stability of pUC8 par

The par locus was isolated from pAT153 par as a 479 bp HindIII – AatII fragment, subsequently cut with TaqI and then cloned into the AccI site within the polylinker region of pUC8 to give pUC8 par. Fig. 3.12 shows the orientation of par within the polylinker region. Analysis of pUC8 par stability in the host W5445 under carbon and phosphate limitation in a chemostat revealed no instability (Fig. 3.13). The instability previously observed in pUC8 under phosphate limitation was reversed by the cloning of the par locus into the plasmid.

3.2 Can the Calculated Copy Numbers Account for the Varying Instabilities of the Plasmid Examined?

The partitioning mechanism of ColEl and its derivatives is unclear but it is probable that they rely upon random partitioning alone to ensure that daughter cells inherit a plasmid copy. Random partitioning will result in an equal probability of any one plasmid copy going to either of the two daughter cells at cell division and if there are 2n copies per cell, the daughter cell will receive any number between 0 and 2n plasmids. Consequently, plasmid free cells will be generated at a frequency which is related to the copy number of the
Segregational Stability of pUC8 in the Host W5445

Grown under carbon (●) and phosphate (○) limiting conditions, pUC8 shown to be stable over 100 generations under carbon nutrient limitation. While pUC8 was rapidly lost from the population, within 15 generations, under phosphate limiting conditions.
Plasmid pUC8_par

Schematic diagram of pUC8_par, showing the location and orientation of par locus within the polylinker region.
Fig. 3.13

Segregational Stability of pUC8 par in the Host W5443

Grown under carbon ( • ) and phosphate ( O ) limiting conditions pUC8 par shown to be stable over 100 generations under both forms of nutrient limiting conditions; in contrast to the instability of pUC8.
plasmids at cell division. The relationship between copy number and segregation frequency is derived from the binomial distribution (Novick et al., 1975) and can be expressed as

\[ P_0 = 2(1-n) \]

where \( P_0 \) = segregation frequency per cell per generation
\( n \) = copy number per dividing cell.

Fig. 3.14 shows a graph of segregation frequency versus copy number and demonstrates that the greater the copy number the less likely it is that a plasmid free segregant will arise. The nature of stability experiments means that segregation frequencies of \( >10^{-5} \) can be detected. Using the above expression, copy number per cell at cell division required for a segregation frequency of \( 10^{-5} \) per cell per generation was calculated to be 18. Therefore, cells at division containing plasmids with a copy number of 18 or greater, would theoretically be expected to be stably maintained. Conversely, plasmid free segregants should be detected when plasmids are maintained at copy numbers less than 18.

Table 3.1 shows the calculated average plasmid copy numbers per chromosome in a bacterial population. For the purpose of calculating segregation frequencies however, plasmid copy number per chromosome is insufficient, since plasmid copy number per dividing cell is required it cannot be assumed that dividing cells possess only 2 chromosomes.

Cooper and Helmstetter (1968) proposed a model to explain the observation that the amount of DNA per cell varies with growth rate. Their experimental data demonstrated that the transit time of
This relationship is derived from the binomial distribution and may be expressed mathematically as $P_0 = (1-n)$, where $P_0$ is the frequency of plasmid-free cell production (per cell per generation) and $n$ is the copy number of the dividing cell. From this it is readily seen that the higher the copy number the lower the probability of plasmid-free segregants being produced.

$$P_0 = 2^{(1-n)}$$
replication forks for cells with doubling times less than 70 min was almost constant in *E. coli* B/r. This indicated that the frequency of initiation was the factor which could account for varying DNA content with doubling time. The time required to replicate one chromosome was around 40 min and cell division did not occur until 20 min after termination of chromosome replication. On this basis they predicted that multiple rounds of replication initiation must occur in cells with doubling times faster than 40 min.

The generation times of the plasmid containing cells during copy number determinations were around 30 min. According to the Cooper-Helmstetter model, cells with a doubling time of this order will possess approximately 3.75 genome equivalents per dividing cell.

The copy number calculation relies upon the ratio of plasmid content to chromosome content and involves the size of the *E. coli* chromosome to calculate the number of copies of a plasmid of known size. Clearly therefore, the number of chromosome equivalents per cell is required for accurate estimation of plasmid copy numbers. Since these copy numbers were carried out on non-synchronous, exponentially growing cells possessing all possible chromosome configurations, the figures in Table 3.1 are at best average values. Therefore in order to obtain estimations of copy number at cell division, the figures in Table 3.1 were multiplied by 3.75 and the results and calculated segregation frequencies are shown in Table 3.2. These segregation frequencies predict that all the plasmids should be very stable but this was evidently not the case.
<table>
<thead>
<tr>
<th>Host</th>
<th>Plasmid</th>
<th>Copy number copies per chromosome</th>
<th>Copy number copies per dividing cell</th>
<th>Segregation frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>W5445 pBR322</td>
<td>50</td>
<td>187</td>
<td></td>
<td>1.02 x 10^{-56}</td>
</tr>
<tr>
<td>W5445 pAT153</td>
<td>147</td>
<td>551</td>
<td></td>
<td>α</td>
</tr>
<tr>
<td>W5445 pAT153 par</td>
<td>195</td>
<td>731</td>
<td></td>
<td>α</td>
</tr>
<tr>
<td>W5445 pNM21</td>
<td>55</td>
<td>206</td>
<td></td>
<td>1.94 x 10^{-62}</td>
</tr>
<tr>
<td>W5445 pNM21 par</td>
<td>44</td>
<td>165</td>
<td></td>
<td>4.28 x 10^{-50}</td>
</tr>
<tr>
<td>W5445 pUC8</td>
<td>495</td>
<td>1856</td>
<td></td>
<td>α</td>
</tr>
<tr>
<td>W5445 pUC8 par</td>
<td>460</td>
<td>1725</td>
<td></td>
<td>α</td>
</tr>
</tbody>
</table>

The copy number values from Table 3.1 were multiplied by 3.75 in order to estimate the copy numbers and segregation frequencies at division.
3.3 Plasmid Copy Number Changes during Cultivation

Throughout the chemostat cultures cell samples were removed and analysed on agarose gels for their plasmid/chromosome content. In the steady state environment of a chemostat where the growth conditions of the host are constant and plasmid-free segregants have not arisen, the plasmid/chromosome ratio can be used to estimate copy number and describe its relative changes during cultivation.

In all cultures examined under continuous culture a gradual drop in copy number was observed (Fig. 3.15). During the same period beta-lactamase activity which is directly related to copy number (Uhlin and Nordstrom, 1978) also fell. This drop in copy number during cultivation is considered to be a phenotypic response to the nutrient limitation rather than selection of mutants with a lower copy number (Jones et al., 1980). This was shown by regeneration of the reduced copy number to its higher level after a single cycle of batch growth in a complex media. Where plasmid loss from the host population occurred the copy number was reduced to approximately 3-4. For plasmids present at such low copy numbers, where no mechanism exists for active segregation of plasmids to the daughter cells during division, the frequency of plasmid-free segregants will be high.

The fraction of plasmid-free segregants should relate directly to the copy number of the plasmids involved (Table 3.2); for plasmids normally present at the high copy numbers of pAT153, the frequency of plasmid-free segregants arising will be low, but where the copy number is reduced (as in pNM21) or with prolonged culturing under nutrient limitation, the frequency of plasmid-free segregants will increase. Plasmids containing par also experience an equivalent drop in copy number, but no plasmid-free segregants occur. Therefore it can
Reduction in Copy Number of pAT153 During Continuous Culture

Copy number analysed on an agarose gel was shown to fall, over 90 generations of continuous culture under carbon nutrient limitation.
be inferred that an active partitioning mechanism must exist for the segregation of plasmids containing par.

3.4 **Effect of par on Copy Number**

Meacock and Cohen (1980) have identified the par function and characterised it as being independent of copy number control and plasmid replication function. The possibility that par could be enhancing stability by increasing the copy number of the plasmids and not as a direct result of active partitioning, was investigated by examining the copy number differences between plasmids with and without par. Table 3.1 shows that no significant difference was identified. Therefore, it can be assumed that the segregational stability produced by par is not due to increased copy number and a decrease in the probability of plasmid free segregants occurring.

3.5 **Effect of par on the 'Fitness' of Host**

The rise of plasmid free segregants in a culture results from the failure of a daughter cell to inherit a plasmid at cell division and the subsequent growth advantage conferred on the plasmid-free cells (Adams et al., 1979). The growth advantage of plasmid free cells, compared to plasmid containing cells is assumed to result from the added metabolic burden of plasmid containing cells, resulting in a competitive disadvantage under growth conditions where the characteristics coded for by the plasmids, are not or do not need to be expressed. The synthesis and replication of a redundant plasmid utilises elemental and energy resources which otherwise might be diverted to biomass production and a high population growth rate (Godwin and Slater, 1979). When investigating plasmid stability it is
difficult to distinguish between segregational instability and the
effect of any competitive advantage or disadvantage conferred by the
plasmid to the host cell. Both are contributing factors to be
considered in the stability of plasmid within the population.

The effect par has on the segregational stability of plasmids
within the population has already been examined. Therefore, the next
stage was to examine the effect that par has on the 'fitness' of the
host cell. The fitness was assessed in a series of competition
experiments between plasmid-bearing cells and isogenic plasmid-free
cells. The competition experiments performed to assess the effect of
par on the fitness of host cells were carried out in response to the
results obtained by Jones (1986), whose initial data indicated that par
may function in plasmids by enhancing the 'fitness' of the host
organism when in competition with isogenic plasmid free cells.

pAT153 and pAT153 par were used to compare the fitness of
W5445 containing the plasmids with and without par in competition with
any plasmid free W5445 cells which may arise. Both plasmids have
previously been determined to be segregationally stable over a 100
generation period under the same conditions (3.1/3.3). In the
following experiments the competition event was initiated by the
addition of varying volumes of inocula to a steady state population in
a chemostat with a dilution rate of 0.1 hr⁻¹, while under carbon or
phosphate limitation. The inoculum was provided from a parallel
chemostat also at steady state with the same dilution rate and nutrient
limitation conditions. The fate of the plasmid containing organisms
was monitored periodically, using a beta-lactamase plate assay, samples
from the chemostat were plated out onto drug free isosensitest agar and
the proportion of plasmid-free cells determined using the beta-
Lactamase assay as an ampicillin resistance marker.

Table 3.3 summarises the competition experiments carried out and the results obtained. Under carbon limitation it appears that the ratio for plasmid containing cells to plasmid free cells determined the final outcome. While under phosphate limiting conditions, plasmid free W5445 predominates irrespective of the proportions of plasmid free cells to plasmid containing which are added. The exception proved to be pAT153 par which remained dominant when challenged by a 5 or 10% inoculum of plasmid free W5445. It required an inoculum addition of 50% plasmid free W5445 to out compete the W5445 containing par. In identical circumstances plasmid free W5445 came to predominance when only a 5% inoculum of plasmid free W5445 was added to W5445 containing pAT153 (Fig. 3.16).

Plasmid-free cells predominate under phosphate limitation, while under carbon limitation the ratio of plasmid containing cells to plasmid-free cells determines the final ratio. Similar findings in competition experiments between Rpl-containing and Kpl-free cells, (Helling et al., 1977) and between pBR322-containing and pBR322-free cells (Primrose et al., 1984) were found.

The results indicate the importance of environment on the survival of plasmid-containing cells in competition with plasmid-free cells, and the relative number of the two strains has a significant effect on the outcome of competition events. The growth advantage of plasmid-free cells was demonstrated under carbon and phosphate limitation where competition between equal numbers of plasmid-containing cells and plasmid-free cells always resulted in takeover by the plasmid free. It is therefore significant that under phosphate limitation W5445 does not outcompete W5445 containing pAT153 par. In
Table 3.3
Summary of Competition Events Carried out in Continuous Culture

<table>
<thead>
<tr>
<th>Inoculum</th>
<th>Predominant strain following competition</th>
</tr>
</thead>
<tbody>
<tr>
<td>pAT153/W5445</td>
<td>5% W5445</td>
</tr>
<tr>
<td>pAT153 para/W5445</td>
<td>5% W5445</td>
</tr>
<tr>
<td>W5445</td>
<td>5% pAT153 para/W5445</td>
</tr>
<tr>
<td>W5445</td>
<td>5% pAT153 para/W5445</td>
</tr>
<tr>
<td>50% W5445</td>
<td>50% pAT153 para/W5445</td>
</tr>
<tr>
<td>pAT153/W5445</td>
<td>5% W5445</td>
</tr>
<tr>
<td>pAT153 para/W5445</td>
<td>5% W5445</td>
</tr>
<tr>
<td>pAT153 para/W5445</td>
<td>10% W5445</td>
</tr>
<tr>
<td>pAT153 para/W5445</td>
<td>50% W5445</td>
</tr>
<tr>
<td>W5445</td>
<td>10% pAT153/W5445</td>
</tr>
<tr>
<td>W5445</td>
<td>10% pAT153 para/W5445</td>
</tr>
<tr>
<td>50% W5445</td>
<td>50% pAT153 para/W5445</td>
</tr>
</tbody>
</table>

Where both cultures were at similar states, cultures grown in parallel chemostats run at the same dilution rates.
Variation in Maintenance of pAT153 par and pAT153 in a Population of W544S at Steady State under Phosphate Limitation Conditions

When challenged with an inoculum of 5% plasmid-free W5445, pAT153 par was maintained (a), while pAT153 was lost from the population (b).

- addition of plasmid free inoculum
identical circumstances W5445 cells containing pAT153 would be outcompeted by plasmid-free W5445 cells. This increased fitness corresponds with the increased growth rate of W5445 containing pAT153 par compared to pAT153 (Table 3.4). This is somewhat unexpected since pAT153 par with its increased molecular weight (+400 bp) over pAT153 would produce a greater metabolic burden in W5445 and thereby reduce W5445 growth rate. This increased fitness cannot be explained in terms of copy number differences since pAT153 and pAT153 par appear to have very similar copy number values (Table 3.1).

Concluding Remarks

For plasmids to be maintained within a population of cells their replication must be controlled so that the replication rate can be adjusted to the growth rate for cells or cell division must be stalled until sufficient replication has occurred. Subsequent partitioning of plasmid copies at cell division by either an active or passive mechanism need only ensure that both daughter cells receive at least one plasmid copy.

Although many low copy number plasmids are actively partitioned, (Meacock and Cohen, 1980; Nordstrom et al., 1980) no active partitioning mechanism has been demonstrated for the multicopy plasmid ColE1. This plasmid is the progenitor of many of the commonly used cloning vectors pBR322 (Bolivar et al., 1977a), pAT153 (Twigg and Sherratt, 1980) and the pUC vectors (Vieira and Messing, 1982) all utilising the ColE1 replicon. ColE1 as a “naturally” occurring plasmid is by definition stable. Difficulty arises in the study of its stability, since it is only possible to study ColE1 in a colicin resistant background to prevent selection against plasmid free
Table 3.4  
Growth Rate Variation Between Organisms Containing pAT153. 

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Host</th>
<th>Growth Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>W5445</td>
<td>0.5</td>
</tr>
<tr>
<td>pAT153</td>
<td>W5445</td>
<td>0.23</td>
</tr>
<tr>
<td>pAT153 par</td>
<td>W5445</td>
<td>0.3</td>
</tr>
</tbody>
</table>

Growth rates obtained by determination of the $\mu_{max}$ of the various microorganisms containing the various plasmids, employing the growth media used for all the continuous culture experiments.
Jones et al. (1980) examined the stability of pDS1109, a derivative of ColEl, which does not produce colicin El. pDS1109 was found to be stably maintained in W5445 under the nutrient limitation of a chemostat.

Because the ColEl derivative plasmids examined in this chapter are all present at high copy number it could be argued that there is no need for an active partitioning function and the instability generated under nutrient limitation is unnatural. However, the finite chance of plasmid-free segregants arising and coming to predominance as a consequence of plasmid-free cells ability to outgrow plasmid-bearing cells (as described in 3.2) make active partitioning of plasmids a requirement for maintenance. The ability of the par locus to stabilise unstable replicons as demonstrated in pNM21 par and pUC8 par make it an essential component in the construction of stable vectors. It should be noted, that recent research has shown that whilst par can improve the stability of some vectors (pBR322 and pACY184 carrying the tryptophan operon in E. coli) it cannot stabilise them completely. The finding that pAT153 par-containing cells have a growth advantage over isogenic plasmid-free cells contradicts the similar situation with pAT153. The increased fitness and growth rate of cells containing pAT153 par compared to pAT153, appears to be linked to the addition of the par function, either in its inactivation of genes which have an adverse effect on growth or on par enhancing the growth rate.

There are precedents for specific DNA sequences dramatically altering the growth properties of E. coli K12 cells. Transposon Tn5 is a composite transposon consisting of a kanamycin neomycin resistance determinant bounded by two inverted 1.5 kb IS50 sequences. The right
hand element, IS50R, encodes the transposase and its inhibitor, while
IS50L has a mutation which results in the production of non-functional
truncated polypeptides. Cells possessing the IS50R element alone have
been shown to have an initial faster growth rate than their IS50-
counterparts in chemostat culture (Biel and Hartl, 1983; Hartl et al.,
1983).

There has also been reports of increased 'fitness' of cells
due to transposition of IS1 into the tetracycline region (Chew et al.,
1986; Deretic et al., 1984). Lee and Edlin (1985) showed that
expression of the tet gene has an adverse effect on the reproductive
fitness of plasmid containing bacteria. It has been proposed by Moyed
and Bertrand (1983) that a mechanism exists by which TcR proteins
adversely affect growth, via alteration to the cell membrane.
Inactivation of the tet gene in pBR322 and its derivatives (including
pAT153) increased the fitness of the host containing the plasmid (Lee
and Edlin, 1985; Chew et al., 1986).

One of the basic requirements of a vector is stability. The
ability of the par locus to enhance the segregational stability of
vectors, therefore made it a requirement for all future
cloning/expression vectors.
CHAPTER 4
CONSTRUCTION AND CHARACTERISATION OF A SERIES OF NOVEL CLONING VECTORS
4. INTRODUCTION

The pUC series of plasmids represents a series of high copy number cloning vectors which offer a wide range of unique cloning sites combined with a simple means of recombinant selection (Vieira and Messing, 1982) (Fig. 4.1). The linker region carrying the multiple cloning sites lies within the alpha-peptide of the lacZ gene (N-terminal amino acids 1 to 145). The functional polypeptide produced, complements a defective lacZ gene (missing amino acids 11 to 41) carried by the host gene in JM83, or resident episome in JM101. Colonies carrying plasmid DNA are therefore blue in the presence of the chromogenic substrate X-Gal. DNA fragments cloned into the linker region produce a non-functional alpha-peptide resulting in colourless (white) colonies in the presence of X-Gal.

In this chapter a series of cloning vectors are described which extend the utility of pUC plasmids. By extending the number of unique cloning sites within the linker region, and rearranging the orientation of certain sites within the linker region a more versatile group of cloning vectors have been produced. During the construction of the pMTL plasmids the high copy number of the pUC plasmids was examined and the pMTL plasmids were made unequivocally nic/bom-. In doing so the copy number and mobilisation properties of the pMTL plasmid were characterised and compared with the pUC plasmids.

RESULTS AND DISCUSSION

4.1 Construction of the pMTL backbone

In designing the strategy for the construction of the proposed vector a number of factors were taken into consideration. Firstly, it was desirable to remove any restriction sites from the
Fig. 4.1
The pUC Series of Plasmids

A schematic diagram of the pUC backbone, showing the linker regions and their location within the alpha-peptide.
vector backbone which were to be inserted into the linker region (i.e., AatII and NaeII). Secondly, investigation of why the copy number of the pUC plasmids is far higher than that of pAT153 (Chapter 3, Table 3.1). Lastly, the vector was designed to be non-mobilised (i.e., nic/bom$^-$), considered to be an asset in terms of biological containment for *in vitro* genetic manipulations of cloning vectors. The simplest way of achieving these aims was to essentially reconstruct the pUC backbone utilising pBR322 as a starting point. The final vector carries an identical deletion to that present in pAT153 (i.e. the two contiguous NaeII B and G fragments of pBR322, carrying the ROP and the nic/bom sites). The exact relationship between pBR322, pAT153 and pMTL and pUC backbones are described in Fig. 4.2. Fig. 4.3 describes schematically the construction of the pMTL backbone; pBR322 was double digested with BamHI - PvuII and the 3' recessed termini created by the digestion was filled in using the 5'-3' polymerase activity of DNA polymerase. The larger (1.69 kb) of the two fragments produced on digestion was isolated by agarose gel electrophoresis and blunt end ligated producing pMTL1. This step removes most of the ROP region, the remainder of the ROP region, with the nic/bom and AccI sites are removed as a NaeII fragment (0.43 kb) from pMTL1 creating pMTL2. Substitution of the SacI region containing the replication origin of pMTL2 with the equivalent region from pUC8 yields pMTL3. pMTL3 digested with AatII and blunted with T4 DNA polymerase exonuclease activity removing the AatII site from the vector backbone. Further digestion with EcoRI produces 3 fragments (1.57 kb, 0.44 kb, 0.24 kb) when resolved by agarose gel electrophoresis, the largest two were isolated and ligated to produce pMTL4. pMTL4 provides the backbone for all subsequent vectors produced in the pMTL series. The peptide and
Fig. 6.2

The Relationship Between pBR322, pAT153 and the pUC and pMTL Backbones

The region of pBR322 deleted to yield pAT153 indicated within the circle, the pBR322 DNA utilised in the construction of the pUC backbone is illustrated by a dashed line outside the circle, while the pBR322 DNA present in pMTL 4 is represented by the bold line. The Hae II site at which the M13 derived lacZ poly linker region was inserted into pMTL and pUC backbones is highlighted by an arrow.
Construction of the pML4 Backbone

The pML4 backbone was derived from pBR322 in the following manner (numbers in parenthesis refer to nucleotide position within pBR322; Balbas et al., 1986). A 1.693 kb region of pBR322 DNA was deleted by cleaving the plasmid with BamHI (375), blunt-ending with polk, cutting with PvuII (2066) and pML4 generated by self-ligation of the larger 2.67 kb plasmid fragment. This particular deletion removed the majority of rop but recreated a BamHI site. The remainder of the rop gene, together with the nic site was removed by deleting a 0.438 kb HaeII (232-2369) fragment. The plasmid obtained, pML2, had also lost an NdeI site (2297). The PstI and HindIII sites within the bla gene of pML2 were removed by substituting the 1.444 kb PstI fragment of the vector with the equivalent region of pUC8, yielding pML3. The final step required was to remove the AntII, EcoRI and HindIII sites of the vector. Although this would have been most easily achieved by deleting the DNA between the AntII (4286) and EcoRV (185) sites, we wished to retain the latter site for future manipulations. The plasmid was therefore cleaved with AntII, blunt-ended with T4 polymerase, cut with EcoRI, ligated and transformed into E. coli JM83. A recombinant plasmid, pML4, was selected which had lost 0.238 kb of DNA between the AntII site and the EcoRI (164) adjacent to the EcoRV site. To create the final vectors the c.540 bp HaeII fragment carrying the lacZ'/polylinker region was inserted into the indicated HaeII site, such that transcription from the lac promoter was from the same DNA strand as the Ap gene. The pML4P" vectors were constructed by inserting a 385 bp EcoRI fragment (blunt-ended with polk) into the EcoRV site as described in the text.
linker regions are then cloned into the Haell site as indicated in Fig. 4.2.

4.2 Derivation of the Novel Linker Regions

Fig. 4.4 shows the range of linkers produced and the restriction sites available.

4.2.1 pMTL 20/21

The initial vectors constructed represented merely an extension of the existing pUC18/19 linker region. Two complementary oligonucleotides were synthesised (Fig. 4.5). After annealing, the linker was cloned into M13mpl8/19 cleaved with SalI and PstI. Vector DNA carrying the synthetic linker were identified using the appropriate oligonucleotide as a radiolabelled probe and the sequence of the positive clones subsequently checked by nucleotide sequencing. The desired linker region, together with the alpha-peptide and lac P0, was isolated from M13 RF DNA as a 544 bp Haell fragment and inserted into pMTL4 at the indicated Haell site (Fig. 4.2). pMTL 20/21 plasmids obtained carried unique sites for AatI, MluI, NcoI, BglII, XhoI and SstI, in addition to those present in pUC18/19 (Fig. 4.6).

4.2.2 pMTL 22/23

In pMTL 22/23 the proposed new linker was synthesised as two complementary pairs of oligonucleotides (Fig. 4.7a/b). The two oligonucleotides of linker a were annealed and inserted into M13 mp 18/19 cleaved with EcoRI and HindIII. The resultant recombinant plaques were white in the presence of X-Gal. In addition, although the linker carries a HindIII sticky-end, the HindIII site was not recreated.
Fig. 4.4

The Multiple Cloning Sites of the pMTL Vectors and their Complementary M13 Vectors

Only those polylinker regions indicated are available within MTL vectors. It should be noted that the EcoRV sites of the relevant pMTL vectors is only unique in those derivatives carrying par. In addition, the BglII, ClaI, EcoRV and NaeI sites of the appropriate M13 mtl polylinkers are not unique.
Fig. 4.5

The Two Complementary Oligonucleotides used in the
Construction of the pML20 and pML21 Linkers

MluI BglII SphI
AatII NcoI XhoI

SalI 5'-TCGACGTCACGCTCCATGGAGATCTCGAGGCCTGCA-3'  PstI
Sticky-end 5'-GGCCTCGAGATCTCCATGGACGCGTGACG-3' Sticky-end
Identification of the orientation of HaeII fragment carrying the alpha-peptide and the linker region in pMTL4. A BglII digest resulted in the production of two fragments (0.88 and 1.59 kb) compatible with the correct orientation required. Also the range of unique sites within the linker are shown.

Lane 1 HindIII-EcoRI digest. Lane 2 pMTL20 - BglII digest (0.88, 1.59 kb). Lane 3-13 pMTL20, AccI, AatII, XbaI, NcoI, BglII, SacI, XhoI, PstI, SalI, MluI and StuI digests.
Fig. 4.7
The Two Pairs of Complementary Oligonucleotides used in the Construction of pMTL22 and pMTL23 Linkers.

A)

<table>
<thead>
<tr>
<th>Stui</th>
<th>BglII</th>
<th>NcoI</th>
<th>Smal</th>
</tr>
</thead>
<tbody>
<tr>
<td>NruI</td>
<td>XhoI</td>
<td>ClaI</td>
<td>SphI</td>
</tr>
</tbody>
</table>

5'-AGCTCGGAGGCTCGAGATCTATCGATGGCATGCCATGGTACCCGGGAGCTCG-3' 
3'-GGCGCTCCGGAGCTCTAGATAGCTACGGTACCATGGGCCCTCGAGCTTAA-5'

B)

<table>
<thead>
<tr>
<th>XbaI</th>
<th>PstI</th>
<th>SalI</th>
<th>NdeI</th>
</tr>
</thead>
<tbody>
<tr>
<td>EcoRI</td>
<td>HindIII</td>
<td>MluI</td>
<td>AatII</td>
</tr>
</tbody>
</table>

5'-AATTCTAGAAGCTTCGAGACACGCTCGACGCTCATATGGATCCGATATCGGCGAC-3' 
3'-GATCTCGAGAGCTGCTCGAGACGCTACGTTACCTAGGCTACGCGGCGCTTTAA-5
as the 5th base of the inserted oligonucleotide linker is C and not the T necessary to produce a HindIII site. RF DNA of recombinant M13 mpl8 and mpl9 was prepared, cleaved with EcoRI and the second linker ligated. The resultant recombinant plaques were blue in the presence of X-Gal. In this case, although the linker has an EcoRI sticky-end at both ends, and EcoRI site was not created at the extremity of the linker region and the nucleotide base at the 55th position was C rather than the G necessary to create an EcoRI site. During this cloning step it became apparent that the annealed oligonucleotides were capable of being inserted in either orientation. The M13 mpl8 derived vector was the only orientation capable of producing a functional alpha-peptide, resulting in the pMTL26 linker. The alternative orientation in the M13 mpl8 derived vector created translational stop codons in the linker regions resulting in a non-functional alpha-peptide. This method generated 3 novel linker regions, MTL22, 23 and 26, which were then isolated from M13 together with the alpha-peptide and the lac PO, as a 544 bp HaeIII fragment and inserted into pMTL4 at the indicated HaeIII site (Fig. 4.2) to yield pMTL 22, pMTL 23, pMTL26.

During dideoxy sequencing of cloned DNA, random subfragments are commonly generated by sonication and subsequently blunt-ended ligated into the Smal site of M13 vectors (Deininger, 1983). A requirement of this procedure is the circularisation of the fragment to be sequenced prior to fragmentation by sonication. The vectors pMTL22/23 should facilitate this process. DNA fragments cloned into the central portion of the linker region may be excised by double digestion with either BamHI/BglII, SalI/KhoI, CiaI/AciI, or combinations of NruI-Stul/EcoRV-Nael, and circularised by self ligation prior to sonication.
4.2.3 **pMTL24/25**

pMTL 24/25 the final pair of vectors constructed were designed to facilitate the conversion of a cloned DNA fragment from one type of sticky end to another. Ideally such a vector requires an inverted duplication of a particular linker region, as would occur when pMTL 20/21 linker regions were inverted and combined. In practice, the existence of such a lengthy palindromic sequence proved to be inherently unstable. To overcome this problem the vectors constructed contained the pMTL20/21 linker region followed by an inverted pUC18/19 linker region. Fig. 4.8 illustrates the procedure adopted to construct pMTL 24/25. To ensure that the final vectors had unique AatII sites in their linker regions, the linker region of pUC 18/19 was first transferred to the pMTL 4 backbone to give pMTL 18/19. In the case of pMTL24, a 0.8 kb Scal-HindIII fragment carrying the polylinker region and 5' end of the bla gene was isolated from pMTL19 and combined with a 2.02 kb Scal-HindIII fragment of pMTL20, which carries the 3' end of the bla gene, the origin of replication and the polylinker region. The plasmid obtained, pMTL24, carries both the pMTL19 and pMTL20 polylinker as adjacent insertions, but inverted relative to one another. A similar strategy was utilised in the derivation of pMTL25 except in this case Scal-EcoRI fragments of pMTL18 and pMTL21 were employed.

4.3 **Mobilisation of pMTL and pUC Vectors**

The Advisory Committee for Genetic Manipulation (ACGM) in the UK have assigned the pUC plasmids an access factor comparable to pBR322 (ACGM Secretariat, 1984). This is in contrast to pAT153 which is nic/bom (Twigg and Sherratt, 1980) and has consequently been assigned
Fig. 4.8
The Construction of pMTL 24 and pMTL 25

The diagram shows the construction of pMTL 24 and pMTL 25. The vectors pMTL18, pMLT19, pMLT20, pMLT21, pMLT22, and pMLT23 are depicted with EcoRI, HindIII, and Scal sites. The lacZ and OriG sites are also indicated.
a lower access factor. This categorisation is presumably made on the assumption that the pUC plasmids can be mobilised in the presence of trans-acting Mob proteins. The plasmid pBR322 lacks the genes coding for the proteins involved in mobilisation and is therefore inefficiently mobilised by large conjugal plasmids. The presence of the site at which these proteins act, the nic/bom site, enables pBR322 to be efficiently mobilised if the mobilisation proteins are supplied in trans i.e., from a co-resident plasmid such as ColK. This nic/bom site is known to lie adjacent to the AccI site at position 2245 (Covarrubias et al., 1981), although no published data appears to be available on the mobilisation frequency of the pUC plasmids. Yanisch-Perron et al. (1985) published nucleotide sequence of pUC18 revealed that the nic/bom site of the vector has been almost totally destroyed during the removal of the AccI site, using Bal31. This suggests that their mobilisation characteristics should closely resemble that of pAT153, which has had its nic/bom site completely removed.

Conjugations by filter mating (Minton et al., 1983b) were carried out to examine the mobilisation frequencies of different cloning vectors, using a Rif-resistant strain of W3110 as recipient and JC2926 transconjugants as donors (VanHaute et al., 1983). Table 4.1 describes the varying mobilisation frequencies of a range of cloning vectors. pUC and pMTL vectors have similar mobilisation frequencies comparable with pAT153, since all are functionally nic/bom- (i.e., non-mobilisable even in the presence of Mob proteins). Therefore, pUC and pMTL should have an ACGM access factor equivalent to pAT153.

4.4 Copy Number of pMTL and pUC Vectors

The copy number of ColEl and its derivative plasmid pBR322 is
## Table 4.1
**Mobilisation Frequencies of the Cloning Vectors**

<table>
<thead>
<tr>
<th>Plasmid(a)</th>
<th>Mobilisation frequency(b)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>from JC2926 R64drdII (pLVC9)</td>
</tr>
<tr>
<td>pBR322</td>
<td>0.9 x 10^-2</td>
</tr>
<tr>
<td>pAT153</td>
<td>5.1 x 10^-6</td>
</tr>
<tr>
<td>pBR328</td>
<td>2.1 x 10^-6</td>
</tr>
<tr>
<td>pUC8</td>
<td>4.2 x 10^-6</td>
</tr>
<tr>
<td>pUC9</td>
<td>6.7 x 10^-6</td>
</tr>
<tr>
<td>pMTL20</td>
<td>3.6 x 10^-6</td>
</tr>
<tr>
<td>pMTL21</td>
<td>4.1 x 10^-6</td>
</tr>
<tr>
<td>pMTL22</td>
<td>2.1 x 10^-6</td>
</tr>
<tr>
<td>pMTL23</td>
<td>5.8 x 10^-6</td>
</tr>
</tbody>
</table>

(a) Cloning vector mobilised

(b) Mobilisation frequency is expressed as frequency of ApR transconjugants per donor cell. Conjugations were carried out by a previously described filter mating technique (Minton et al., 1983b), using a RifR mutant of *E. coli* W3110. The donor strain JC2926 carries the conjugal mobiliser R64drdII and plasmid pLVC9. This latter plasmid carries the P15A replicon (Chang and Cohen, 1978), the CoIE1 mob region (but nic-), and encodes resistance to Ca.
negatively regulated by two different gene products. Regulation is mediated through both the RNAI molecule and the ROP polypeptide (Conrad and Campbell, 1979; Twigg and Sherratt, 1980). RNAI is a small untranslated RNA molecule encoded within the DNA region that is used to transcribe the RNA pre-primer, (RNAII) utilised as the primer in DNA replication (Cesaranii, 1982). Because RNAI is complementary to RNAII the two can hybridise, an event which interferes with RNAIIs role in initiation of replication. The ROP protein is believed to stabilise the interaction. The elevated copy number of pAT153 (150 copies per chromosome) compared to pBR322 (50 copies per chromosome) is explained on the basis of the deletion of DNA which encodes the ROP protein (Twigg and Sherratt, 1980). The pUC plasmids have also effectively deleted the rop gene, however copy number analysis has shown that the copy number of these vectors are considerably higher (500 - 700 copies per chromosome) than pAT153 (as stated in Chapter 3; Table 3.1).

The strategy utilised in the construction of pMTL 4 was designed to test the supposition that some other factor is contributing to the higher copy number of the pUC plasmids. In removing the PstI and HincII sites from pMTL2, a 1.44 kb TagI fragment carrying the RNAI/RNAII region was substituted with the equivalent region from pUC8 to give pMTL3 (Fig. 4.3). Copy number determination on cells carrying pMTL2 and pMTL3 (Table 4.2, Fig. 4.9) clearly indicate that the copy number of the plasmids carrying the pUC derived replicon (pMTL3-658 copies per chromosome) was significantly higher than the plasmids carrying the pBR322 derived replicon (pMTL2 - 197 copies per chromosome). The most obvious cause of such an increase would be a mutation within the RNAI/RNAII region.
Table 4.2
Calculated Copy Numbers from Plasmid/Chromosome Ratios, Obtained From Densitometry of EtBr-Stained Electrophoretic Gels

<table>
<thead>
<tr>
<th>Host</th>
<th>Genotype</th>
<th>Plasmid</th>
<th>Copy Number</th>
<th>% Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>(a)</td>
<td>(b)</td>
</tr>
<tr>
<td>JM83</td>
<td>-</td>
<td>pMTL2</td>
<td>197 ± 5.0</td>
<td>4</td>
</tr>
<tr>
<td>JM83</td>
<td>-</td>
<td>pMTL3</td>
<td>658 ± 9.5</td>
<td>4</td>
</tr>
<tr>
<td>GM271</td>
<td>dam-</td>
<td>pMTL2</td>
<td>113 ± 13.2</td>
<td>3</td>
</tr>
<tr>
<td>GM271</td>
<td>dam-</td>
<td>pMTL3</td>
<td>355 ± 15.1</td>
<td>3</td>
</tr>
<tr>
<td>GM2163</td>
<td>dam-, dam-</td>
<td>pMTL2</td>
<td>28 ± 4.0</td>
<td>3</td>
</tr>
<tr>
<td>GM2163</td>
<td>dam-, dam-</td>
<td>pMTL3</td>
<td>158 ± 16.4</td>
<td>3</td>
</tr>
<tr>
<td>GM2199</td>
<td>dam-</td>
<td>pMTL2</td>
<td>55 ± 8.6</td>
<td>3</td>
</tr>
<tr>
<td>GM2199</td>
<td>dam-</td>
<td>pMTL3</td>
<td>228 ± 12.2</td>
<td>3</td>
</tr>
<tr>
<td>GM161</td>
<td>dam-</td>
<td>pMTL2</td>
<td>30 ± 4.0</td>
<td>3</td>
</tr>
<tr>
<td>GM161</td>
<td>dam-</td>
<td>pMTL3</td>
<td>184 ± 11.1</td>
<td>3</td>
</tr>
</tbody>
</table>

Plasmid copy numbers were calculated per chromosome. Assuming one plasmid per chromosome equivalent exists and based on a figure of $3.8 \times 10^3$ Kb for the size of a non-replicating E. coli chromosome (Shepard and Polisky, 1979). The calculated copy numbers per chromosome equivalent are given in column (a), the standard deviations in column (b) and the number of independent trials in column (c).
Fig. 4.9

Whole Cell Lysates Run on an Agarose Gel

Identifying the variation in copy number between pMTL2 and pMTL3 in the host JM83.
4.5 Sequence Variation between the pBR322 and pUC8 RNAI/RNAII Regions

Having previously established that the increase in copy number of the pUC/pMTL plasmids is greater than that resulting from the deletion of rop as in pAT153, the RNAI/RNAII region, with numerous mutations already identified as increasing copy number (Davison, 1984) in particular was examined for sequence anomalies, along with the rest of pBR322 and pMTL4. The published nucleotide sequence of pUC18 (Yanish-Perron et al., 1985) implies no such difference from the pBR322 sequence (Sutcliffe, 1979).

Random fragments of pMTL4 and pBR322 were generated by sonication, cloned into M13 mp8 and sequenced. The templates carrying the RNAI/RNAII region after identification by sequencing were accurately sequenced using synthetic oligonucleotide primer 5'ATC-CCT-TAA-CGT-GAG-TT 3', annealing at position 1293-1310 downstream from the RNAI/RNAII region. Subsequent examination of the sequences showed variation in the nucleotide sequence between pMTL/pUC and pBR322 (Fig. 4.10). In pMTL4 a single G to A base change was discovered immediately preceding the 5' end of the RNAI region. There are a number of reasons why the single base change identified as the only variation in sequence between pBR322 and pMTL/pUC causes an increase in plasmid copy number. The simplest explanation would be that the level of RNA I transcription has been reduced, however as a complementary change has occurred within RNAII the possibility cannot be ruled out that it is the C to T base change within the RNAII transcript which causes the observed copy number increase. Although other studies have shown that DNA methylation can regulate DNA replication (Smith et al., 1985; van Putten et al., 1986), there appear to be no specific sequences present
DNA Sequencing Gel of the 3' Coding Sequence of the RNAI/RNAII Region in pBR322 and pMTL/pUC

Sequence of the non-coding strand 3'-5' is shown, identifying the single base difference between pBR322 and pMTL/pUC.
which are known to be acted upon by either adenosine and cytosine methyltransferases. This was confirmed by the observation that pMTL3 still exhibiting a 3-fold higher copy number than pMTL2, when both were examined in dam- and dcm- E. coli hosts (Table 4.2).

4.6 Construction of pMTL par

Having previously established in Chapter 3 that the par locus of pSC101 enhances the stability of cloning vectors, it was therefore decided to make par an integral part of any subsequent cloning or expression vectors constructed.

The strategy utilised in the construction of the pMTL backbone had ensured the existence of a unique EcoRV cloning site into which additional DNA fragments could be cloned without impairment of essential plasmid functions. The plasmid pMTL20 and pMTL21 were therefore cleaved with EcoRV and the par locus inserted as a blunt-ended 380 bp EcoRI fragment. The resultant vectors were designated pMTL20 par and pMTL21 par. Deletion derivatives of these two vectors were then obtained by deleting the 350 bp PvuII fragment carrying the polynucleo/LacZ region and the equivalent PvuII fragment from pMTL 18, 19, 22, 23, 24, 25 and 26 inserted to yield their par carrying equivalents. An important consequence of these manipulations was that by destroying the backbone EcoRV site, the EcoRV site of the pMTL22 par, 23 par and 26 par polylinkers were made unique.

4.6.1 Stability of pMTL20 par

The improved stability of the pMTL derivatives containing par was evaluated by comparing them with their pMTL equivalents without par. As with previous stability work, segregational stability was
qualitatively assessed in the steady state population provided by a chemostat. The stability of pMTL20 and pMTL20\textsubscript{par} was monitored in the host JM83 grown at a dilution rate of 0.1 hr\textsuperscript{-1} under carbon and phosphate nutrient limitation. Under carbon nutrient limitation no plasmid instability was observed in either pMTL20 or pMTL20\textsubscript{par}, over a 100 generation period, while under phosphate nutrient limitation pMTL20 was lost from the population after 15 generations, while pMTL20\textsubscript{par} exhibited no instability and was in the population for over a 100 generations (Fig. 4.11). These results confirm that the presence of the \textit{par} locus in pMTL20\textsubscript{par} improves its stability over pMTL20.
Fig. 4.11

Segregational Stability of pMTL20 par, (■), pMTL20 (□) and pUC8 (○) in the Host JMB3

Grown under phosphate limiting conditions pMTL20 par shown to be stable over 100 generations, while pUC8 and pMTL20 are lost after 15 generations.
CONCLUDING REMARKS

The construction of the pMTL cloning vectors provided a series of cloning vectors, which can be usefully employed in the generation of fragments with compatible sticky ends, required when M13 templates are produced for dideoxy sequencing by the sonication procedure.

Additionally, the construction and subsequent characterisation of the pMTL vectors examined features which were to be exploited in the high expression of CPG2. The high copy number of the pUC plasmids and the segregational stability exhibited by plasmids containing the par locus. Finally, the pUC and pMTL vectors were shown to be definitively nic/bom-.

An essential feature of the dideoxy nucleotide sequencing procedure is the prior generation of a set of M13 templates carrying random DNA fragments representative of the region to be sequenced. One particularly effective way of generating such templates is by the sonication procedure described by Deininger (1983). A crucial requirement of this procedure is that the fragment to be sequenced must be circularised by self-ligation prior to fragmentation. It follows that the DNA fragment must have compatible cohesive ends. The vectors pMTL22, pMTL23 and pMTL26 were specifically designed to aid in the conversion of DNA fragments with incompatible sticky-ends to fragments with cohesive termini. Thus, dependent on the site of insertion, cloned DNA may be excised by double digestion with either BamHI/BgllI, SalI/XhoI, Clal/AccI, or combinations of NruI-Stul/EcoRI-Nael (pMTL22 and pMTL23 only), and circularised by self-ligation. The vectors pMTL24 and pMTL25 present still further opportunities for ensuring that
DNA fragments possess self compatible cohesive ends. In both cases fragments which carry different termini compatible with any two of the six sites AatII, MluII, NcoI, BglII, XhoI and StuI may be cloned and subsequently excised by digestion with either EcoRI, SstI, KpnI, SmaI, BamHI, XbaI, Sall or AccI in the case of pMTL24, and with HindIII, SphI or PstI in the case of pMTL25. In addition DNA fragments generated by double digestion which possess one termini compatible with the above six sites and the other compatible with any of the additional sites present in the pMTL24/25 linker may also be converted to a restriction fragment with self-compatible cohesive ends. Thus in the case of pMTL24 the second cloning site used could be either SphI, PstI or HindIII with subsequent excision by digestion with either EcoRI, SstI, KpnI, SmaI, BamHI, XbaI, Sall or AccI, while the use of the above sites as secondary cloning or excision sites would be reversed in the case of pMTL25.

The availability of the vectors with and without par has important implications with regard to the overall size of the plasmid. Occasionally the excised cloned DNA fragment has an identical electrophoretic mobility to the linearised vector. This can present problems when preparing g quantities of the cloned DNA by electroelution, as is the case in the sonication procedure. In the majority of cases, if a cloned fragment comigrates with the linearised pMTL vector, then separation will be achieved with the corresponding pMTL par plasmid, or vice versa.

In addition to their use in the sonication procedure the pMTL vectors have a number of additional advantages over the pUC plasmids. The basic vectors are smaller than the pUC plasmids by 0.3 kb, while the pMTL par plasmids exhibit improved plasmid stability due to the
presence of the pSC101 par function. Both types of vector carry a
greater number of unique restriction enzyme cloning sites. In the
case of pMTL20/21 there are 6 additional sites to those present in
pUC18/19 (AatII, MluI, NcoI, BalII, XhoI and StuI), while pMTL22/23/26
carries 11 new cloning sites (AatII, MluI, NcoI, BalII, XhoI, StuI,
NaeI, EcoRV, ClaI, NdeI, and NruI). More importantly, the relative
order of the sites within these latter vectors has been rearranged.
This new order of sites will overcome, to a large extent, the limits
imposed by previously available polylinker regions on the choice of
restriction enzyme that can be used to excise DNA fragments. Aside
from their use in the sonication procedure, the pMTL24/25 vectors can
be employed to convert a particular restriction fragment from one type
of sticky end to another. From a safety point of view the pMTL vectors
are nic- and should therefore be awarded the lowest categorisation in
genetic manipulation experiments e.g. UK ACCM access factor of 10^-9.
In this respect we have also shown that the pUC plasmids are also
functionally nic-.

The strategy utilised in the construction of the pMTL vectors
has also identified the reason for the abnormally high copy number of
the pUC plasmids. There could be a number of reasons why the base
change identified causes an increase in copy number. The simplest
explanation would be that the level of RNA I transcription has been
reduced, however as a complementary change has occurred within RNA II
the possibility cannot be ruled out that it is the C to T base change
on the opposite strand which causes the observed copy number increase.
Other studies have shown that DNA methylation can control DNA
replication (Smith et al., 1985; van Putten et al., 1986). There
appear to be no specific sequences present which are known to be acted
upon by either adenosine or cytosine methyltransferases.

The series of pMTL cloning vectors constructed retain all the advantages of the pUC plasmids but are smaller (by 300 bp), and possess more versatile polylinker regions while the par containing variants exhibit improved plasmid stability. The pUC and pMTL plasmids have been shown to be functionally nic/bom, and on the basis of these findings the ACGM have now reassigned the access factor of the pUC plasmids with the pMTL to $10^{-9}$. The copy number variation between pAT153 and pUC plasmids has been identified as a single base change within their pBR322 derived replication regions which results in a 3-fold increase in copy number.

Finally, the pMTL cloning vectors were designed with the construction of a series of expression vectors as the final goal. pMTL 4 provided the starting point in the construction of the basic expression vector backbone. Also the new linker regions constructed for the pMTL cloning vectors aided the localisation of the various components of the expression vector as separate DNA restriction fragmenta "cassettes" which may be combined (by virtue of homologous sticky ends) in any number configurations in the backbone of the expression vector.
Fig. 4.12

Nucleotide Sequence of pMTL20 par

The total length of pMTL 20 par is 2854 nucleotides; in the print out of the sequence, the relevant restriction sites, the origin of the RNA1/KNAII replication regions, the par locus and the promoter regions of beta-lactamase and beta-galactosidase are indicated.

```
NCCTGCTCGCTCGCTCGGCTGCGCCCACCGGTATCACCTCACTCAAACGCGGTAATACGGTTATCCACAGAATCACGCCATAACCCAGGAAAGAACATG
```

```
RNA II
```

```
TCCTCGCTCTGCTGAAGCCACTTACCTTCCAAAAAAGAGTTCOTACCTCTTGATCCCGCAAACAAACCArCGCTGGTACrGGTGCTTTT
```

```
GCTTAAOCOATTTTGGTCATGACATTATCAAAAACOATCTTCACCTAGATCCTTTTAAATTAAAAATOAAOTTTTAAATCAATCTAAAI'.TATATATGAr.TA I
```

```
CTTAAOCOATTTTGGTCATGACATTATCAAAAACOATCTTCACCTAGATCCTTTTAAATTAAAAATOAAOTTTTAAATCAATCTAAAI'.TATATATGAr.TA I
```

```
TCCTGCTCGCTCGCTCGGCTGCGCCCACCGGTATCACCTCACTCAAACGCGGTAATACGGTTATCCACAGAATCACGCCATAACCCAGGAAAGAACATG
```

```
RNA II
```

```
TCCTCGCTCTGCTGAAGCCACTTACCTTCCAAAAAAGAGTTCOTACCTCTTGATCCCGCAAACAAACCArCGCTGGTACrGGTGCTTTT
```

```
GCTTAAOCOATTTTGGTCATGACATTATCAAAAACOATCTTCACCTAGATCCTTTTAAATTAAAAATOAAOTTTTAAATCAATCTAAAI'.TATATATGAr.TA I
```

```
CTTAAOCOATTTTGGTCATGACATTATCAAAAACOATCTTCACCTAGATCCTTTTAAATTAAAAATOAAOTTTTAAATCAATCTAAAI'.TATATATGAr.TA I
```

```
TCCTGCTCGCTCGCTCGGCTGCGCCCACCGGTATCACCTCACTCAAACGCGGTAATACGGTTATCCACAGAATCACGCCATAACCCAGGAAAGAACATG
```

```
RNA II
```

```
TCCTCGCTCTGCTGAAGCCACTTACCTTCCAAAAAAGAGTTCOTACCTCTTGATCCCGCAAACAAACCArCGCTGGTACrGGTGCTTTT
```

```
GCTTAAOCOATTTTGGTCATGACATTATCAAAAACOATCTTCACCTAGATCCTTTTAAATTAAAAATOAAOTTTTAAATCAATCTAAAI'.TATATATGAr.TA I
```

```
CTTAAOCOATTTTGGTCATGACATTATCAAAAACOATCTTCACCTAGATCCTTTTAAATTAAAAATOAAOTTTTAAATCAATCTAAAI'.TATATATGAr.TA I
```

```
TCCTGCTCGCTCGCTCGGCTGCGCCCACCGGTATCACCTCACTCAAACGCGGTAATACGGTTATCCACAGAATCACGCCATAACCCAGGAAAGAACATG
```

```
RNA II
```

```
TCCTCGCTCTGCTGAAGCCACTTACCTTCCAAAAAAGAGTTCOTACCTCTTGATCCCGCAAACAAACCArCGCTGGTACrGGTGCTTTT
```

```
GCTTAAOCOATTTTGGTCATGACATTATCAAAAACOATCTTCACCTAGATCCTTTTAAATTAAAAATOAAOTTTTAAATCAATCTAAAI'.TATATATGAr.TA I
```

```
CTTAAOCOATTTTGGTCATGACATTATCAAAAACOATCTTCACCTAGATCCTTTTAAATTAAAAATOAAOTTTTAAATCAATCTAAAI'.TATATATGAr.TA I
```

```
TCCTGCTCGCTCGCTCGGCTGCGCCCACCGGTATCACCTCACTCAAACGCGGTAATACGGTTATCCACAGAATCACGCCATAACCCAGGAAAGAACATG
```

```
RNA II
```

```
TCCTCGCTCTGCTGAAGCCACTTACCTTCCAAAAAAGAGTTCOTACCTCTTGATCCCGCAAACAAACCArCGCTGGTACrGGTGCTTTT
```

```
GCTTAAOCOATTTTGGTCATGACATTATCAAAAACOATCTTCACCTAGATCCTTTTAAATTAAAAATOAAOTTTTAAATCAATCTAAAI'.TATATATGAr.TA I
```

```
CTTAAOCOATTTTGGTCATGACATTATCAAAAACOATCTTCACCTAGATCCTTTTAAATTAAAAATOAAOTTTTAAATCAATCTAAAI'.TATATATGAr.TA I
```

```
TCCTGCTCGCTCGCTCGGCTGCGCCCACCGGTATCACCTCACTCAAACGCGGTAATACGGTTATCCACAGAATCACGCCATAACCCAGGAAAGAACATG
```

```
RNA II
```

```
TCCTCGCTCTGCTGAAGCCACTTACCTTCCAAAAAAGAGTTCOTACCTCTTGATCCCGCAAACAAACCArCGCTGGTACrGGTGCTTTT
```

```
GCTTAAOCOATTTTGGTCATGACATTATCAAAAACOATCTTCACCTAGATCCTTTTAAATTAAAAATOAAOTTTTAAATCAATCTAAAI'.TATATATGAr.TA I
```

```
CTTAAOCOATTTTGGTCATGACATTATCAAAAACOATCTTCACCTAGATCCTTTTAAATTAAAAATOAAOTTTTAAATCAATCTAAAI'.TATATATGAr.TA I
```
CHAPTER 5
EXAMINATION OF $P_L$ PROMOTER SYSTEMS
INTRODUCTION

The expression vector pMM21, previously used to express the *cub* gene in *E. coli* employed the plasmid pAT153. A 4.8 kb *Bul* I fragment carrying the *cub* gene was cloned into the BamHI site of pAT153. The level of expression obtained was of the order of 1-2.5% of the cell soluble protein. This relatively high level of expression (compared with the native promoter, <0.1%) was due to the transcriptional readthrough from the promoter of the upstream *tet* gene plus the *cub* promoter (Minton et al., 1983a).

It should be possible to increase the level of *CGP2* expression by replacing the relatively weak *tet* promoter with a 'stronger' promoter, such as that of the *E. coli* lac operon. Accordingly, a 2.03 kb *Bam* HI fragment carrying the *cub* gene was therefore cloned into the *Bam* HI site of pUC8, giving pKM830. The lac promoter, however, proved difficult to regulate in the resultant clones. Complete repression was not possible when employing the lacI*G* gene in the host chromosome, the high copy number of the vector, and therefore the lac operator site, titrates out the effectiveness of the repressor molecule (Backman et al., 1976). However, complete repression was possible when the lacI*G* gene was carried on a compatible plasmid (Fig.5.1.)

Unregulated overexpression of *CGP2* drastically affected the physiological characteristics of the host cell, manifested as a reduction in cell viability, growth rate and a tendency for plasmid instability; all previously reported as the detrimental consequences of overexpression (Caulcott et al., 1985; Garrett et al., 1981). In the case of *CGP2* this was predictable since high levels of this protein produced "constitutively" or in a non-regulated manner would deplete
Fig. 5.1

Expression of CPG2 by pNM830, when regulated by the compatible plasmid pNM52 (pACYC184+ lacIq ) in the host JM103.

Δ represents expression while catabolite repressed.
□ " " " " " " with IPTG.
▲ " " when not catabolite repressed.
■ " " " " " " with IPTG.
the cell's folic acid levels primarily therefore affecting DHA synthesis as a result of the lack of methyl folate.

This inability to achieve tight regulatory control of metabolic promoters, such as lac and trp, on multicopy plasmids (O'Farrell et al., 1978; Lee and Bailey, 1984) has resulted in them being superseded by temperature-controlled promoter systems (Remaut et al., 1981; Wright et al., 1986). Temperature-controlled systems mainly utilize the powerful transcriptional promoters obtained from bacteriophage lambda P_l or P_r controlled by the temperature-sensitive repressor gene product cl. This gene product denatures at temperatures above 37°C providing a mechanism by which the promoter can be switched on, by shifting the temperature from 28°C to 42°C. This method of induction was preferable to induction of the lac promoter using the lactose analogue IPTG. Even if tight regulation of the lac promoter was possible, the high cost of the gratuitous inducer IPTG would make its use prohibitive on the large scale.
RESULTS AND DISCUSSION

5.1 Expression Vector Construction Examined

Two basic vectors are examined in this Chapter; pPlc28, a Pl inducible expression vector derived from pAT153 and constructed by Remaut et al. (1981). The other pNM501, derived from pUC8 and constructed by N. Minton. pNM501 was obtained by isolating the Pl promoter from pPlc28 as a HaeII fragment, which was subsequently cleaved with Sau3A and the resultant sub-fragment carrying Pl cloned into M13mp7. Pl was then cloned as a 256bp EcoRI fragment into pUC8 to give pNM501, this resulted in the Pl promoter running in tandem with the lac promoter (lac-Pl). Remaut et al. (1981) had previously expressed 40% of the cells soluble protein as beta-galactosidase when directed by Pl in a vector with a copy number similar to pAT153. It was anticipated that increased levels of expression would be obtained utilising a vector with a higher copy number.

The cpg gene was cloned as a 2.03kb BamHI fragment into pPlc28 and pNM501 to give pNM232 and pNM502 respectively. The designations Pla and Plc (Pl anti-clockwise and clockwise) refer to the direction of transcription from the cloned Pl promoter relative to the orientation of the origin of replication. pNM232 derived from pPlc28 carries its Pl promoter in the opposite direction to pNM502 (Fig. 5.2a/b).

The cloning of the cpg gene, with its own promoter, into pNM501 downstream of the lac and Pl promoters results in all 3 promoters running in tandem. Previous vectors where multiple promoters have been constructed have not always led to a combined level of expression equal to the sum of the expression due to the individual
Schematic Diagram of (a) pNM232 and (b) pNM502

Showing the position and orientation of the cpg gene and the $P_L$ promoter within the vector.
promoters. deBoer et al. (1982) produced a trp-lac tandem promoter. It appeared from the expression data obtained that the frequency of transcriptional initiation was determined solely by the downstream lac promoter irrespective of the number of RNA polymerase molecules that had started at the upstream trp, and entered the lac promoter region. Transcription of the entire lac operator resulted in a messenger RNA with a large hairpin structure (G = -22 Kcal/mol). Such structures in RNA have been shown to cause pausing of the RNA polymerase that subsequently can lead to premature termination of transcription. Examination of the cpg promoter (Minton and Clarke, 1985) and the P_1 promoter (Remaut et al., 1981) sequences revealed no secondary structure which would inhibit transcriptional readthrough. The expression work carried out by Minton et al. (1983a) indicated no termination of transcription within the cpg promoter, when the tet promoter was used, to express CPG_2 in pNM21.

The expression work carried out on pNM830, in cultures with and without glucose revealed that CPG_2 expression was subject to catabolite repression, indicating the probable effect of the lac promoter upstream of the P_1 promoter, in pNM502, could have on CPG_2 expression. Therefore, to obtain an accurate measure of the contribution of the P_1 promoter to the expression of CPG_2, the cultures carrying pNM502 were supplemented with glucose to switch off the lac promoter. The contribution of the cpg promoter made towards expression was considered negligible (less than 0.2% of the cells soluble protein expressed as CPG_2).

5.2 Host Strains for Induction

The P_1 plasmids constructed do not contain the cl gene coding
for the repressor. Transcription from the $P_l$ promoter must therefore be repressed by maintaining the plasmid in an *E. coli* strain that can synthesize repressor from a chromosomal $cI$ gene. The strain K-12AH1Atrp (Castellazzi et al., 1972; Bernard et al., 1979) harbours a defective, non-excisable lambda prophage carrying a mutant $cI$ gene that codes for a temperature-sensitive repressor ($cI857$). This allows transcription from the $P_l$ promoter to be activated by adjusting the temperature from 28°C to 42°C. K12AH1Atrp is deleted for the $cro$ gene. Secondary repression of $P_l$-mediated expression by the $cro$ gene product is therefore avoided (Takeda et al., 1977), also the K12AH1Atrp lysogen does not express a functional $N$-gene product ($Nom-7am53$). The alternative to using K12AH1Atrp was to employ an *E. coli* host with a dual plasmid system operating; two compatible plasmids, one containing the $P_l$ promoter, while the other carrying the $cI$ repressor gene. Another possibility was to clone the $cI$ gene onto the plasmid carrying the $P_l$ promoter.

5.3  $P_l$ Directed Expression of CPC$_2$ in the Host K12AH1Atrp

Induced expression of CPC$_2$ was examined through a series of shake flask experiments. Cultures of K12AH1Atrp containing the expression vectors under investigation, pNM232 and pNM502, were grown in 2 x TY (supplemented with 1% glucose, to repress the lac promoter, and ampicillin, for plasmid maintenance). 1 l batch cultures at 28°C, were grown in a 2 l shake flask, shaken at 200 rpm on a bench top orbital shaker, to an OD$_{450}$ nm of 0.4. At this point the culture was divided, half raised to 42°C, the remainder maintained at 28°C. The cultures were maintained at these temperatures while samples were removed hourly for analysis. The results of the enzyme assays from
these shake flask experiments are described in Fig. 5.3, and show an increase in CPG2 expression on induction. The level of expression obtained (of the order of 2% of the cells soluble protein), however, was not comparable with the 30% and 40% levels of expression achieved by Remaut et al. (1981) with beta-lactamase and tryptophan synthetase A respectively. Also, CPG2 was being expressed prior to induction, possibly due to the cpg promoter, or incomplete repression of the lac or PL promoters.

Subsequently a number of 5 l batch cultures were used to examine more fully the expression of CPG2, pre- and post-induction, employing the same media and growth conditions as in the shake flask experiments. These cultures were carried out in LH 500 series fermenters, which provided a more accurate means of induction due to the temperature control available, compared to the shake flask experiment.

Samples were removed every hour for analysis. The results obtained for CPG2 expression were similar to those previously obtained in shake flasks. CPG2 was produced throughout the culture, with temperature induction only slightly increasing its expression. The results obtained are described in Fig. 5.4. Beta-lactamase expression was also monitored in these cultures; its expression appeared to parallel that of CPG2. A possible explanation for beta-lactamase and CPG2 expression running concurrent could be provided for pNM502. The induction of the PL promoter upstream of the beta-lactamase gene (bla) could jointly affect CPG2 and beta-lactamase expression. Unfortunately no such scenario could be envisaged for pNM232 where the PL promoter and the bla gene are on opposite strands. An alternative explanation could be that the parallel levels of expression achieved by CPG2 and
Expression of $\text{CPG}_2$ by pNM232 and pNM502 in the Host $\text{K12}\Delta\text{HI}\Delta\text{trp}$ carried out in 1 L Batch Cultures

The $\text{CPG}_2$ activity attained by pNM232 and pNM502 induced are represented by ■ and ● respectively. Analogous values for uninduced cells are represented by □ and ○. Arrow represents point of induction.
Expression of CPG2 by pNM232 and pNM502 in the Host K12 H1 trn

Carried out in 5 L Batch Cultures

The CPG2 activity attained by pNM232 and pNM502 induced are represented by ■ and ● respectively. Analogous values for uninduced cells are represented by □ and ○. Arrow represents the point of induction.
beta-lactamase are not the result of any directed promoter activity, but simply a gene dosage response. Unlin and Nordstrom (1978) had previously related beta-lactamase activity directly to copy number. Therefore the slight increase in expression on induction could be a result of an increase in copy number.

Also, following induction of both constructions, beta-lactamase levels in the culture medium increased, suggesting that the normally cell-bound beta-lactamase was being released. This, coupled with the observation that cell viability falls rapidly following induction, suggests that cell lysis may be occurring.

5.3.1 Effect of Copy Number on Expression

Analysis pre- and post-induction of both cultures shows a dramatic increase in the copy number of both vectors immediately following induction (Fig. 5.5). Following induction the copy number of pNM502 increased from 150 to 700, while that of pNM232 rose from 50 to 116. In both cases the elevated copy number was maintained over the period of induction.

Although Remaut et al. (1981) did not carry out any copy number analysis on the original vectors, Siegel and Ryu (1985), working with similar $P_l$ plasmid constructions, observed an increase in plasmid content associated with a reduction in growth rate following induction. They also showed that the host strain, K12 ΔH1Δtrp, containing the $P_l$ vectors exhibited a slower deceleration of specific growth rate and plasmid loss from the population under the derepressed condition compared to the host containing $P_{M15}$ vectors. This is in agreement with the observation of Remaut et al. (1981) that the anti-parallel "a" configuration of the operator, with respect to the origin, has an
Copy number of pNM232 and pNM502 pre-induction are represented by and respectively, while the copy numbers of pNM232 and pNM502 post-induction are represented by ■ and ○ respectively. Arrow represents point of induction.
adverse effect on protein synthesis. The host strain, K12ΔH1Δtrp containing the vectors under examination also exhibited a reduction in growth rate upon induction (Table 5.1) with the host containing pNM232, (a pPlc orientation) exhibiting a slower deceleration of growth rate compared to the host containing pNM502, (pPla orientation) following derepression. Stueber and Bujnard (1982) have demonstrated that an inverse correlation between copy number and growth rate exists. The increase in copy number following the temperature induction of Pp, expression could therefore be explained in terms of a decrease in the growth rate. This would be caused by the increased production of the Pp, promoted gene product, cumulating in the observed cell lysis and release of beta-lactamase into the culture medium which accompanied induction. The elevated copy number of pNM502 following induction could also arise from transcription readthrough from the Pl promoter to RNAII, which would produce an increase in copy number.

The increase in copy number and gene dosage accompanying induction, prevents an accurate measure of the induced Pl promoter activity being obtained. McKenney et al. (1982) attempted to eliminate the problem when comparing promoters with different transcriptional activities on different multicopy plasmids. By relating specific activity to copy number, a measure of the promoter strength of different plasmid constructions exhibiting different copy numbers could be obtained. Applying the same relationship to the varying levels of expression and copy numbers of the vectors under investigation it was possible to show that the relative promoter strength did not increase with induction (Table 5.2). Therefore, the slight increase in CPG2 activity could be solely attributed to the increase in copy number and
Table 5.1
Variation in Growth Rates and Copy Number Pre- and Post-Induction of
pNM232 and pNM502 in the Host K12AH1Atrp.

Growth rate determined by the rate of change of optical density (hr⁻¹) during batch growth. Copy number values used in table were mean values taken during the pre and post-induction phases.

<table>
<thead>
<tr>
<th></th>
<th>Pre-induction</th>
<th>Post-induction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth rate</td>
<td>Copy Number</td>
<td>Growth rate</td>
</tr>
<tr>
<td>pNM232</td>
<td>0.81</td>
<td>0.40</td>
</tr>
<tr>
<td>pNM502</td>
<td>0.77</td>
<td>0.25</td>
</tr>
</tbody>
</table>
Table 5.2

Variation in Expression and Relative Promoter Strengths of
P\text{\textsubscript{J}} Promoter in the Constructions and Hosts under Examination

<table>
<thead>
<tr>
<th>HOST</th>
<th>PLASMID</th>
<th>GROWTH TEMPERATURE</th>
<th>CPG\text{\textsubscript{2}} ACTIVITY U mg\textsuperscript{-1}</th>
<th>COPY NUMBER</th>
<th>PROMOTER ACTIVITY</th>
<th>RELATIVE PROMOTER STRENGTH</th>
<th>% SOLUBLE PROTEIN OBTAINED ON INDUCTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>K\textsubscript{1} \text{2ΔAαβrp}</td>
<td>pNM232</td>
<td>28°C</td>
<td>0.362</td>
<td>50</td>
<td>0.007</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>K\textsubscript{1} \text{2ΔAαβrp}</td>
<td>pNM232</td>
<td>42°C</td>
<td>1.779</td>
<td>116</td>
<td>0.015</td>
<td>2.1</td>
<td>2.1</td>
</tr>
<tr>
<td>K\textsubscript{1} \text{2ΔAαβrp}</td>
<td>pNM502</td>
<td>28°C</td>
<td>1.281</td>
<td>150</td>
<td>0.008</td>
<td>1.1</td>
<td>-</td>
</tr>
<tr>
<td>K\textsubscript{1} \text{2ΔAαβrp}</td>
<td>pNM502</td>
<td>42°C</td>
<td>6.281</td>
<td>750</td>
<td>0.008</td>
<td>1.1</td>
<td>2.2</td>
</tr>
<tr>
<td>W5445/pc1</td>
<td>pNM232</td>
<td>28°C</td>
<td>0.200</td>
<td>100</td>
<td>0.002</td>
<td>1.0</td>
<td>-</td>
</tr>
<tr>
<td>W5445/pc1</td>
<td>pNM232</td>
<td>42°C</td>
<td>8.33</td>
<td>120</td>
<td>0.008</td>
<td>4.0</td>
<td>1.9</td>
</tr>
<tr>
<td>W5445/pc1</td>
<td>pNM502</td>
<td>28°C</td>
<td>7.1</td>
<td>400</td>
<td>0.017</td>
<td>2.4</td>
<td>-</td>
</tr>
<tr>
<td>W5445/pc1</td>
<td>pNM502</td>
<td>42°C</td>
<td>29.1</td>
<td>460</td>
<td>0.064</td>
<td>9.1</td>
<td>7.4</td>
</tr>
<tr>
<td>W5445</td>
<td>pNM502c1</td>
<td>28°C</td>
<td>2.5</td>
<td>400</td>
<td>0.006</td>
<td>0.9</td>
<td>-</td>
</tr>
<tr>
<td>W5445</td>
<td>pNM502c1</td>
<td>42°C</td>
<td>84.9</td>
<td>400</td>
<td>0.213</td>
<td>30</td>
<td>17.1</td>
</tr>
<tr>
<td>W5445</td>
<td>pCM10</td>
<td>28°C</td>
<td>2.0</td>
<td>395</td>
<td>0.005</td>
<td>0.7</td>
<td>-</td>
</tr>
<tr>
<td>W5445</td>
<td>pCM10</td>
<td>42°C</td>
<td>85.2</td>
<td>400</td>
<td>0.213</td>
<td>30.4</td>
<td>17.7</td>
</tr>
</tbody>
</table>
5.4 P_l Directed Expression of CPG_2 in the host W5445

Further work on the P_l induced expression of CPG_2 was carried out employing a dual plasmid system in the host W5445. In this system the cl repressor gene was carried on pACYC177 (designated pcI857), which is compatible with the ColEl derived vectors under examination. Shake flask and 5 l batch cultures identical to those previously carried out (in Section 5.3) were used to compare the levels of expression obtained with these vectors in W5445 to those previously obtained in the host K12ΔH1 Δtrp. The results obtained are described in Fig. 5.6. The results show that pNM502, post-induction produces higher levels of expression, 7% of the cells soluble protein expressed as CPG_2 in W5445, compared to the 2% expressed in K12ΔH1Δtrp. The level of expression of CPG_2 produced by pNM232 in W5445 was 3-fold lower than that expressed by pNM502, consistent with the copy number variation between the two plasmids. The level of expression directed by pNM232 in W5445 (2% of the cells soluble protein expressed as CPG_2), was comparable with the level of expression obtained in K12ΔH1Δtrp.

Relatively high levels of CPG_2 expression were exhibited prior to induction, indicating inefficient repression of the P_l promoter by the cl gene product. This inefficiency can be attributed to the high copy number of pNM232 and pNM502. pcI857, derived from pACYC177 has a copy number of 20 (Chang and Cohen, 1978) and therefore is unlikely to produce enough cl repressor. This contradicts Remaut et al. (1981) claim that a single copy of the cl gene produces enough repressor to completely repress the P_l promoter, even when present on a multicopy plasmid.
Expression of CPG$_2$ by pNM232 and pNH502 in the Host W5445

Carried out in 5 L Batch Cultures

The CPG$_2$ activity attained by pNM232 and pNH502 induced are represented by ■ and □ respectively. Analogous values for uninduced cells are represented by ○ and □. Arrow represents point of induction.
Analysis of the copy number of pNM502 and pNM232 in W5445, pre and post-induction revealed, in contrast to K12 ΔH1Δtrp, only a slight increase in copy number. Also, measurement of CPG2 expression in terms of promoter strength indicated that the increase in CPG2 expression on induction was due to an increase in promoter strength and not as a consequence of a gene dosage (Table 5.2).

5.4.1 What Effect does Cloning the cl Gene into pNM502 have on Induced Expression?

The combination of pNM502 in the host W5445 had provided the highest expression levels of all the vector systems examined. The combination of a strong promoter on a high copy number vector resulted in about 10% of the cell soluble protein being expressed as CPG2. This level of expression, although high compared to the other constructions available, is not as high as expected. The PL promoters had previously been used to express heterologous protein at levels of 30 to 40% soluble protein (Remaut et al., 1981). One possible explanation for the relatively low levels of CPG2 expressed could be attributed to the PL promoter not being fully repressed. CPG2 expression taking place throughout growth, induction only slightly increasing the level of expression. Such constitutive expression would be detrimental to the host, resulting in poor expression. The PL promoter-operator located on the high copy number vector is repressed by the cl gene product located on the low copy number plasmid pC1857. The high copy number of the vector is probably sufficient to titrate out the repressor molecule, resulting in constitutive expression from the PL promoter.

Tight regulation of the PL promoter could be achieved by
cloning the cl gene onto the expression vector itself, thereby providing sufficient cl gene product to fully repress the PL promoter. The cl gene was isolated from pC1857 as a 0.94 kb TaqI fragment, and cloned into the Accl site of pKMS02. The resulting plasmid designated pKMS02cl should provide enough repressor molecules to repress the downstream PL promoter when maintained at 28°C. When pKMS02cl was used in conjunction with the strain W5445, reduced levels of CMP activity were recorded, compared to the dual plasmid system, prior to induction. On raising the temperature from the permissive 28°C to 42°C, induction of CMP activity occurs. The difference in the level of expression obtained was 17% of the cells soluble protein expressed as CMP, for the single plasmid system compared to 7% obtained in the dual plasmid system. The difference appears to be due to the effectiveness of the repression when the cl gene co-exists on the same plasmid with PL. The results obtained from this experiment are described in Fig. 5.7.

5.5 \( P_L \) Directed Expression of Chloramphenicol Acetyl-Transferase (CAT) in the Host W5445

The need for CMP for clinical trials in tumour therapy had led to its choice as the major gene product within the laboratory to be expressed in a high expression vector. As a model system, it is a poor choice. The codon utilisation of the cmp gene (Table 5.3) is unusual and predictably influenced by the high G-C content (67%) of the UNA. The bias towards G or C nucleotide is most marked in the wobble position where there is a 93% preference for codons ending in a G or C. Grosjean et al. (1978) have described how optimal expression of genes in enterobacteria is achieved by the selection of codon on the basis of an optimal codon-anticodon interaction energy. The observed
Expression of CPG₂ by pNM502 cl in the Host W5445

Carried out in 5 L Batch Cultures

The CPG₂ activity attained by pNM502 cl induced are represented by ●. Analogous values for uninduced cells are represented by ○. Arrow represents point of induction.
<table>
<thead>
<tr>
<th>Codon</th>
<th>aa</th>
<th>Codon</th>
<th>aa</th>
<th>Codon</th>
<th>aa</th>
<th>Codon</th>
<th>aa</th>
<th>Codon</th>
<th>aa</th>
<th>Codon</th>
<th>aa</th>
<th>Codon</th>
<th>aa</th>
<th>Codon</th>
<th>aa</th>
<th>Codon</th>
<th>aa</th>
<th>Codon</th>
<th>aa</th>
<th>Codon</th>
<th>aa</th>
</tr>
</thead>
<tbody>
<tr>
<td>UAU</td>
<td>Tyr</td>
<td>UGU</td>
<td>Cys</td>
<td>UUU</td>
<td>Phe</td>
<td>UCU</td>
<td>Ser</td>
<td>UCC</td>
<td>Ser</td>
<td>UCA</td>
<td>Leu</td>
<td>UCG</td>
<td>Leu</td>
<td>UAA</td>
<td>Leu</td>
<td>UAG</td>
<td>Leu</td>
<td>UAC</td>
<td>Leu</td>
<td>UAA</td>
<td>Leu</td>
</tr>
<tr>
<td>UAC</td>
<td>13</td>
<td>9</td>
<td>98</td>
<td>65</td>
<td>UGC</td>
<td>1</td>
<td>23</td>
<td>39</td>
<td>UUC</td>
<td>12</td>
<td>11</td>
<td>113</td>
<td>102</td>
<td>UCC</td>
<td>Ser</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UAA</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>UGA</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>UUA</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>71</td>
<td>UCA</td>
<td>Leu</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UAG</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>UGG</td>
<td>Trp</td>
<td>1</td>
<td>5</td>
<td>25</td>
<td>66</td>
<td>UUG</td>
<td>0</td>
<td>5</td>
<td>16</td>
<td>64</td>
<td>UCG</td>
<td>Leu</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAU</td>
<td>His</td>
<td>1</td>
<td>5</td>
<td>19</td>
<td>95</td>
<td>CGU</td>
<td>0</td>
<td>5</td>
<td>223</td>
<td>99</td>
<td>CUU</td>
<td>0</td>
<td>1</td>
<td>26</td>
<td>73</td>
<td>CCC</td>
<td>Pro</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAC</td>
<td>4</td>
<td>10</td>
<td>75</td>
<td>59</td>
<td>CCG</td>
<td>Arg</td>
<td>18</td>
<td>9</td>
<td>101</td>
<td>133</td>
<td>CUC</td>
<td>10</td>
<td>8</td>
<td>33</td>
<td>69</td>
<td>CCC</td>
<td>Pro</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAA</td>
<td>Gln</td>
<td>0</td>
<td>1</td>
<td>38</td>
<td>90</td>
<td>CGA</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>27</td>
<td>CUA</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>22</td>
<td>CCA</td>
<td>Pro</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAG</td>
<td>7</td>
<td>6</td>
<td>169</td>
<td>166</td>
<td>CCG</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>42</td>
<td>CUG</td>
<td>30</td>
<td>18</td>
<td>345</td>
<td>294</td>
<td>CCG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAU</td>
<td>Asn</td>
<td>1</td>
<td>3</td>
<td>13</td>
<td>101</td>
<td>AGU</td>
<td>Ser</td>
<td>0</td>
<td>2</td>
<td>10</td>
<td>56</td>
<td>AAU</td>
<td>2</td>
<td>2</td>
<td>67</td>
<td>156</td>
<td>ACU</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAC</td>
<td>17</td>
<td>5</td>
<td>159</td>
<td>98</td>
<td>ACC</td>
<td>Arg</td>
<td>6</td>
<td>1</td>
<td>49</td>
<td>61</td>
<td>AUC</td>
<td>1</td>
<td>21</td>
<td>5</td>
<td>262</td>
<td>118</td>
<td>ACC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAA</td>
<td>Lys</td>
<td>1</td>
<td>5</td>
<td>259</td>
<td>163</td>
<td>AGA</td>
<td>Arg</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>28</td>
<td>AUA</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>27</td>
<td>ACA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAG</td>
<td>31</td>
<td>9</td>
<td>106</td>
<td>44</td>
<td>AGG</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>17</td>
<td>AUG</td>
<td>Met</td>
<td>7</td>
<td>10</td>
<td>140</td>
<td>130</td>
<td>ACC</td>
<td>Thr</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAU</td>
<td>Asp</td>
<td>1</td>
<td>8</td>
<td>116</td>
<td>183</td>
<td>GGU</td>
<td>Gly</td>
<td>2</td>
<td>7</td>
<td>226</td>
<td>124</td>
<td>GUU</td>
<td>0</td>
<td>2</td>
<td>192</td>
<td>108</td>
<td>GCU</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAC</td>
<td>24</td>
<td>19</td>
<td>204</td>
<td>106</td>
<td>GCC</td>
<td>Gly</td>
<td>43</td>
<td>12</td>
<td>174</td>
<td>140</td>
<td>GUC</td>
<td>Val</td>
<td>10</td>
<td>4</td>
<td>41</td>
<td>66</td>
<td>GCA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAA</td>
<td>Glu</td>
<td>10</td>
<td>11</td>
<td>373</td>
<td>210</td>
<td>GGA</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>42</td>
<td>GUA</td>
<td>0</td>
<td>2</td>
<td>119</td>
<td>48</td>
<td>GCA</td>
<td>Ala</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAG</td>
<td>16</td>
<td>9</td>
<td>106</td>
<td>98</td>
<td>GGG</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>46</td>
<td>GUG</td>
<td>20</td>
<td>11</td>
<td>83</td>
<td>123</td>
<td>GCG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.3.
A comparison of the codon usage of the cpg gene and the cat gene.
A comparison of the codon usage pattern in highly and weakly expressed E.coli genes is also shown: H:total number of codons compiled from 64 highly expressed genes. W: total number of codons compiled for 18 weakly expressed genes (Grojean and Fiera, 1982).
basis of an optimal codon-anticodon interaction energy. The observed preference of the cpg gene for the C or G in the wobble position would appear to restrict adherence to these rules. The preference for NNC codons in the case of Phe, Ile, Tyr, and Asp suggest that cpg gene is highly expressed, while the preference for NNU codons in the case of Pro, Ala, Arg and Gly indicates that the cpg gene is weakly expressed. Other features apart from codon preferences, also play a role in the level of product produced, such as the rate of transcription and mRNA stability. Also high levels of expression of CPG2 would result in a drastic reduction in the folate pool in the cell, thus affecting both cell metabolism and DNA synthesis. The most common gene product used to examine promoter activity and expression is chloramphenicol acetyltransferase (CAT) (Close and Rodriguez, 1982; Cohen et al., 1980), an enzyme which despite having poor codon utilisation, is known to be expressed and tolerated at high levels in E. coli (Bennett and Shaw, 1983). To examine if the relatively poor expression of CPG2 was due solely to the nature of the enzyme, the cat gene was substituted for the cpg gene. CAT was cloned as a 0.77 kb TaqI fragment from pBR325 into the AccI site of pNM502. The resulting plasmid, with cat now under the transcriptional control of the P7 promoter, was designated pNM503. When examined in batch culture pNM503 in W5445 exhibited similar expression levels, (7% of the cells soluble protein expressed as CAT) to pNM502 where CPG2 was expressed. Thus, although the codon utilisation of CPG2 is probably poor, it does not appear to make much difference to the levels of expression achieved, in comparison to CAT expression.
The final construction examined utilising the \textit{Pl} promoter, incorporated all the factors that have been examined in this thesis; a high copy number plasmid (in this case pUC8) was used as the backbone to carry the \textit{Pl} promoter, the \textit{par} locus and the \textit{cl} gene.

A description of the isolation of the \textit{Pl} promoter from the plasmid \textit{P}lc28 constructed by Remaut \textit{et al.} (1981) and its subsequent cloning into pUC8, to produce pNMS01 has already been described previously in this chapter (section 5.1). The additional fragments inserted into pNMS01 are all cloned into the pUC8 polylinker region. A schematic representation of the construction of pCM10 is described in Fig. 5.8. The \textit{par} locus was the first addition to pNMS01. Isolated as a 479 bp AatII-HindIII fragment from pAT153 \textit{par}, which was further redigested with TaqI and subcloned into the \textit{AccI} site of pNMS01 to give the plasmid designated pNMS01 \textit{par}. Cloning the \textit{par} fragment in the correct orientation resulted in the \textit{AccI} site being recreated within the linker; into which the \textit{cl} gene, isolated as a 940 bp TaqI fragment from pcI857, can be cloned. The final construction designated pCM10 (Fig. 5.8), contained the \textit{cpg} gene as a 2.03 Kb BamHI fragment cloned into the BamHI site in the polylinker, downstream of the \textit{Pl} promoter.

The levels of \textit{CPG} expression obtained from pCM10 was examined as previously in 1 l shake flasks and 5 l batch cultures. The results for the 5 l batch experiments are described in Fig. 5.9. As expected the levels of expression obtained using pCM10, pre- and post-induction are similar to those achieved by pNMS02\textit{cl}. The maximum level of \textit{CPG} expression produced by pCM10 was of the order of 17% of cells soluble protein; comparable with the maximum level of \textit{CPG2} expression obtained by pNMS02\textit{cl}.
Fig. 5.8

Construction of pCM10
Fig. 5.9

Expression of CPG

in 5 L Batch Cultures

The CPG activity attained by pCH10 induced are represented by •. Analogous values for uninduced cells are represented by O. Arrow represents point of induction.
5.7 Effect of Induction on Plasmid Stability

Seigel and Ryu (1985) reported plasmid loss following induction. Therefore, as part of our investigation into $\text{Pl}$-induced expression it was decided to examine the effect of temperature-induced derepression on the segregational stability of $\text{pPl}$ expression vectors.

The strains $\text{pNM502/K12AH1A trp}$ and $\text{pNM232/K12AH1A trp}$ were grown under carbon limitation in a chemostat at a dilution rate of 0.05 h$^{-1}$. The two cultures were originally maintained at 30°C, but once a steady state was obtained the temperature of the cultures was raised to 42°C. This induction resulted in plasmid instability, $\text{pNM502}$ being lost from its population after 15 generations; with $\text{pNM232}$ plasmid free segregants did not occur until after 13 generations and complete plasmid loss was only obtained after 27 generations (Fig. 5.10). This result confirms the observation made by Seigel and Ryu (1985) that $\text{pPlC}$ plasmids experience slower loss from the population under the derepressed condition compared to $\text{pPlA}$ plasmids.

When examining the dual plasmid system for segregational stability in the strain W5445, the $\text{pPl}$ plasmids were found to be very unstable. They were so unstable that continuous culture could not be employed, since the plasmids were lost before a steady state was achieved. As a result batch cultures were used exclusively to monitor the stability.

The batch cultures were sampled hourly, stability of plasmid $\text{pcI857}$ being monitored on Km plates, while $\text{pPl}$ plasmids were monitored on isomeric plates. The results as described in Fig. 5.11 show that while $\text{pcI857}$ was maintained throughout the cultures, $\text{pNM502}$ and $\text{pNM232}$ were lost rapidly from the culture, prior to induction. $\text{pNM502c1}$ when monitored for segregational stability in batch culture,
Segregational Stability of pNM232 and pNM502 when Induced in the Host K12ΔHla{trp}

The constructions under examination were grown in K12ΔHla{trp} under carbon nutrient limitation at a dilution rate of 0.05 hr⁻¹. Once a steady state was achieved at 30°C, the temperature was increased to 42°C. pNM502 was lost rapidly from the population after 15 generations (●), while pNM232 was lost from the population after 27 generations (■).
showed a marked increase in stability compared to pNM502 in a dual plasmid system presumably due to increased repression of the \( P_l \) promoter which results when the plasmid carries the \( cI \) gene.

When similar batch cultures were used to examine the segregational stability of pCM10, they revealed that the addition of the \( \text{par} \) locus to the backbone had improved the segregational stability, compared to pNM502cI (Fig. 5.11). The combination of repression of the \( P_l \) directed expression by the \( cI \) gene and the addition of the \( \text{par} \) locus appear to improve the segregational stability of the vectors under examination.
Segregational Stability of \( pNM232 \) and \( pNM502 \) with \( pcI857 \) and \( pNM502cl \) and \( pCM10 \) in the Host \( W5445 \)

The segregational stability of \( pNM232 \) (□), \( pNM502 \) with \( pcI857 \) (○), \( pNM502cl \) (△) and \( pCM10 \) (★) in \( W5445 \) were monitored in batch cultures. The cultures were initially grown at 28°C (open symbols) and then during the experiment the temperature was raised to 42°C (closed symbols).
CONCLUDING REMARKS

In the series of experiments examining $P_l$ directed CPG2 expression a number of previously unreported observations have been made. Remaut et al. (1981) based much of their expression work in K12 ΔH1 Δtpr and were able to achieve 30% soluble protein of beta-lactamase expressed by the $P_l$ promoter. This level of expression was never achieved in the expression vectors employed in this chapter, although the means of estimating levels of expression are not the same. Also, Remaut et al. (1981) observed no copy number variation or fall in cell viability on induction. Both features dramatically observed in our expression system, have also been identified by Seigel and Ryu (1985) who also observed a similar phenomenon following induction.

The increase in copy number observed was linked directly to the decrease in growth rate following induction. Stueber and Bujard (1982) and Lin-Chao and Bremer (1986) put forward experimental data showing the inverse proportionality between copy number and growth rate. The coupling of frequency of initiation of replication and growth rate was initially put forward by Prichard (1969) in the inhibitor dilution model, although based on chromosomal replication it could be extended to plasmid replication control. The inhibitor dilution model was initially proposed to explain the observation that chromosomal replication initiation always occurs at constant cellular mass/chromosome ratio (Donachie, 1968).

Basically, Prichard predicted the existence of an inhibitor which was, either synthesised in a burst during replication, or synthesised constitutively and was unstable. Cell growth leads to an increase in cell volume and therefore a progressive dilution of the
inhibitor substance, whereupon replication initiation occurs and more inhibitor is synthesised. This model can be applied to plasmid replication where the synthesis rates of RNAI and RNAII, as well as the stability of RNAI and the efficiency with which it prevents priming of replication, remain constant. Any drop in copy number due to random fluctuations, would lead to a temporarily increased rate of plasmid replication, while a higher than normal copy number would have the opposite effect. i.e., a temporarily decreased rate of replication. For any given growth rate, an equilibrium plasmid copy number exists at which the rate of replication is equal to the rate of plasmid dilution by cell growth. Without this negative control, plasmids would either be diluted out by cell growth or exhibit a runaway type of replication which would be lethal to the cell.

This phenomenon is most dramatically seen when growth rate was reduced to zero by the inhibition of protein synthesis, and thereby cell division, by the addition of chloramphenicol (Clawell, 1972). Lin-Chao and Bremer (1986) showed that RNAI inhibition of plasmid replication was at a minimum under such conditions, while Stueber and Bujard (1982) showed that the addition of chloramphenicol resulted in the amplification of plasmid copy number. The addition of chloramphenicol has long been used to amplify the copy number of ColEl derived plasmids prior to plasmid isolation (Maniatis et al., 1982).

The increase in copy number following induction could possibly explain the low levels of expression obtained when Pl was used in Ki2 ΔH1Δtrp. The expenditure of so much energy on plasmid replication would reduce the cell's ability to carry on transcription and translation functions. This is a problem previously observed in runaway replication vectors (Uhlin et al., 1979).
More success in expressing CPG$_2$ was achieved employing a dual plasmid system, one vector carrying the $P_L$ promoter and gene product, the other co-resident plasmid contained the $cI$ gene. Expression in this system was constitutive, full repression was not achieved due to copy number variation between the vector carrying the promoter and the one carrying the repressor. The results obtained with the dual plasmid system emphasise the need to provide $cI$ repressor protein at levels appropriate to the copy number of the promoter operator sequences; if complete repression of the promoter system is to be achieved.

Incorporation of the $cI$ gene into the backbone of the vector in pNM502cI and pCM10 increased the degree of temperature regulation over the $P_L$ promoter. This, and the addition of the $par$ locus in the vector, resulted in the highest level of CPG$_2$ expression from all the vectors examined (Fig. 5.9). In conclusion, it was anticipated that pCM10 would help meet the demands for the high expression of CPG$_2$ in production, providing the system was capable of being scaled-up.
CHAPTER 6

SCALE UP OF PRODUCTION OF CPG₂ UTILISING pCM10
INTRODUCTION

The next step in the development of this expression system was the scale-up of the fermentation of W5445, utilising pCM10 for the large scale production of CPG2. The process had to be scaled up to 400 l, the capacity of the vessel to be used at the Microbial Technology Laboratory's Pilot Plant. Preliminary scale up work was carried out in a 15 l open frame New Brunswick Scientific (NBS) industrial fermenter. Scale-up was simplified since the volumetric oxygen transfer coefficient (KLa) of all the vessels used was known, and had similar geometry. The plans and construction details of the fermentation vessels used have previously been described by Elsworth and Stockwell (1968). The seed vessel (20 l) was denoted BR, while the main production vessel (400 l) was denoted ER.

RESULTS AND DISCUSSION

6.1.1 Fermenter Medium

The fermentation medium used was described in Chapter 2, it comprises of glycerol, casamino acids, yeast extract and simple salts. The media was designed to meet the needs of the expression system. Glycerol as the major component in the medium plays a key role in the fermentation. As the main carbon source, catabolite repression of the lac promoter which runs in tandem with the Fl promoter is minimised, so maximising expression. The growth rate of cells when grown on glycerol was lower than comparable cells grown on glucose (Lin-Chao and Bremer, 1986). This results in an increase in copy number and gene dosage thereby increasing expression. A similar glycerol based fermentation medium has previously been used in a high expression system for the
production of protein A (Shuttleworth et al., 1987). It appears that the rheology of the fermentation medium, possibly due to the high concentration of glycerol, protects the inherent fragility of the overproducing cells. The ingredients of the fermentation medium were dissolved to make up a concentrated broth to enable easy transfer to the vessel, where the medium was made up to its correct volume with distilled water. The medium was then sterilised in situ by heating the vessel with steam under pressure to 121°C for 30 min. After sterilisation, the vessel's water jacket was used to bring the temperature down. Once the temperature was about 40°C sterile solutions of ampicillin (100 μg/ml) and MgCl₂ were added aseptically to the media via an additional port on the vessel's top plate.

6.1.2 Seed Preparation

A single colony of pCM10/W5445 from an L-agar plate containing ampicillin (100 μg ml⁻¹) incubated at 28°C was used to inoculate a 10 ml L-broth culture containing ampicillin at the same concentration. This culture was incubated at 28°C in a bench top shaker, shaken at 260 rpm for 6 to 8 hr. This in turn was used to inoculate a sterile 2 l flask containing 1 l of L-broth with ampicillin at the appropriate concentration. After incubation at 28°C for 6 to 8 hr (shaking at 160 rpm), this culture was transferred aseptically, in a MDH laminar flow cabinet, to a sterile seed bottle. The seed bottle provided a convenient method of transferring the inoculum into the vessel aseptically.

6.1.3 Fermentation

The 15 l and the 20 l seed culture were both inoculated by
the aseptic addition of a 1 l seed, and aerated by the addition of filter-sterilized bottom air at a rate of 1 vvm. The temperature of the culture was maintained at a constant 28°C with a thermostatically controlled water jacket. The pH of the culture was maintained at 7.0 with the automatic additions of 1 M NaOH or 20% (w/v) H₃PO₄. Dissolved oxygen tension (DOT) was maintained at or above 30% with the use of variable impeller speed. Foaming was controlled as necessary by the addition of PPG-2000 antifoam and the use of back pressure.

At the mid to late stage of exponential growth, as estimated from OD 600 nm readings and CO₂ production, the seed culture was transferred to the main vessel. Control in this vessel was identical to the seed vessel and the NBS 15 vessel. The 15 l and 400 l cultures were both grown for approximately 8 to 9 hr (late exponential growth, as estimated from OD readings and monitoring CO₂ production). At this stage of growth, induction was initiated by raising the temperature to 42°C by a steam injection into the jacket of the vessel. After maintaining the temperature at 42°C for 30 min the culture was harvested.

6.1.4 Harvesting

The culture was harvested 30 min after induction. The 15 l culture was crash cooled by passing cooling water through the vessel jacket. At the same time pH and antifoam additions were isolated and air to the vessel was switched off. The culture was then transferred directly to a Westfalia separator centrifuge, using back pressure where necessary. The centrifuge was run at 8000 rpm and flow rate was approximately 15 l hr⁻¹. The 400 l culture was similarly crash cooled and centrifuged through two De Laval centrifuges running in parallel at
a flow rate of 50-60 l hr\(^{-1}\). The cell paste was removed from the centrifuge bowls, packed in 500 g lots in plastic bags, squeezed flat using a roller and snap frozen in cardice at -80°C. The frozen cell paste was then stored at -20°C until used.

6.1.5 Monitoring of Culture

6.1.5.1 Culture conditions

As stated in 6.1.3, pH was monitored and controlled between preset parameters by automatic acid/alcohol additions. Foaming was controlled by additions of PPG 2000 antifoam and backpressure. Temperature was maintained until induction at 28°C by a thermostatically controlled water jacket. DOT was maintained where possible above 30% using stirrer speed and monitored using an Ingold dissolved oxygen probe. The CO\(_2\) content of the effluent gas from the culture was monitored using an Analytical Development Corporation Ltd CO\(_2\) analyser model SS-200. The growth of the culture was monitored by readings made at 600 nm using a Pye Unicam SP6 spectrophotometer. A more accurate means of measuring the progress of the culture was by carrying out viable counts on tryptone soya broth agar and lab dry weight (LDW) measurements on culture samples. The viable counts and LDW measurements were carried out after the culture on samples stored up to a maximum of 8 hr at +4°C. Viable count plates were also examined for evidence of contamination. Any morphological variation between colonies were further examined by Gram strain and API kits to determine the purity of the culture.
6.1.5.2 **Biochemical analysis**

The level of CPG\(_2\) activity was monitored using the CPG\(_2\) assay described in Chapter 2. Determination of the total protein content allowed the specific activity of CPG\(_2\) to be calculated. From the specific activity of pure CPG\(_2\) the percentage of enzyme as soluble protein was calculated. Confirmation of this value was made by analysis of culture sample run on an SDS-PAGE gel, which also gave a 'crude' estimate to the level of CPG\(_2\) as percentage of the total soluble protein.

6.1.5.3 **Genetic Analysis**

Segregational stability of the plasmids in the culture was monitored by the presence of Ap\(^R\) colonies producing beta-lactamase on isosensitit agar and CPG\(_2\) activity in colonies on folate plates.

Structural stability was monitored by restriction analysis of the plasmid throughout the culture with reference to a control sample of the plasmid. Screening for the physical presence of the plasmid on agarose gels allowed copy number estimations to be made on the plasmid in the culture.

6.2 **15 l Small scale culture**

The fermentation took place over 8.5 hr. Following inoculation the culture grew rapidly, accompanied by an increase in the production of CO\(_2\) and fall in DOT. The pH of the culture rose during the log phase of growth, uncharacteristic of most E. coli fermentations. Usually the production of acids from the tri-carboxylic acid and glycolytic cycle results in a decrease in pH during fermentation. This increase in pH was possibly due to the utilisation
and deamination of the casamino acids in the medium as the carbon source. This observation was supported by the finding that the level of glycerol in the medium remained constant. This suggested that glycerol did not act as the primary carbon source, but was superseded by the more readily metabolised casamino acids. Induction took place 8 hr after inoculation, judged to be near the end of the log phase. The temperature jump to 42°C required for induction was achieved within 6 min. After induction a rapid increase in CO₂ production took place, coupled with a decrease in DOT.

During the induction period, the stirrer speed was unable to maintain DOT levels above 30% as a result 15 min after induction the DOT fell below zero for 5 min. Increasing the air flow rate to 2 vvm enabled the DOT to be restored to 30%. The fermentation parameters of this culture are illustrated in Fig. 6.1.

Assays on the culture showed that a certain amount of CPG₂ was produced prior to induction. Immediately after induction there was a dramatic increase in CPG₂ levels, this increase peaked 20 min after induction, at 85 U mg⁻¹. The levels of CPG₂ expressed during this culture are described in Fig. 6.2. Calculation of the % total soluble protein from SDS gels and specific activity levels showed that CPG₂ was expressed at 17% soluble protein following induction. The culture was harvested 30 min after induction, yielding a total of 770 g of cell paste, wet wt, corresponding to a dry wt of 240 g.

Monitoring the segregational stability of the plasmid during the culture revealed only slight plasmid loss, the greatest loss being experienced during the induction phase. The segregational stability of the plasmid during the culture is illustrated in Fig. 6.3. The structural stability of the plasmid was maintained throughout the
Fig. 6.1

Fermentation Parameters of a 15 l Small Scale Culture

This fermentation took place in a 15 l KRS vessel. Over 8.5 hr induction taking place after 8 hr. The following parameters were monitored during this culture: absorbance at 600 nm (●), % (v/v) CO₂ in effluent gas (▲), Σ DOT (■) and LDW (△).
Levels of CPG\textsubscript{2} Expression During Fermentation of 15 l Small Scale Culture

Expression was induced after 8 hr by raising the temperature of the culture in 6 min to 42\textdegree C from 28\textdegree C for 30 min before harvesting.
Segregational stability was followed by monitoring the % Ap^R colonies and % CPG2 expressing cells in the population during fermentation.
culture.

6.3 **Large Scale Production Culture**

The main culture was inoculated, with a 6 hr old seed culture, judged to be at a mid log phase of growth, thereby ensuring that the main culture commenced growth immediately and reduced the lag phase. The fermentation parameters of both the seed and main culture mirrored that experienced by the 15 l culture, prior to induction.

Induction of the main culture was initiated after 9 hr of fermentation, the temperature jump was achieved within 15 min. A rapid increase in CO$_2$ production and fall in DOT was monitored, but unlike the 15 l culture DOT was maintained above 30% by supplementing the inlet air with oxygen. The fermentation parameters of the seed and main culture are illustrated in Figs. 6.4 and 6.5 respectively.

CPG$_2$ activity in samples from the seed and main culture prior to induction were modest in comparison to the CPG$_2$ levels following induction. After induction CPG$_2$ levels rose immediately from 9.3 U mg$^{-1}$ to 74 U mg$^{-1}$ after 15 min. The amount of CPG$_2$ as % total soluble protein was calculated to be 14. The levels of CPG$_2$ expression in the main culture are illustrated in Fig. 6.6a/b. The culture was harvested 30 min after induction, yielding 10.12 Kg of cell paste, wet wt, corresponding to a dry wt of 3.1 Kg.

Monitoring the segregational stability of the plasmid during the fermentation revealed a fall in the number of plasmid-containing cells in the main vessel, falling rapidly post induction. The degree of plasmid loss is illustrated in Fig. 6.7. In contrast to this plasmid loss, the copy number of the remaining plasmids appeared to increase, as illustrated in Fig. 6.8.
Fig. 6.6

Fermentation Parameters of Seed Culture

This fermentation took place in a 20 l seed vessel (BR) over 6 hr.

The following parameters were monitored during this culture:

- absorbance at 600 nm (●)
- % (v/v) CO₂ in effluent gas (▲)
- DOT (■)
- LDW (△).
Fig. 6.5

Fermentation Parameters of Production Culture

This fermentation took place in a 400 l vessel (ER) over 9.5 hr induction taking place after 8 hr. The following parameters were monitored during this culture: absorbance at 600 nm (∙), % v/v CO₂ in effluent gas (∆) % DOT (■) and LDH (▲).
Expression was induced after 9 hr, by raising the temperature of the culture in 10 min to 42°C from 28°C and maintaining it for 30 min before harvesting.

Fig. 6.6a
Levels of CPG2 Expression During Fermentation of A001 Production Culture
A series of samples were taken during the course of the fermentation and run on a 11% SDS-PAGE. The samples show the increase in CPG$_2$ during the fermentation and particularly the increase, post-induction, in the final sample. The CPG$_2$ band is indicated by the arrow.
Segregational Stability of pCH10 in W5443 During Production Culture

Segregational stability was followed by monitoring the % Ap^R colonies and % CPG2 expressing cells in the culture during fermentation.
Variation of Copy Number During Fermentation

Plasmid content observed in whole cell lysates run on agarose gels.

Fig. 6.8
CONCLUDING REMARKS

The scale up of the process was successful in that it provided valuable cell paste from which CPG2 could be extracted fulfilling some of the demands for CPG2 made by clinical trials. The expression system employed previously, pNM21 in the host W5445, had at best only ever expressed CPG2 as 2% of total soluble protein. This was in comparison to the 14% soluble protein expressed in this production culture. This increased expression was due to the combination of a number of features: the higher copy number of pCM10 (600) compared to pNM21 (150), the more powerful promoter employed P_L compared to tet, the enhanced stability due to par and high CPG2 expression being limited to a short period at the end of the culture.

The 17% level CPG2 expression produced at 15 l was not previously duplicated in scale-up, presumably due to plasmid loss in the production culture, 20% ApR of 50% ApR indicating that the addition of par to the expression vector had not totally stabilized the plasmid in the population. Subsequent purification of CPG2 was as described by Sherwood et al. (1985). The host W5445 is not an ideal host due to its high natural endotoxin production, although it is possible to remove this during purification.
CHAPTER 7
GENERAL DISCUSSION
7. GENERAL DISCUSSION

The aim of the work described in this thesis was to improve the expression of CPG\textsubscript{2}, previously obtained using pNM21 in the host W5445. The potential value of CPG\textsubscript{2} in cancer chemotherapy had necessitated the need for a high expression system for CPG\textsubscript{2}.

The work was originally directed at the instability of pNM21 in W5445 observed in large scale fermentations. This plasmid loss results in a decrease in CPG\textsubscript{2} yields obtained from the fermentation. The natural parental plasmid pAT153 was subsequently shown to be maintained in W5445 under continuous culture, while pNM21 was shown to be unstable under identical conditions. It was inferred from this observation that the plasmid instability must be a consequence of the insertion of the cpg gene into pAT153 and/or CPG\textsubscript{2} expression. In an attempt to resolve this problem the par locus was used to stabilise pNM21 and other pAT153 derivatives. The par locus proved successful in stabilising pNM21, although it could not selectively maintain a plasmid in a population only improve the partitioning mechanism. Also the par locus did not improve CPG\textsubscript{2} expression above that which could be achieved by pNM21 when fully maintained in the population.

One of the simplest ways of achieving high expression is by utilising a high copy number vector. In the course of examining the copy number of a range of plasmids, the pUC plasmids were shown to exhibit a significantly higher copy number than their nearest derivative pAT153. In the construction of the pMTL cloning vector, a series of vectors analogous to pUC, the opportunity arose to examine the difference between pAT153 and pUC plasmids. Sequence analysis of pAT153 and pUC8 identified a single base change within the RNAII
region, immediately preceding the RNAI region. This region is known to play a key role in replication control, and several copy number mutations have been identified within this region (Davison, 1984). It was therefore concluded that the single base change identified could account for the elevation of copy number observed between pAT153 and pUC8.

It has recently been communicated to us by S. Cole that the RNAI transcripts produced by pUC plasmids are 3 bp smaller than those produced by pBR322. It is therefore likely that the G to A substitution has altered the position at which transcription of RNAI initiates. The 5' single strand domain of RNAI has been shown to be crucial for the interaction of RNAI with RNAII (Dooley and Poliaky, 1987) and thereby, regulation of plasmid copy number.

The low levels of CPG2 expressed obtained from pNM21 are due to the transcriptional readthrough from the relatively weak tet promoter, situated upstream from the cpg gene. Previous attempts to increase the expression of CPG2 utilising a more powerful promoter had involved cloning the cpg gene downstream of the lac promoter in pUC8. The subsequent levels of CPG2 expression obtained were disappointing. One possible explanation for the low levels of expression obtained was that CPG2 was being overexpressed, and that the high levels of CPG2 were proving to be detrimental to the host. Although the lac promoter is defined as inducible, full repression is not obtainable. An alternative inducible promoter Pλ, used in conjunction with the cl gene was therefore investigated. The Pλ promoter fully repressed at 28°C by the cl gene product would minimise the adverse effects of over expression by separating cell growth and cell biosynthesis into two distinct phases. Induction carried out at a stage in the growth cycle
when cell densities were at their highest and therefore the protein synthesising capacity of the culture will be at its maximum.

\( P_L \) directed expression of CPG\(_2\) when examined in the host K12 \( \Delta HI\Delta trp \) revealed unexpectedly low levels of CPG\(_2\) expression, compared to previous reports where \( P_L \) expression vectors have been used. The levels of CPG\(_2\) expressed did increase following induction, but this appeared to be as a result of an increase in copy number. Higher levels of expression were achieved in the host W5445, when the ci gene was carried on a separate plasmid which co-exists with the expression vector. Complete repression of \( P_L \) was not achieved, due to the copy number difference between the two plasmids. The higher copy number of the expression vector titrates out the ci repressor molecule, reducing the effectiveness in repressing the \( P_L \) promoter. Repression of the \( P_L \) promoter was achieved by cloning the ci gene into the expression vector, thereby increasing its gene dosage, producing more effective repression and regulation of expression. The tighter regulatory control of CPG\(_2\) expression using this method resulted in the highest levels of CPG\(_2\) expression being achieved.

A basic expression vector was constructed employing the features explored; the par locus, the \( P_L \) promoter and the ci gene on a high copy number backbone. The final test for this expression vector was the scale up and large scale production fermentation. This expression vector system when used in large scale production of CPG\(_2\), produced an 8-fold increase in CPG\(_2\) levels over the previous vector system employed for production. Therefore the aim of this thesis has in part been achieved.

The work carried out in this thesis has also provided much of the ground work and a starting point for the construction of a series
of expression vectors based on the pMTL cloning vectors, employing many of the features examined. The strategy revolves around the construction of vector systems as 'cassettes' in which the individual units combine in unique sites in the backbone in the necessary configuration. It will be possible to construct a cassette carrying different replicons, (eg ColEl, RSF1010, a runaway replicon and a temperature sensitive replicon) different repressor genes (eg, cl gene, lacI), different promoter systems (eg, lac, trp, tac and P1) and different signal peptides giving a choice of product localisation either cytoplasmic or secreted.

The work presented has examined ColEl replicon vectors; their stability, fitness, copy number, mobilisation and use in expression. The improved segregational stability and fitness afforded to the vectors examined, by the par locus, necessitated its incorporation in subsequent vectors constructed and highlighted its importance in plasmid maintenance. In the construction of the pMTL cloning vectors the high copy number of the pUC series of plasmids was explained and exploited in the expression vectors constructed. Although, the pMTL vectors were specifically designed to be nic+, experimental evidence was obtained that clearly demonstrated that the pUC plasmids were also functionally nic-. On the basis of this result the ACGM have now recategorised the pUC plasmids from \(10^{-6}\) to \(10^{-9}\), when used in specially disabled hosts.


Jones, I.M., Primrose, S.B., Robinson, A. and Ellwood, D.C. (1
Maintenance of Some Col El-type plasmids in continuous culture.

plasmids in E. coli grown in chemostat cultures. FEMS Microbiol. l
22, 239-243.

review with special emphasis on carboxypeptidase G2. In: Enzyme
Wiley, New York.


Kawaguchi, Y., Shimizu, N., Nishimon, K., Hozumi, T. and Beppi
(1984). Renaturation and activation of calf prochymosin produced

mutants for the replication of plasmids in K. coli : requirement
deoxyribonucleic acid polymerase I in the replication of the pl


Stewart, G.S.A.B., Lubinsky-Mink, S., Jackson, C.G., Cassel, A. and

Stewart, G.S.A.B., Lubinsky-Mink and Kuhn, J. (1986b). pHG276: A
multiple cloning site pBR322 copy number vector expressing a functional,
lac alpha peptide from the bacteriophage lambda Pr promoter. Plasmid.
15, 182-190.

Stueber, D. and Bujard, H. (1982). Transcription from efficient
promoters can interfere with plasmid replication and diminish expression
of plasmid specified genes. EMBO J. 1, 1399-1404.

number plasmids causes instability: ColEl encodes a determinant
essential for plasmid monomerisation and stability. Cell. 36, 1097-
1103.

Sutcliffe, G. (1979). Complete nucleotide sequence of the E. coli

Swamy, K.H. and Goldberg (1982). Subcellular distribution of various

specified by bacteriophage lambda. Structure, DNA-binding and

(USA). 77, 3369-3373.


The pMTL nic cloning vectors. L improved pUC polylinker regions to facilitate the use of sonicated DNA for nucleotide sequencing

(Recombinant DNA; nic/bom site; mobilisation; par function; copy number; multiple cloning sites; pUC plasmids)

Steve P. Chambers, Sue E. Prior, David A. Barlow and Nigel P. Minton

Microbial Technology Laboratory, PHLS Centre for Applied Microbiology and Research, Porton Down, Salisbury, Wiltshire SP4 0JG (U.K.) Tel. 0980610391

Received 25 November 1987
Revised 11 February 1988
Accepted 15 February 1988
Received by publisher 8 April 1988

SUMMARY

A series of nic cloning vectors have been constructed analogous to the pUC plasmids but which are smaller in size and carry more extensive polylinker regions within the lacZ' gene. The vectors pMTL20 and pMTL21 carry six additional sites (AatII, MfeI, NcoI, BglII, XhoI and Sall) to those present in pUC18 and pUC19, while pMTL22 and -23 possess eleven new cloning sites (AatII, MfeI, NcoI, BglII, XhoI, Sall, Nael, EcoRV, Chrl, Ndel and NruI). More importantly, the relative order of the restriction sites within the polylinker of these latter vectors has been totally rearranged, relative to pUC18 and pUC19, to facilitate the conversion of DNA fragments with incompatible ends to fragments with compatible termini. The availability of such DNA fragments is a crucial requirement when M13 templates are generated for dideoxy sequencing by the sonication procedure. Derivatives of these vectors have also been constructed which demonstrate improved segregational stability by incorporation of the pSC101 par locus. During the construction of these new vectors data were obtained which demonstrated that the pUC and pMTL plasmids contain a previously unreported single base pair difference within the RNA I/RNA II region (compared to pBR322) responsible for a three-fold increase in plasmid copy number. The pUC and pMTL plasmids were also shown to be functionally nic, thus affording the lowest categorisation in genetic manipulation experiments.

Correspondence to Dr. N.P. Minton, Head of Molecular Genetics, Division of Biotechnology, PHLS Centre for Applied Microbiology and Research, Porton Down, Salisbury, Wiltshire SP4 0JG (U.K.) Tel. (0980)610391, ext. 562, 308

Abbreviations: aa, amino acids(s); ACGM, United Kingdom Advisory Committee for Genetic Manipulation; Ap, ampicillin; bp, base pair(s); Cm, chloramphenicol; kb, kilobase(s) or 1000 bp; lacZ', promoter and operator of the lac operon; Mob, mobilisation; ml, molecular weight; oligo, oligodeoxynucleotide; par, partition function; PolIII, Klenow (large) fragment of E. coli DNA polymerase I; RF, replicative form; RIF, resistance; RFΔ, refractory; XGal, 5-bromo-4-chloro-3-indoly-β-D-galactoside; XhoI, enzyme is truncated on the indicated side (3' or 5'); par designation as specified in the carrier state

0278-1101 88 $03.50 © 1988 Elsevier Science Publishers B.V. (Biomedical Division)
Fig. 1. Construction of the pMTL backbone. The pMTL backbone (pMTL4) was derived from pBR322 in the following manner (numbers in parentheses refer to nucleotide position within pBR322; Balbas et al., 1986). A 1693-bp region of pBR322 DNA was deleted by cleaving the plasmid with BamHI (375), blunt-ending with PolI, cutting with PvuII (2066) and then generating pMTL1 by self-ligation of the larger 2.67-kb plasmid fragment. This particular deletion removed the majority of rop but recreated a BamHI site. The remainder of the rop gene, together with the nor site was removed by deleting a 438-bp flaerI (232–2349) fragment. The plasmid obtained, pMTL2.
INTRODUCTION

The pUC series of plasmids are high-copy-number cloning vectors which offer a wide range of unique cloning sites combined with a simple means of recombinant selection (Vieira and Messing, 1982). The linker region carrying the multiple cloning sites lies within the lacZ' gene (encoding N-terminal aa 1-145). The functional polypeptide produced complements a defective lacZ gene (missing aa 11-41) carried by the host genome (i.e., Escherichia coli JM83) or resident episome (i.e., E. coli JM101). Colonies carrying plasmid DNA are therefore blue in the presence of the chromogenic substrate XGal. DNA fragments cloned into the linker region cause insertional inactivation of lacZ' resulting in colourless (white) colonies in the presence of XGal.

In this study we have sought to increase the utility of these vectors, primarily by the provision of a more extensive range of polylinker regions, but at the same time we have effected the reconstruction of the vector backbone. The relative orders of the restriction sites within the polylinkers have been designed to aid the conversion of cloned fragments with dissimilar cohesive ends into compatible termini. Such a conversion allows self-ligation of the DNA fragment, an essential first step in the generation of random M13 templates by the sonication procedure (Deininger, 1983). Reconstruction of the vector backbone was undertaken for three reasons. First, we wished to remove restriction sites from the vector which we proposed to add to the polylinker region. Second, we wished to make our vector unequivocally nic', thus affording the lowest ACGM categorisation in genetic manipulation experiments. Plasmids which lack a functional nic' site cannot be mobilised, even in the presence of mobilising proteins. Finally, we wished to investigate the reasons for the observed (Minton, 1984) high copy number of the pUC plasmids (500-700 copies per cell) relative to pAT153 (150 copies per cell).

MATERIALS AND METHODS

Restriction endonucleases, T4 DNA ligase, calf intestinal alkaline phosphatase, and Polk were purchases from BCL, and T4 DNA polymerase, T4 polynucleotide kinase and pUC8 plasmid DNA from BRL. [a-32P]dATP was obtained from Amersham International Ltd. Oligodeoxynucleotides were synthesised by solid phase synthesis using phosphite triester method (Atkinson and Smith, 1984) with an Applied Biosystems 380A DNA synthesiser. Plasmid DNA isolation, transformation and M13 manipulations have been previously described (Minton et al., 1986a). Compilation of nucleotide sequencing data was undertaken using the computer software of DNASTAR Inc., Madison, WI.

EXPERIMENTAL AND DISCUSSION

(a) Construction of the pMTL backbone

The principal desirable features of the proposed vector were that it should carry an identical deletion to that present in pAT153 (i.e., the two contiguous HaeII-B and HaeII-G fragments of pBR322, carrying rop and the nic' hom site; Twigg and Sherrat, 1980) and that the new restriction sites to be inserted into the linker region should not occur within the vector backbone (e.g., Aaill and AMrl, nt positions 65 and 2499 respectively in pUC19). The simplest way of achieving these aims was to reconstruct the

had also lost an Ndel site (2297) The Psi and HaeII sites within the Ap gene of pMTL2 were removed by substituting the 1.44-kb PstI fragment of the vector with the equivalent region of pUC8, yielding pMTL4. The final step required was to remove the Aaill and EcoIIII sites of the vector. Although this would have been most easily achieved by deleting the DNA between the Aaill (4216) and EcoIIII (115) sites, we wished to retain the latter site for future manipulations. The plasmid was therefore cleaved with Aaill, blunt-ended with T4 polymerase, cut with Ndel, ligated and transformed into E. coli JM83. A recombinant plasmid, pMTL4, was selected which had lost 236-bp of DNA between the Aaill site and the EcoIIII site adjacent to the EcoIIII site. To create the final vector, the approx. 1400-bp HaeII fragment carrying the lacZ' polylinker region was inserted into the induced HaeII site, such that transcription from the lac promoter was from the same DNA strand as the Ap gene. The different polylinker regions inserted are indicated in Fig 3. The pMTL vectors were constructed by inserting a 385-bp EcoRI fragment (blunt-ended with Polk) into the EcoIIII site as described in EXPERIMENTAL AND DISCUSSION, section a.
pUC backbone utilizing pBR322 as a starting point. The region of pBR322 deleted to yield pAT153 is indicated within the circle, the pBR322 DNA utilized in the construction of the pUC backbone (pUR222; Ruthcr et al., 1980) is illustrated by a dashed line outside of the circle, while the pBR322 DNA present in pMTL4 is represented by the bold line. The Hael site at which the M13 derived lacZ' polylinker region was inserted into the pMTL and pUC backbones is highlighted by a wavy arrow.

(b) Derivation of novel linker regions

Novel polylinker regions were derived by inserting complementary pairs of annealed oligos into the appropriate restriction sites of M13mpl8 and M13mpl9. Recombinant clones were detected using the appropriate oligo as a radiolabelled probe and their authenticity was confirmed by nucleotide sequencing. The new polylinker region, together with the x-peptide and lacZ' polylinker, was isolated from M13 RF DNA as an approx. 550-bp HaelI fragment and inserted into pMTL4 at the indicated HaelI site (Fig. 2). Recombinant clones were detected by their blue colouration on agar plates containing XGal. The orientation of insertion of the HaelI fragment was determined by digestion with BglI.

The polylinker regions of pMTL20 and pMTL21 were generated using the oligos illustrated in Table I (A and B), inserted between the SmaI and PvuI sites of M13mpl8 and M13mpl9. They therefore carry unique sites for AatII, MluI, Ncol, BglII, XbaI and Smal, in addition to those present in pUC18 and pUC19. Two complementary pairs of oligos (C/D, and E/F; Table I) were used to generate the polylinkers of pMTL22 and pMTL23, by their insertion between the EcoRI and HindIII sites of M13mpl8 and M13mpl9. The plasmid pMTL26 carried the same DNA insert as pMTL22 except that the second pair of oligos (E and F, Table I) was inserted in the opposite orientation. An equivalent version of pMTL23 was not obtained as the polylinker generated resulted in translational stop codons within the
Fig. 3. The multiple cloning sites of the pM13 vectors and their complementary M13 vectors (M13m1). Only those polylinker regions indicated are available within M13 vectors. It should be noted that the EcoRV sites of the relevant pM13 vectors is only unique in those derivatives carrying par, i.e., pM1322P, pM1323P and pM1326P. In addition, the BglII, ClaI, EcoRV, NdeI, and NsiI sites of the appropriate M13m1 polylinker are not unique. The coding sequence of the α-peptide is indicated below the illustrated sequence. The single-letter aa code is used beneath the second site of the appropriate codon.
lacZ’ gene. Vectors carrying these polylinker regions possess eleven additional cloning sites (AatII, MluI, NcoI, BglII, XhoI, SmaI, NcoI, EcoRV, Clal, NdeI, and NsiI) to those present in pUC18 and pUC19. More importantly, the relative order of the sites has been rearranged. Although the previously available pUC polylinker regions have proved extremely valuable (Yanisch-Perron et al., 1985), the rigidity of the relative order of the sites imposes limits on the choice of restriction enzymes that can be used to resolute the cloned DNA. This is particularly frustrating in the case of the two commonly used sites at the extremes of the linker region, EcoRI and HindIII. DNA fragments cloned into these sites may only be excised by subsequent digestion with these restriction enzymes, either singularly or in combination with another enzyme. The new order of sites in the pMTL vectors largely overcomes the above limitations.

The vectors pMTL24 and 25 carry the polylinkers of pMTL20 and pMTL21 followed by an inverted pUC18 and pUC19 polylinker region. These particular polylinkers should prove useful in converting a cloned DNA fragment from one type sticky/blunt end to another, and complement those polylinkers already available in the pIC series of plasmids (Marsh et al., 1984). The complete range of polylinker regions available in the pMTL series of vectors are shown in Fig. 3.

e) Insertion of the pSC101 par determinant

Previous studies (e.g., Skogman et al., 1983; Zunta et al., 1984) have demonstrated that the segregational stability of cloning vectors may be enhanced by provision of partition functions. The nucleotide sequence of a 375-bp EcoRI-Aval fragment carrying the partition function (par) of the plasmid pSC101 has previously been determined (Miller et al., 1983). The strategy utilized in the construction of the pMTL4 backbone had ensured the existence of a unique EcoRV cloning site into which additional DNA fragments could be inserted without impairment of essential plasmid functions. The plasmids pMTL20 and pMTL21 were therefore cleaved with EcoRV and the pSC101 par locus inserted as a blunt-end-ed 180-bp EcoRI fragment (derived from the above EcoRI-Aval fragment by addition of an EcoRI linker). The resultant vectors were designated pMTL20P and pMTL21P. Deletion derivatives of these two vectors were then obtained by deleting the 350-bp PstI fragment carrying the polylinker/lacZ’ region and the equivalent PstI fragments from pMTL18, -19, -22, -23, -24, -25 and -26 inserted to yield pMTL18P, -19P, -22P, -23P, -24P, -25P, and -26P. An important consequence of these manipulations was that by destroying the backbone EcoRV site, the EcoRV site of the pMTL22P, -23P and -26P polylinkers became unique.

To confirm that the presence of par improved the segregational stability of the pMTL plasmids, comparative experiments were undertaken with pUC8, pMTL20 and pMTL20P. The stability of the three plasmids in E. coli JM83 was quantitatively assessed in a steady-state chemostat population by monitoring the level of Ap-resistant cells. Under conditions of carbon limitation no segregational instability was exhibited by any of the plasmids. Under the more stringent conditions of phosphate limitation, pUC8 and pMTL20 were lost from the population after 15 generations. In contrast, no detectable loss of pMTL20P within the population was observed over a period of 100 generations (Fig. 4).
(d) Mobilisation frequencies of the pUC and pMTL plasmids

The plasmid pBR322 lacks the genes coding for proteins involved in mobilisation and is therefore inefficiently mobilised by large conjugative plasmids such as F. The presence of the site at which these proteins act, the nic/bom site, enables pBR322 to be efficiently mobilised if Mob proteins are supplied in trans, i.e., from a co-resident plasmid such as ColK (Finnegan and Sherrat, 1982). Accordingly, the ACGM have assigned pBR322 an access factor of $10^{-9}$, when used in an especially disabled E. coli host. The plastid pAT153 does not carry the nic/bom site, cannot be mobilised if Mob proteins are supplied in trans, and has therefore been awarded a lower access factor of $10^{-9}$. The pUC plasmids were previously assigned an identical access factor to pBR322 ($10^{-9}$), presumably on the assumption that their mobilisation could be effected by trans-acting Mob proteins.

The construction of the pMTL plasmids was designed to ensure that an equivalent deletion to that present in pAT153 was introduced into the pBR322 derived DNA. The plasmids should therefore be

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Mobilisation frequency$^a$ from JC2926 R64eWll(pLVC9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBR322</td>
<td>$0.9 \times 10^{-4}$</td>
</tr>
<tr>
<td>pAT153</td>
<td>$2.1 \times 10^{-4}$</td>
</tr>
<tr>
<td>pBR128</td>
<td>$2.1 \times 10^{-4}$</td>
</tr>
<tr>
<td>pUC19</td>
<td>$4.5 \times 10^{-4}$</td>
</tr>
<tr>
<td>pUC24</td>
<td>$6.7 \times 10^{-5}$</td>
</tr>
<tr>
<td>pMTL16</td>
<td>$3.6 \times 10^{-5}$</td>
</tr>
<tr>
<td>pMTL11</td>
<td>$4.1 \times 10^{-4}$</td>
</tr>
<tr>
<td>pMTL22</td>
<td>$3.1 \times 10^{-4}$</td>
</tr>
<tr>
<td>pMTL23</td>
<td>$5.8 \times 10^{-5}$</td>
</tr>
</tbody>
</table>

$^a$ Cloning vector mobilised.

$^b$ Mobilisation frequency as assessed as frequency of Ap$^{r}$ transconjugants per donor cell. Conjugations were carried out by the previously described donor mating technique (Meselson et al., 1961), using a Rf$^{m}$ mutant of E. coli W3110. The donor strain IC2926 carries the conjugative mobiliser R64eWll and plasmid pLVC9. This latter plasmid was kindly supplied by Dr. O.J. Warren, and carries the P15A replicon (Chang and Cohen, 1978). The ColE1 mar region (but not nic) and encodes resistance to Cm.

| Unambiguously nic/bom$^*$ and could be awarded an ACGM access factor of $10^{-9}$. Confirmation that the pMTL plasmids were not mobilised even when Mob proteins were supplied in trans was obtained experimentally (Table II). It was apparent, however, that the pUC plasmids were not mobilised under the same experimental conditions. The nic/bom site is known to lie adjacent to the AccI site (Covarrubias et al., 1981) at nt 2245 of pBR322. The recently determined nucleotide sequence of pUC18 (Yanisch-Perron et al., 1985) demonstrated that the nic/bom site of the vector was almost totally destroyed during removal of the AccI site (Vieira and Messing, 1982). On the basis of our findings the ACGM have now reassigned the access factor of the pUC plasmids to $10^{-9}$, when used in an especially disabled E. coli host.

(e) Copy number of the pUC and pMTL plasmids

The copy number of ColE1 and its derivative plasmid pBR322 is negatively regulated by two different gene products (see Scott, 1984). Regulation is mediated through both the RNA I molecule and the Rop (Rom) polypeptide. RNA I is a small untranslated RNA molecule encoded within the DNA region that is used to transcribe the RNA preprimer (RNA II) utilised as the primer in DNA replication. Because RNA I is complementary to RNA II, the two are suggested to hybridise, an event which should interfere with RNA II's role in the initiation of replication. The Rop protein is believed to facilitate this interaction (Tomizawa and Som, 1984). The higher copy number of pAT153 (150 copies per cell), compared to pBR322 (50 copies per cell) is explained on the basis of deletion of DNA which encodes the Rop protein (Twigg and Sherrat, 1980). The pUC plasmids have also effectively deleted the rop gene, however, data obtained in our laboratory have shown that the copy number of these vectors is considerably higher (500-700 copies per cell) than pAT153, and cannot be explained by the deletion of rop alone.

The strategy utilised in the construction of pMTL4 was designed to test our supposition that some other factor, besides deletion of rop contributes to the high copy number of the pUC plasmids. To remove the Parl and HincII sites from pMTL2 we elected to substitute the 1.44-kb TagI fragment carrying the RNA I/RNA II region with the equivalent region...
Fig. 5. The complete nucleotide sequence of pMTL20P. Generation of random templates by the transposon procedure and nucleotide sequencing methodology have been previously described (Blumenthal et al., 1984a). The observed sequence begins at the 4th nt of the HaeIII site at position 0 of the pMTL20 map illustrated in Fig. 1. Unlike pUC, the lacZ' polypeptide terminates immediately after the HaeIII site (nt 252). The position at which DNA was deleted from between the SalI and XhoI sites of pBR322 (i.e., during the construction of pMTL4; Fig. 1) is indicated by an upward arrow. The single base change between pBR322 and pMTL (and the pUC plasmids) is represented by a lower-case g above the sequence at nt position 1036. The single-taster aa codon is used for the α-peptide and β-lactamase proteins, with each aa symbol positioned beneath the 2nd nt of the relevant codon. The translational stop codon of lacZ is indicated by an asterisk. It should be noted that translation of lacZ' will not terminate at the indicated UAA codon in host bacteria (i.e., β-3 and JM110), but at the nonsense UGA codon within the β fragment at nt 2741. The position at which initiation of DNA synthesis occurs during plasmid replication is marked by 'Origin'. The position at which the RNA II primer begins is indicated by a curved arrow. Processing of this primer occurs at the Origins. The transcriptional start point of RNA I is also indicated by a curved arrow. The data shown in this transcr is shown by a horizontal arrowed line.

TABLE III

<table>
<thead>
<tr>
<th>Plasmid*</th>
<th>Size (kb)</th>
<th>Copy number*</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBR322</td>
<td>4.36</td>
<td>57 ± 3</td>
</tr>
<tr>
<td>pAT151</td>
<td>3.68</td>
<td>130 ± 13</td>
</tr>
<tr>
<td>pUC18</td>
<td>3.72</td>
<td>494 ± 30</td>
</tr>
<tr>
<td>pMTL2</td>
<td>2.25</td>
<td>197 ± 17</td>
</tr>
<tr>
<td>pMTL3</td>
<td>2.25</td>
<td>658 ± 14</td>
</tr>
<tr>
<td>pMTL20</td>
<td>2.47</td>
<td>630 ± 10</td>
</tr>
</tbody>
</table>

* All plasmids were transformed into E. coli JM101 (Vieira and Messing, 1982). The amount of pBR322 derived DNA in plasmids pAT151, pUC18 and pMTL20 is indicated in Fig. 1. The derivation of pMTL2 and pMTL3 is shown in Fig. 1.

Copy number determinations were performed by maxima scanning of whole-cell lysates (Preysing et al., 1983) and are expressed as plasmid copies per cell.

from pUC18 to yields pMTL2 (Fig. 1). Copy number determinations on cells carrying pMTL2 and pMTL3 (Table III) clearly indicated that the copy number of the plasmid carrying the pUC-derived replicon (pMTL2, 658 copies per cell) was significantly higher than the plasmid carrying the pBR322-derived replicon (pMTL3; 197 copies per cell). The most obvious cause of such an increase would be a mutation within the RNA I/RNA II region of the plasmid. The recently determined nucleotide sequence of pUC18 (Yanisch-Perron et al., 1985) implied no such difference from the pBR322 sequence (Sutcliffe, 1979). However, when the complete nucleotide sequence of pMTL20P was determined, a single G-to-A substitution was discovered immediately preceding the 5' end of the RNA I coding region (Fig. 5). This single base change repre-
sents the only difference between pMTL2 and pMTL3.

There are a number of reasons why the base change identified causes an increase in plasmid copy number. The simplest explanation would be that the level of RNA transcript has been reduced, however, as a complementary change has occurred within RNA II the possibility cannot be ruled out that it is the C-to-T base change within the RNA II transcript which causes the observed copy-number increase. Although other studies have shown that DNA methylation can regulate DNA replication (Smith et al., 1985; Putten et al., 1986), there appear to be no specific sequences present which are known to be acted upon by either adenosine or cytosine methyltransferase. Indeed the plasmid pMTL3 still exhibited a three-fold higher copy number than pMTL2 in both a dam and a dem E. coli hosts (not shown).

(f) Application of the pMTL vectors to the sonication procedure

Rapid nucleotide sequencing of DNA fragments by the dideoxy procedure is dependent on the prior generation of a set of M13 templates carrying random DNA subfragments representative of the region to be sequenced. One particularly effective way of generating such templates is by the sonication procedure described by Deaninger (1983). A crucial requirement of this procedure is that the DNA fragment to be sequenced must have compatible ends. The vectors pMTL22, pMTL23 and pMTL26 were specifically designed to aid in the conversion of DNA fragments with compatible ends to fragments with compatible termini. Thus, dependent on the site of insertion, cloned DNA may be excised by double digestion with either BamHI/BglII, SalI/XhoI, CiaI/Accl, or combinations of NruI/SalI/EcoRV-NcoI (pMTL22 and pMTL23 only), and circularised by self-ligation prior to fragmentation. It follows that the DNA fragment must have compatible ends. The vectors pMTL24 and pMTL25 present still compatible ends to fragments with compatible termini. Thus, dependent on the site of insertion, cloned DNA may be excised by double digestion with either BamHI/BglII, SalI/XhoI, CiaI/Accl, or combinations of NruI/SalI/EcoRV-NcoI (pMTL22 and pMTL23 only), and circularised by self-ligation. This strategy has been successfully employed in this laboratory to sequence two PstI-HindIII fragments of the Clostridium butyricum plasmid pCB102 (Minton et al., 1988b), a 1.85-kb Pma1-TaqI fragment carrying part of the gene encoding the type A neurotoxin of C. botulinum (D. E. Thompson, J. K. Brehm, J. D. Outram and N. F. M., unpublished data) and a 2.2-kb Spel-ScaI fragment carrying the E. coli mfd gene (D. Nichols, personal communication).

The vectors pMTL24 and pMTL25 present still further opportunities for ensuring that DNA fragments possess self-compatible ends. In both cases DNA fragments which carry different termini compatible with any two of the six sites AarI, MluI, NcoI, BglII, XhoI and SmaI may be cloned and subsequently excised by digestion with either EcoRI, SstI, KpnI, Smal, BamHI, XbaI, Sall or Accl in the case of pMTL24, and with HindIII, Spel or PstI in the case of the pMTL25. In addition, DNA fragments with one terminus compatible with the above six sites and the other compatible with any of the additional sites present in the pMTL24 and pMTL25 linker may also be converted to a DNA fragment with compatible ends. Thus in the case of pMTL24 the second cloning site used could be either Spel, PstI or HindIII with subsequent excision by digestion with either EcoRI, SstI, KpnI, Smal, BamHI, XbaI, Sall or Accl, while the use of the above sites as secondary cloning or excision sites would be reversed in the case of pMTL25. The polylinkers of the analogous pC plasmids may also be used in this manner (Marsh et al., 1984).

The availability of the vectors with and without par has important implications with regard to the overall size of the plasmid. Occasionally the excised cloned DNA fragment has an identical electrophoretic mobility to the linearised vector. This can present problems when preparing μg quantities of the cloned DNA by electrophoresis, as is the case in the sonication procedure. In the majority of cases, if a cloned fragment comigrates with the linearised plasmid vector, then separation will be achieved with the corresponding pMTL24 plasmid, or vice versa.

(g) Conclusions

The series of pMTL vectors constructed retain all the advantages of the pUC plasmids but are smaller (by 300 bp), exhibit improved plasmid stability, and possess more versatile polylinker regions. Both series of plasmids are functionally minicell−, allowing the lowest categorisation in genetic manipulation experiments, and carry a single base change within their pBR322-derived replication regions resulting in a three-fold increase in copy number. These vectors
ACKNOWLEDGEMENT

We thank Tony Atkinson for initial discussions on concept and critical reading of the manuscript, Roy Hartwell for synthesis of oligodeoxynucleotides and would like to acknowledge financial support from the Health and Safety Executive. The authors would also like to thank Stewart Cole and Thierry Garnier for communicating their unpublished observations.

NOTE ADDED IN PROOF

Recent studies at the Institut Pasteur by T. Garnier and S.T. Cole clearly demonstrate that in vitro RNA I transcripts produced using pUC18 template DNA were 3'nt smaller than the equivalent transcript from pBR322. It is therefore likely that the G to A substitution identified in this study alters the position at which initiation of RNA I transcription occurs.

REFERENCES


Plasmid pMTL153: a high copy number version of pAT153 and its use to obtain high expression of the Pseudomonas carboxypeptidase G2 gene

Sieve P. Chambers, Sue E. Prior, Rachell A. Evans, Roger F. Sherwood, and Nigel P. Minton

Molecular Genetics Group, Division of Biotechnology, Centre for Applied Microbiology and Research, Porton Down, Salisbury, Wiltshire, SP4 OJG, UK

Summary. High level expression (27% of the cells soluble protein) of the potentially important therapeutic enzyme carboxypeptidase G2 (CPG2) has been achieved by placing the encoding gene under the transcriptional control of the E. coli lac promoter and using a high copy number version of the plasmid pAT153. The high copy number vector, pMTL153, was constructed by replacing the replication region of pAT153 with the equivalent region from pUC8. The resultant 3-fold increase in copy number was due to a previously unreported nucleotide substitution in the RNA I/RNA II region of pUC plasmids. Evidence is also presented that production of β-lactamase from the vector encoded bla gene dramatically reduces the levels of the periplasmically located CPG2 product. Substitution of the vector bla gene with tet (non-periplasmic gene product) resulted in a significant increase in the levels of CPG2 obtained. This finding may be relevant to the expression of other secreted heterologous proteins.

Introduction

The Pseudomonas derived enzyme carboxypeptidase G2 (CPG2) belongs to a class of enzymes which hydrolyze the C-terminal glutamate moiety from folic acid and its analogs (Sherwood et al. 1985). The key role of reduced folates as co-enzymes in many biochemical pathways has led to the folic acid molecule becoming a prime target in cancer chemotherapy (Bertino et al. 1971). Folate depletion has most commonly been achieved using the folic acid antagonist methotrexate to inhibit dihydrofolate reductase (Bleyer 1978), but may also be achieved directly using the carboxypeptidase G class of enzymes (Kalograti and Bertino 1981). In addition, these enzymes may be used in rescue therapy following high-dose methotrexate regimes (Chabner et al. 1972). More recently an antibody-CPG2 conjugate has been shown to be cytotoxic to a choriocarcinoma cell line in vitro (Searie et al. 1986), and to be selectively targeted to tumour tissue in a choriocarcinoma xenograft in nude mice (Melton et al. 1986). The potential of CPG2 in cancer chemotherapy has brought forward the requirement for large quantities of purified protein for experimental and clinical testing.

Recombinant DNA methodology has meant that high level production of scarce, but therapeutically important proteins, is most often achieved by cloning the encoding gene and directing its high expression in a foreign host, most commonly Escherichia coli. Accordingly the gene encoding CPG2 (cpg) has been cloned, its entire nucleotide sequence determined, and the mature protein shown to reside in the periplasmic space of both E. coli and Pseudomonas putida (Minton et al. 1983; Minton et al. 1984).

Numerous factors are known to affect yields of recombinant proteins, including plasmid copy number, transcriptional/translational efficiency, and product localization, stability and lethality within the cell. In the present study we additionally show that choice of the antibiotic resistance gene carried by the recombinant plasmid vector, can significantly affect the level of recombinant protein attained. Specifically, the use of plasmid vectors carrying the β-lactamase gene (bla), as opposed to the tetracycline resistance gene (tet), dramatically reduces the amounts of CPG2 product measured.

We also describe the construction of a high copy number derivative of plasmid pAT153 which
incorporates a previously undescribed copy number mutation from within the RNA I/RNA II region of pUC plasmids. Recombinant plasmids based on this new vector, pMT153, facilitated high level expression of CPG.

Materials and methods

Bacteria and plasmids. The Escherichia coli strain JM103 (Varma and Manning 1982) was used as the host in the construction of all recombinant plasmids. Plasmids utilised were pUC8 (Varma and Manning 1982), pAT153 (Turgay and Sherratt 1980), pBR322 (Bolivar et al. 1977), pNYM12 (Gibson et al. 1984) and pNYM15 (Minton and Clarke 1983). CPG expression studies were undertaken in the E. coli strain RV308 (pUC8 rpm L str). Media and growth of bacteria. Bacterial strains were routinely cultivated in LB medium, containing 15 g tryptone, 5 g yeast extract, and 5 g NaCl per litre, pH 7.2. L agar was composed of LB medium solidified with 2% agar. The presence of the CPG, was detected using 369 minimal medium (Miller 1972) supplemented with 0.1% lactose and 0.9% glucose as previously described (Minton et al. 1983). Selection for plasmid maintenance was imposed by incorporation of ampicillin (100 μg/ml) or tetrazycline (15 μg/ml) into the medium. Expression studies were conducted in 2XYT liquid medium (16 g tryptone, 10 g yeast extract, 5 g NaCl). 1 litre batch cultures were inoculated to an OD₆₀₀ of 0.05 and shaken at 37°C for 6 hours. Cells harvested at this point were previously shown to be producing maximal levels of CPG (Minton et al., 1983). For in vitro propranolol (15% glycerol/15% methanol/15% glycerol. 2% SDS, 2% mercaptoethanol. 1.5% bromophenol blue) Samples were then sonicated (20 k/s, 2 ampsi) for 3 s and boiled for 5 min. Insoluble material was removed by centrifugation at 10,000 g for 10 min and the supernatant loaded onto a 15% SDS-polyacrylamide slab gel in the presence of 0.1% SDS. The percentage of heterogenous protein was estimated by scanning the Coomassie blue stained gel with a Joyce Loebl Scanning Densitometer.

Results and discussion

Expression of CPG, from thelac promoter of pUC8

Previous studies (Minton et al. 1983) had demonstrated that the recombinant plasmid pMN21, constructed by inserting a 3.1 M dal BglII fragment carrying cpg into the BamHI site of plasmid pAT153, directed the expression of CPG, to between 1% and 2.5% of the soluble protein. This relatively high level of expression (compared to the native promoter, 0.1%) was due to transcription through the upstream iap gene. Subsequently, a smaller version of pMN21 was constructed by localising the gene to a 2.03 kb BamHI fragment (Minton and Clarke 1985). This new plasmid, pNM15, effectively lacked 1 kb of extraneous DNA between the initiation codon of the CPG, gene and the iat promoter. The level of CPG, expression directed by pNM15, however, was equivalent to that obtained with pMN21 (Table 1). In terms of its ability to direct the expression of heterologous genes, the iat promoter may be considered of "moder-
Fig. 1. Construction of recombinant plasmids. The position of the lac promoter/operator region is indicated by lac po, and the direction of transcription by the open arrow. The position of the base change (see Fig. 2) within the replica of pUC8 (and pMTL153), relative to pBR322/pAT153, is indicated by an asterisk. The gene encoding the protein responsible for positive regulation of plasmid copy number is indicated by opr. The pUC8 derived (A) 1.37 kb Scal-NspI fragment used in the construction of pMTL153 (C) is by the dotted, double line. Plasmids pCPM2 and pCPM3 were constructed in the same manner as pCPM1 (B), but using the vectors pAT153 and pMTL153 in place of pBR322, respectively.
ate" strength. One way to improve on the level of cpg expression would be to use a "strong" promoter such as that of the E. coli lac operon.

The 2 kb BamHI fragment carrying cpg was isolated from pNM15, ligated to BamHI cut pUC8 DNA, transformed into E. coli JM83 and transformants selected on M9 minimal medium supplemented with Ap (100 µg/ml) and 0.1% (w/v) folic acid. Clones carrying recombinant plasmids in which the CPG gene was being transcribed from the vector lac promoter (i.e. pNM830, Fig. IA) were readily identified by the appearance of concentric rings of precipitated pteriic acid around individual colonies (Minton et al. 1983). When such clones were restreaked onto L agar the bacterial growth produced was of an atypical nature, being consistently less than expected from the inoculum used while individual colonies had an amorphous appearance. If these colonies were then stabbed into appropriately supplemented M9 minimal agar containing 0.1% folate, precipitation of pteriic acid was seen around the stab within 2 h, long before bacterial growth was visible to the naked eye. Cultivation of these clones in liquid medium was characterised by a considerable lag phase (2 to 3 h) and the levels of CPG attained were extremely variable (0.1% to 1.0% cell soluble protein). The highest yields observed were lower than those previously obtained (Minton et al. 1983) when expression was from the promoter of the tet gene (2.5% cell soluble protein). Examination of the plasmid content of stationary phase cells indicated that deleted forms of the recombinant plasmid had arisen.

Regulation of the lac promoter

One interpretation of these observations was that the level of expression of CPG, directed by pNM830 was so high as to drastically effect the physiological characteristics of the host cell. These effects manifest themselves as a reduction in cell variability and growth rate, and a tendency for plasmid instability. Attempts to culture cells carrying pNM830 would result in the outgrowth of cells carrying deleted forms of the plasmid and consequently low production levels of CPG.

High expression would be a consequence of two factors, the relatively high copy number of pUC8 (600-700 copies per cell), and the relative strength of the lac promoter. One way to alleviate the effects of the latter would be to repress transcription from the lac promoter during early stages of growth and then "switch on" transcription with an appropriate inducer. Such an experiment was undertaken by introducing a compatible plasmid carrying the lacZ gene, pNM42 (Gilbert et al. 1986), into the cell with pNM830. The cells were then grown to mid-log phase and induction of CPG expression elicited by the addition of IPTG. The results are illustrated in Fig. 2, and show that, 6 h after the addition of inducer the level of CPG attained represented 3.2% of the cells' soluble protein, with an indication that continued growth would have resulted in even higher levels of expression. This experiment confirmed that pNM830 directed the expression of higher levels of CPG than pNM21/pNM15, but required the exertion of regulatory control of transcription from the lac promoter during the early stages of growth.

Expression of CPG, from the lac promoter using lower copy number plasmids

Although higher, more consistent, yields of CPG, may be obtained when tight regulatory control of the lac promoter is applied, using IPTG, such a strategy would be prohibitive on the large scale due to the high cost of this gratuitous inducer. In order to reduce the metabolic stress on the cell by
alternative means, we elected to construct a recombinant plasmid with a lower copy number. A 10-fold reduction on pUC8 may be most simply achieved by utilizing pBR322 (50 copies per cell). A new recombinant plasmid was constructed by isolating the CPG2 gene from pNM830, together with the lac promoter, as a 2.33 kb PvuII fragment and inserting it between the ScaI and EcoRI (blunt-ended with Klenow polymerase) sites of pBR322 (Fig. 1B). The resultant plasmid (pCPM1) carried, therefore, a functional lac gene as the selectable marker rather than the bla gene. Insertional inactivation of the bla gene as opposed to the lac gene was preferred as there is now an increasing requirement by regulatory and licensing authorities to avoid the use of recombinant plasmids encoding β-lactamase in the large scale production of human therapeutic agents. Cells carrying pCPM1 exhibited improved growth characteristics and an improved level (2.3% soluble protein) of synthesis of CPG2 (Table 1), compared to cells carrying pNM830.

**Construction of a high copy number version of pAT153**

Having attained moderate yields of CPG2 using an unregulated lac promoter on a lower copy number plasmid, we determined how successive increases in copy number would effect expression levels. An analogous plasmid to pCPM1 was constructed, but using the plasmid pAT153 (150 copies per cell). Cells carrying the resultant plasmid, pCPM2, exhibited improved yields of CPG2 (Table 1), without any discernable effect on cell viability. The increase in level of expression over pCPM1 (3-fold) was consistent with the difference in copy number between the two plasmids. A further recombinant plasmid was constructed using pMTL153.

The cloning vector pMTL153 is a modified version of plasmid pAT153 which has a comparable copy number to pUC8 (600-700 copies per cell). The increase in copy number was achieved by substituting the 1.37 kb Scal/NspH1 fragment of pAT153 carrying ColE1 replication functions with the equivalent region from pBR322 (Fig. 1C). The ensuing copy number increase is due to the presence of a single base difference within the replication regions of pBR322/pAT153 and pUC8.

A preliminary report on this G to A substitution (Fig. 3), immediately preceding RNA I, has already appeared (Chambers et al. 1988). The existence of this copy number mutation within pUC
plasmids was not indicated in the recently published sequence of plasmid pUC19 (Yanisch-Perron et al. 1985), and is now known to affect the size of the RNA I transcript produced (S. T. Cole and T. Garnier, personal communication.) Cells carrying the new recombinant plasmid, pCPM3, produced up to 26.8% of their cell soluble protein as CPGJ (Table 1), compared to 7.1% by pCPM2. This increase could be directly attributed to the higher copy number of pCPM3 relative to pCPM2 (3-fold higher). An SDS-polyacrylamide gel of extracts from E.coli RV308 cells carrying the various recombinant plasmids is illustrated in Fig. 4.

Conclusion

The experiments undertaken in this study have shown that problems encountered with cell viability and plasmid stability when a gene is cloned into an expression vector are not just a consequence of the high expression of that gene. The plasmids pNM830 and pCPM3 are remarkably similar both in their copy number and the promoter utilised to elicit high expression of CPGJ. Their principal difference resides in the vector encoded antibiotic resistance gene. It is apparent that the high, unregulated expression of CPGJ, in cells carrying pNM830 coupled with synthesis of β-lactamase is detrimental to cell growth and viability. Such a situation may arise if the cell is unable to fully export the high levels of CPGJ being produced. The presence of cytoplasmic CPGJ, could then cause intracellular folate depletion. It is tempting to speculate that as both proteins are transported to the periplasmic space of E.coli, their combined high expression overloads the protein export machinery of the cell, resulting in the deleterious presence of CPGJ in the cytoplasm. Indeed, as it is generally believed that protein export through the cytoplasmic membrane occurs at specialised sites (see Randall et al. 1987), the two proteins might be in direct competition for translocation.

A comparable situation does not occur when the resistance gene employed is tetr. rather than tsr. This appears not to be a consequence of lower expression of the tet protein relative to β-lactamase, as both proteins were expressed at similar levels (data not shown). Although nucleotide sequencing has demonstrated the presence of a typical signal peptide at the 5' end of the tet gene (Peden 1983), its gene product is not an exported protein but has been found to be associated with the cytoplasmic membrane (Tait and Boyer 1978). Thus competition between CPGJ and the tet protein for localisation beyond the cytoplasmic membrane does not occur.

The observations made in this study have important implications with regard to the choice of plasmid vector utilised for high expression of recombinant genes encoding exported proteins. Currently, the bia gene is the most frequently used resistance gene in the many available E.coli expression vectors (Pouwels et al. 1985). Its presence may in part account for previous observations of lack of cell viability and low production levels. In the future it may be more advantageous to utilise plasmid expression vectors carrying non-exported resistance gene products (e.g. Tc, Km, and Km') when the recombinant protein of interest is being secreted through the cytoplasmic membrane. In this respect Cm' and Km' genes may prove to be preferable as the latter has been shown to effect the reproductive fitness of E.coli cells (Whan Lee and Eddin 1985). New expression vectors are currently under construction in this laboratory which carry various promoter systems and offer a choice of antibiotic resistance gene.

Acknowledgements. We would like to acknowledge financial support from the Health and Safety Executive.

References


Reference


S. P. Chambers et al. - High expression of carboxypeptidase G,.


Received: April 24, 1985 / Accepted: August 9, 1985.
Attention is drawn to the fact that the copyright of
this thesis rests with its author.
This copy of the thesis has been supplied on condition
that anyone who consults it is understood to recognise
that its copyright rests with its author and that no
information derived from it may be published without
the author's prior written consent.