MANIPULATION OF GENE EXPRESSION USING THE 434 REPRESSOR

THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY AT THE UNIVERSITY OF LEICESTER

BY

CARL INNES WEBSTER B.Sc. (HONS) DEPARTMENT OF BIOCHEMISTRY UNIVERSITY OF LEICESTER

JUNE 1992
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATP</td>
<td>Adenosine 5’-triphosphate</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>CTAB</td>
<td>Hexadecyltrimethylammonium bromide</td>
</tr>
<tr>
<td>dATP</td>
<td>Deoxyadenine 5’-triphosphate</td>
</tr>
<tr>
<td>dCTP</td>
<td>Deoxycytosine 5’-triphosphate</td>
</tr>
<tr>
<td>dGTP</td>
<td>Deoxyguanine 5’-triphosphate</td>
</tr>
<tr>
<td>dTTP</td>
<td>Deoxythymine 5’-triphosphate</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxiribonucleic acid</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>IAA</td>
<td>Isoamyl alcohol</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl β-D-thiogalactoside</td>
</tr>
<tr>
<td>LMP</td>
<td>Low melting point</td>
</tr>
<tr>
<td>MCK</td>
<td>Muscle creatine kinase</td>
</tr>
<tr>
<td>MOPS</td>
<td>3-(N-morpholino)propanesulphonic acid</td>
</tr>
<tr>
<td>OD</td>
<td>Optical density</td>
</tr>
<tr>
<td>ONPG</td>
<td>Ortho nitrophenylgalactoside</td>
</tr>
<tr>
<td>PC</td>
<td>Personal computer</td>
</tr>
<tr>
<td>PEG</td>
<td>Polyethylene glycol</td>
</tr>
<tr>
<td>PGK</td>
<td>Phosphoglycerate kinase</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulphate</td>
</tr>
<tr>
<td>SSC</td>
<td>Standard sodium citrate</td>
</tr>
<tr>
<td>tk</td>
<td>Thymidine kinase</td>
</tr>
<tr>
<td>TPNPG</td>
<td>Paranitrophenylthiogalactoside</td>
</tr>
<tr>
<td>UAS</td>
<td>Upstream activator sequence</td>
</tr>
<tr>
<td>UV</td>
<td>Ultra violet</td>
</tr>
</tbody>
</table>
Abstract

The interactions of the 434 repressor with its operator DNA have been extensively studied in vitro. The aims of the project described in this thesis have been to characterise the activity of the 434 repressor, the product of the coliphage 434 cl gene, and its variants in vivo.

Work presented characterises an optimum, symmetrical 434 operator which demonstrates high levels of repression in combination with the 434 repressor in a reporter system. The combination of one half-site of this operator with a P22 operator half-site, to create a hybrid 434/P22 operator, is used to demonstrate the efficient repression by a 434 repressor/434P22 repressor heterodimer in vivo.

The sequence of the non-contacted bases in the centre of the 434 operator are important for determining the affinity of the repressor in vitro and in vivo. The effect of the sequence of these bases is further characterised in vivo, demonstrating that the presence of G and C bases reduces the degree of repression in a reporter system. A variant of the 434 repressor, the 434Ala44 repressor, is shown to be ineffective in vivo at relieving the deleterious effect of G and C bases in the centre of the operator, despite its improved characteristics in vitro.

The efficiency of the 434 repressor in a reporter system in Saccharomyces cerevisiae is investigated, demonstrating that a functional 434 repressor can be produced in a eukaryotic cell and that this repressor is able to repress expression from a highly active eukaryotic promoter transcribed by RNA polymerase II. Expression of the 434P22 repressor Saccharomyces cerevisiae has a deleterious effect on the viability of transformed yeast cells.

A selection system for altered binding specificity 434 repressors is used to attempt the isolation of a repressor which will recognise a 434 pseudo-operator within the Saccharomyces cerevisiae PGK promoter. No repressors active against this operator have been detected.
CONTENTS

Chapter 1
Introduction. 1
1.1. Repressors and Repression 1
1.11. Repression by E. coli and Phage Repressors 1
1.2. The 434 Repressor 8
1.2.1. The Structure of the 434 Operator/Repressor Complex 8
1.2.2. Altering the Binding Specificity of the 434 Repressor 15
1.3. Heterodimeric DNA Binding Proteins 18
1.3.1. 434 Repressor Heterodimers 19
1.3.2. Heterodimeric Eukaryotic Transcription Factors 20
1.4. The use of Bacterial Repressors to Regulate Gene Expression in Eukaryotes 22
1.5. Aims of the Project 24

Chapter 2.
Repression by a Heterodimeric Repressor in E. coli. 26
2.1. Introduction 26
2.2. Results 27
2.2.1. Plasmids Constructions 27
2.2.1.1. Reporter Plasmids Containing the LacZ Gene 27
2.2.1.2. Construction of pADΔTc 27
2.2.1.3. pAD Plasmids Containing Operators 30
2.2.1.4. Plasmids Expressing the 434 Repressor 30
2.2.1.5. Plasmids Expressing Two Repressor Genes 32
2.2.2. Western Analysis of 434 Repressor Expression 34
2.2.3. Experimental Design 35
2.2.4 Measurement of Repression In Vivo 35
2.2.4.1. Repression by Heterodimeric Repressors 41
2.3. Discussion 44

Chapter 3.
The Effect of the Non-contacted Bases of the 434 Operator on the Levels of Repression In Vivo. 47
3.1. Introduction 47
3.2. Results 49
3.2.1. Plasmid Constructions 49
3.2.1.1. Repressor Plasmids 49
3.2.1.2. Reporter Plasmids Carrying Operators 49
3.2.2. Western Analysis of Strains Expressing Repressor 50
3.2.3. Experimental Design 51
3.2.4. Measurement of Repression In Vivo 51
3.3. Discussion 60

Chapter 4.
The 434 Operator/Repressor System as a Means of Regulating Eukaryotic Gene Expression. 64
  4.1. Introduction 64
  4.2. Results 65
  4.2.1. Plasmid Constructions 65
  4.2.1.1. Location of Operators in the PGK Promoter 65
  4.2.1.2. Construction of Reporter Plasmids 68
  4.2.1.3. Construction of a Plasmid Expressing the 434 Repressor 73
  4.2.2. Western Analysis of 434 Expression in Yeast 75
  4.2.3. Experimental Approach 76
  4.2.4. Measurement of Repression 76
  4.2.4.1. Effect of Operators on Expression From the PGK promoter 76
  4.2.4.2. Repression of Reporter Constructs 76
  4.3. Discussion 81

Chapter 5.
The Effect of the 434$^{P22}$ repressor on the growth of Saccharomyces cerevisiae. 85
  5.1. Introduction 85
  5.2. Results 86
  5.2.1. Plasmid Constructions 86
  5.2.2. Experimental Approach 90
  5.2.3. The Effect of the 434$^{P22}$ Repressor on Transformation Efficiency 90
  5.3. Discussion 97

Chapter 6.
Selection of Altered Binding Specificity Repressors. 99
  6.1. Introduction 99
  6.2. Results 101
  6.2.1. Screening for Pseudo-operators 101
  6.2.2. Plasmid Constructions 101
  6.2.2.1. Plasmids Containing Operators 101
  6.2.2.2. Construction of the 434 Expression Vector 103
6.2.3. Screening the N29 Library with Operator Plasmids on TPNPG and Analysis of Transformants 107

6.3. Discussion 108

Chapter 7.
Discussion and Conclusion. 111
7.1. Discussion 111
7.2. Conclusion 115

Chapter 8.
Materials and Methods. 116
8.1. Transformation of E. coli. 116
8.1.1. Preparation of Competent Cells 116
8.1.2. Introduction of DNA 116
8.1.3. Treatment of Cells Transformed with Plasmids 117
8.1.4. Treatment of Cells Transformed with M13 RF DNA 117
8.2. Miniprep 118
8.3. Midiprep 118
8.4. Preparation of Phage DNA 119
8.4.1. For RF DNA 120
8.4.2. For Single Stranded DNA 120
8.5. Sequencing DNA 120
8.5.1. DNA preparation for Plasmid Sequencing 121
8.5.2. DNA Preparation for Single Stranded Phage Sequencing 121
8.5.3. Sequencing Reaction 121
8.6. Purification of Restriction Fragments from Agarose 122
8.7. Radiolabelling of DNA 123
8.7.1. Restriction Fragments of DNA 123
8.7.2. End Labelling of Oligonucleotides 124
8.8. Transformation of Yeast 124
8.8.1. Preparation of Competent Cells 124
8.8.2. Introduction of DNA 124
8.9.8.8. Transformation of Yeast 124
8.8.1. Preparation of Competent Cells 124
8.8.2. Introduction of DNA 124
8.9. Isolation of Genomic DNA from Saccharomyces cerevisiae 125
8.10. Southern Blotting 126
8.11. Hybridisation of Probes to DNA Immobilised on Nylon Membranes 126
8.12. β-galactosidase Assay 127
8.12.1. For *E. coli* 127
8.12.2. For *Saccharomyces cerevisiae* 128
8.13. Preparation of Protein from *E. coli* and *Saccharomyces cerevisiae* 128
8.15. Western Analysis of Protein from *E. coli* and *Saccharomyces cerevisiae* 130
8.15.1. SDS Polyacrylamide Gel Electrophoresis 130
8.15.2. Western Blotting 130
8.15.3. Detection of Antigens 131
8.16. Site Directed Mutagenesis 132
8.16.1. Preparation of Template 132
8.16.2. Preparation of Mutagenic Oligonucleotide 133
8.16.3. Mutagenesis Reaction 133
8.17. Statistical Calculations 133
8.17.1. Linear Regression 133
8.17.2. Standard Deviation of Repression 134
8.17.3. Students t-Test 135
8.18. Description of Plasmids 136
8.18.1. *E. coli* Plasmids 136
8.18.2. Yeast plasmids 136
8.19. Strains 138
8.19.1. *Escherichia coli* 138
8.19.2. Yeast 138

Chapter 9.

References. 139

Appendices.

The DNA sequence and deduced amino acid sequence of the 434 CI gene encoding the 434 repressor.
Program for the processing of data from a microtitre plate reader for \(\beta\)-galactosidase assays.
CHAPTER 1.

Introduction
1.1. Repression and Repressors

Repression is the reduction of gene expression by the inhibition of transcription. In terms of the repression to be discussed in this thesis repression is brought about by the binding of a protein, termed a repressor, to a specific DNA sequence, termed an operator, the result of which is a significant reduction in the transcription of a gene associated with the operator.

1.1.1. Repression by Escherichia coli and Phage Repressors

Repression is a phenomenon which occurs widely in the regulation of gene expression. The first examples of this have come from Escherichia coli and the bacteriophages, such as lambda, which infect it.

The regulation of the lac operon of Escherichia coli is a classic example of the negative control of gene expression at the level of transcription (Jacob and Monod, 1961) and a considerable amount of data has been accumulated in elucidating the mechanism by which the expression of this gene is regulated. During growth on a carbon source consisting of lactose a 1000 fold higher level of lac operon-encoded enzymes are expressed. This apparent activation of expression is brought about by a relief of repression, in response to carbon source. The repression is brought about by a DNA-binding protein, the lac repressor, which is the product of the lacI gene. This binds to a specific DNA sequence, the lac operator O₁, located immediately downstream of the -10 binding site for RNA polymerase within the lac promoter. It was initially thought that the binding of the lac repressor to its operator prevented transcription by steric hindrance of the binding of RNA polymerase (Galas & Schmitz, 1978) although it has now been shown that the affinity of the promoter for RNA polymerase is increased in
the presence of the repressor (Straney and Crothers, 1987) and that the lac repressor acts by halting RNA polymerase in a pretranscriptional complex (Straney and Crothers, 1987, Lee & Goldfarb, 1991). Two other operators for the lac repressor have also been described within the lac operon, one downstream of O₁, O₂ (Rezinkoff et al., 1974) and one upstream of O₁, O₃ (Gilbert et al., 1976). These operators have a relatively low affinity for the repressor (Pfahl et al., 1979, Winter and von Hippel, 1981, Fried and Crothers, 1981) and due to their location their contribution to repression was thought to be relatively small. However, detailed studies in vitro have shown that a lac repressor tetramer is able to bind two of the lac operator sequences simultaneously, either on two separate DNA fragments (O'Gorman et al., 1980, Culard and Maurizot, 1981) or on a single DNA fragment, where the intervening DNA is forced into a loop structure (Krämer et al., 1987; 1988). The ability to form a loop depends on the relative molarities of DNA and repressor. In vivo it has been demonstrated that all three lac operators greatly contribute to the overall level of repression: the loss of either O₂ or O₃ reduces repression 2-3 fold and the loss of both reduces repression >50 fold (Oehler et al., 1990), thus demonstrating that the three operators act together to regulate the expression of the lac operon.

The closely related temperate phages λ, 434 and P22 (Hershey and Dove, 1971; Botstein and Herskowitz, 1974) encode a number of sequence-specific DNA-binding proteins, including a repressor and a cro protein (Ptashne et al., 1980, Johnson et al., 1981), which regulate transcription of phage genes. The cro proteins are small, single domain proteins whereas the repressors contain an N-terminal domain which binds DNA and a C-terminal domain which mediates dimerisation (Pabo and Sauer, 1979; Sauer et al., 1979; Johnson et al., 1981). Analysis
of the protein sequences of cro and repressor proteins reveals a marked conservation of amino acid sequence suggesting a common ancestry. The crystal structures of the DNA binding domain of the \( \lambda \) repressor and \( \lambda \) cro (Anderson et al., 1981; Pabo and Lewis, 1982) have been determined, as have the structures for the DNA binding domain of the 434 repressor both in complex with its operator (Anderson et al., 1985, 1987, Aggarwal et al., 1988), and as a free protein (Mondragon et al., 1989b) and 434 cro with its operator (Mondragon and Harrison, 1991) and as a free protein (Mondragon et al., 1989a). The analysis of the crystal structures has revealed that these proteins all recognise and bind to DNA by a similar mechanism, the so called helix-turn-helix motif. This motif is indeed widespread in its use as a DNA-binding motif in both prokaryotes (Sauer et al., 1982) and eukaryotes, including the homeodomain proteins (see for example Kissinger et al., 1990; Scott et al., 1989).

Coliphages have a mechanism of regulating gene expression, similar to the lac operon, which is critical to the life cycle of the phage. The best characterised is that of the phage lambda, \( \lambda \). Like the lac repressor/operator system, gene regulation is brought about by the specific interaction of protein and DNA, again a repressor/operator combination. This is explained in some detail by Ptashne (1986). The life cycle of the phage is precisely regulated and may follow either of two paths: A lytic path, controlled by the ordered expression of phage genes where the phage chromosome is highly replicated, new head and tail proteins are synthesised, new phage particles are produced and the host cell lyses releasing them: or a second path, lysogeny, where all but one of the phage promoters is turned off, the phage chromosome is integrated into the host chromosome and is passively replicated with the host DNA. The single active promoter in lysogeny is the \( \text{p}_\text{RM} \)
promoter which drives expression of the phage *cl* gene which encodes a repressor. The transcript from the $p_{RM}$ promoter also includes a second gene, *rex*, which confers resistance to superinfection of $\lambda$ infected cell by *rII* mutants of T4 phage. The $\lambda$ repressor binds to its operators as a dimer and therefore, like the *lac* operator, an ideal or consensus operator is a palindrome. Unlike *lac* however, the operators in $\lambda$ are arranged in equally spaced groups of three, which permits interaction between repressors bound at adjacent operators. The repressors in fact bind cooperatively, i.e. the binding of a repressor at one operator apparently increases the affinity of an adjacent operator for a second repressor dimer. This is comparable to the system in *lac* where, instead of looping to increase the degree of repression, cooperative binding between adjacent operator sites increases the degree of repression relative to the binding of a repressor at a single operator site (Figure 1.1).

![Figure 1.1](image)

**Figure 1.1** Cooperative vs non-cooperative repression. The graph demonstrates the difference in the rise in repression on increasing repressor concentration between operators acting cooperatively and a single operator acting non-cooperatively. Non-cooperative repression increases slowly with increasing repressor concentration whereas cooperative repression demonstrates a rapid rise in repression with only a small increase in repressor concentration.
Figure 1.2. The three OR operators of λ overlapping the pR and pRM promoters. Binding of a repressor at OR1 represses expression of pR, the binding of a repressor at OR2 enhances expression from pRM.

In contrast to the lac operator, the operators in λ overlap with the binding site for RNA polymerase, thus preventing the formation of a transcription complex. The three operators in λ actually overlap two promoters, pRM and pR (Figure 1.2). The affinity of each of the three operators for the repressor is not equal, the first operator OR1 having the highest affinity, then OR2, then OR3. The binding of the λ repressor to the OR operators switches off pR but enhances expression from pRM, the promoter for the cl gene encoding the repressor. This is brought about because OR1 overlaps the pR promoter and the binding of the repressor under normal conditions is such that only OR1 and OR2 are occupied. This switches off pR but the position of the repressor occupying OR2 is such that it makes contact with RNA polymerase binding at pRM, enhancing the transcriptional activity from this promoter. A second DNA-binding protein, cro, is expressed from the gene controlled by pR and is hence switched off during lysogeny. This
protein binds to the same set of operators as the repressor but unlike the repressor has its highest affinity for or3. Certain environmental stimuli, such as exposure to ultra-violet light, result in the degradation of repressors within the cell, by activation of the RecA protease which specifically cleaves the repressor protein between the DNA-binding domain and the dimerisation domain. This relieves repression at pR allowing expression of cro. Cro then binds at or3, switching off the expression of the cl gene from pRM and hence the synthesis of repressor. Unlike the repressor, cro does not bind cooperatively. Thus the first effects of cro are the switching off of pRM preventing the expression of the repressor. The levels of cro build up until all of the or operators are occupied, switching off both pR and pRM, thus cro eventually inhibits its own synthesis and reduces expression of λ early genes.

The cooperative binding of the repressor to the or operators demonstrates a rapid reduction in gene expression from only small increases in protein concentration, analogous to a switching effect rather than a slow reduction in gene expression from a single operator on increasing repressor concentration (Fig 1.1).

The λ repressor also binds to 3 similarly arranged operators in a second promoter, pL, switching off the expression of the N gene, the product of which is also important for lytic growth. The binding of cro to the 0L operators in pL slowly inactivates the expression of N, switching gene expression from the early to late genes.

Many other genes in E. coli utilise repressors as a means of regulating gene expression, for example Met where cooperative binding of repressor with corepressor to tandem operators also enhances repression (Philips et al., 1989). The trp operon is also regulated by a repressor, the trp repressor (Reviewed in, Perutz, 1989).
This repressor exists as an inactive apo-repressor until it binds L-tryptophan: it then becomes an active repressor and binds to operators in the genes of the \textit{trp} operon and the gene \textit{aroH} repressing their expression. Thus the product of the metabolic pathway encoded by the genes in the \textit{trp} operon, L-tryptophan, inhibits its own synthesis.

Perhaps one of the more complex operator/repressor systems is that of the \textit{E. coli} \textit{lexA} repressor, which is involved in the SOS regulatory system (reviewed in, Little and Mount, 1982). In response to environmental stimuli such as exposure to ultra-violet light the expression of several \textit{E. coli} genes is enhanced. This leads to the phenomena of enhanced DNA repair, enhanced DNA mutagenesis, inhibition of cell division and prophage induction (as mentioned earlier with reference to \textit{\lambda} phage). This increase in gene expression is again a relief of repression. The product of the \textit{lexA} gene, the \textit{lexA} repressor, binds to a number of operators in the promoters of a wide range of genes which are induced during the SOS response, including the \textit{recA} gene. The product of the \textit{recA} gene, the RecA protein, under normal cellular conditions is involved in recombination events. However, at the onset of the SOS response it undergoes a change which alters its activity to that of a protease. This specifically cleaves certain repressor proteins, including the \textit{lexA} repressor and also the \textit{\lambda} repressor, between their DNA binding domain and dimerisation domain. This greatly reduces their effectiveness as DNA-binding proteins thus relieving the repression of genes containing operators bound by these repressors. This then allows increased expression of the genes involved in the SOS response.
1.2. The 434 Repressor

1.2.1. The Structure of the 434 Operator/Repressor Complex

In order to understand the mechanism involved in DNA-binding the precise structure of the repressor must be known. The λ repressor has been extensively studied and a clear picture of its interactions with DNA described. However a more detailed picture of the repressor from the phage 434 has emerged, making this system a good candidate for further work.

The coliphage 434 is a phage related to λ and is extremely similar to λ in its genetic organisation and the way it regulates its life cycle using repressor and cro proteins. Like the λ repressor, the repressor from the cl gene of 434 has been extensively characterised. The repressor from phage 434 binds to six operator sites on the phage chromosome. Comparison of these sequences reveals a consensus pattern with the outer 4 bases conserved in five of the 6 operators and the sixth differing at only one position. In contrast the inner six bases are somewhat variable in sequence, although they are predominantly T/A-rich, with exceptions in some positions (Wharton, 1985), as shown in figure 1.3.

The DNA-binding domain of the 434 repressor, amino acids 1-69, has been co-crystallised with its operator, both a synthetic operator (Anderson et al., 1985; 1987) and the wild-type operator Or1 (Aggarwal et al., 1988).

A high resolution view of the interactions between the repressor and operator reveals a complex interaction of many amino acids within the protein with the base-pairs and, the phosphate backbone of the operator DNA. The fragment 1-69 was found to consist of 5 α-
### Introduction

<table>
<thead>
<tr>
<th>OR1</th>
<th>ACAAGAA AGTTTGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>OR2</td>
<td>ACAAGAT ACATTGT</td>
</tr>
<tr>
<td>OR3</td>
<td>ACAAGAA AAACTGT</td>
</tr>
<tr>
<td>OL1</td>
<td>ACAAGGA AGATTGT</td>
</tr>
<tr>
<td>OL2</td>
<td>ACAATAA ATATTGT</td>
</tr>
<tr>
<td>OL3</td>
<td>ACAATGG AGTTTGT</td>
</tr>
</tbody>
</table>

Figure 1.3. The operators of the 434 repressor. As in λ the operators are arranged in two groups of three, OR and OL. For each group the operator with the highest affinity is numbered 1. A gap in the sequence represents the division between the two half-sites of the operator.

Helices linked by turns of varying length (Aggarwal et al., 1988). The first helices of the 434 repressor closely correspond to the first four helices in λ, although λ contains additional protein structures (Pabo and Lewis 1982; Pabo et al., 1982). A comparison of the DNA-bound fragment 1-69 and the free fragment reveals no significant change in the conformation of the α-carbon backbone on binding to DNA (Mondragon et al., 1989b), but significant adjustments of side chain orientations do occur. The repressor binds as a dimer and although dimerisation is predominantly linked to the C-terminal domain, contacts occur between the DNA-binding domains involving several hydrophobic amino acids (Leu\(^{45}\), Pro\(^{46}\), Val\(^{56}\) and Leu\(^{60}\)) as well as a salt bridge from Arg\(^{41}\) to Glu\(^{47}\).

The 1-69 dimer is bound to DNA with the α-3 helix lying in the major groove of the operator and with the amino ends of helices 2 and 4 making contact with the phosphate backbone of the DNA stabilising the sequence-specific interactions (Fig.1.4). Direct contacts occur between the protein and positions 1 to 4 of the operator, as well as to position -1, and there are similar contacts for both half-sites. At
position 1 a bidentate hydrogen bond links Gln$^{28}$ to the adenine: this is the only base bound by a single amino acid, although the side-arm of Gln$^{28}$ is also involved in van der Waals interactions with thymine at position -1 of the OR1 operator and is stabilised by hydrogen bond interactions with Gln$^{17}$. The van der Waals interaction between Gln$^{28}$ and the 5-methyl group of thymine at -1 may explain some of the differences in affinity between some of the wild-type 434 operators which have apparently no other difference in sequence (Wharton, 1985). The C:G base-pair at position 2 of the operator is contacted by Gln$^{29}$, with a bidentate hydrogen bond linking the ε-NH$_2$ group of the Gln to O6 in the guanine ring. At position 3 the A:T base pair is not involved in hydrogen bonding with the protein, a hydrophobic pocket is created by the side chains of Gln$^{29}$ and Thr$^{27}$ which is orientated towards the 5-methyl group of thymine and direct contacts are non-polar van der Waals interactions. At the A:T base-pair at position 4 of the operator a single hydrogen bond links the O2 of the thymine with the ε-NH$_2$ group of Gln$^{33}$. Interactions between position 4 of the operator and Gln$^{33}$ are influenced by the base-pair at position 5. In one half-site of OR1 position 5 is occupied by a G:C base-pair, here the ε-oxygen of Gln$^{33}$ is orientated towards C5 of cytosine and is stabilised by hydrogen bonding to a solvent molecule bridging the phosphate backbone of the operator to Ser$^{30}$. In the other half-site of OR1 position 5 is occupied by a T:A base pair. Here the the ε-carbonyl of Gln$^{33}$ is rotated away from the position which is now occupied by the methyl group on thymine. This does not alter the hydrogen bond length to the thymine at position 4, but lengthens the bond to the water molecule and also influences the orientation of the side chains of amino acids 26 to 30, and may alter the strength of the interaction between protein and DNA.
Figure 1.4. The interactions between the 434 repressor and one half-site of the operator or1. The α-helices of the repressor are represented as cylinders. Hydrogen bonds linking the repressor and operator are represented as fine dotted lines. The methyl groups of thymine are represented as circles. Phosphate groups on the backbone of the DNA are represented as open ovals. A solvent molecule is represented as a filled circle.
The importance of each amino acid/base-pair interaction has been illustrated by Wharton (1985), who showed that if any of the consensus bases of the 434 operator are altered the affinity of the operator for the repressor is lowered at least 10-fold.

Extensive interactions also occur between the repressor and the phosphate backbone of the operator DNA, helping to position the α-3 helix in the major groove. Residues involved include Arg_{10}, Gln_{17}, Asn_{36}, Asn_{16}, Glu_{32}, Lys_{40}, Arg_{41} and Arg_{43}. Thus the repressor molecule is able to bind the operator by a complex interaction between the amino acids and the bases, sugars and phosphates of the DNA.

In addition to the bases in positions 1 to 4 in both half-sites the bases in the central region of the operator also influence the repressor/operator complex. Only one amino acid within the 1-69 fragment used for the X-ray crystal studies makes any contact with the central region of the operator, Arg_{43}. However, this is a water-mediated hydrogen bond to base-pair 7 in the minor groove of the operator and these bonds could be made with other sequences. Indeed the influences of base-pairs 6 and 7 have been shown to be independent of Arg_{43}. Koudelka et al. (1987) mutated Arg_{43} to an alanine and although the mutant protein showed vastly reduced affinity for the operator it showed the same dependence on the sequence of the centre bases, implying that the sequence-specific affinities are not due to sequence-specific contacts through Arg_{43}. The influence of the centre 4 bases on the affinity of the operator for the repressor has been extensively studied in vitro with reference to a synthetic operator, ACAATATATATTGT (Koudelka et al., 1987; 1988). The results clearly demonstrate that the presence of G:C or C:G base-pairs in any position in the centre 4 bases of the operator reduces the affinity of the operator for the repressor. The precise sequence of A:T or T:A base-pairs at
positions 7 or 8 has no negative effect on the affinity. However, a number of sequences have a slightly higher affinity, notably a 7A operator, ACAATAAATATTGT, has a three-fold higher affinity for the repressor. The inclusion of G:C or C:G base-pairs at either of these two positions has a large effect on the affinity of the operator for the repressor. A symmetrical operator, ACAATAGCTATTGT, or its counterpart ACAATACGTATTGT, reduced the affinity of the operator for the repressor by 50-fold. A single introduction of a G:C base-pair to give the operator ACAATAGATATTGT had a less dramatic effect, lowering the affinity by 6-7 fold. Similar changes at positions 6 and 9 show the same preferences for A:T or T:A base-pairs but the effect of G:C or C:G base pairs is reduced. The inclusion of G:C and C:G base-pairs to give ACAATGTACATTGT or its counterpart ACAATCTAGATTGT, reduces the affinity of the operator for the repressor by 5 fold. The introduction of a single G:C base-pair to give ACAATGTATATTGT reduced the affinity of the operator for the repressor by 2.5 fold. The X-ray crystal studies of the 434 operator/repressor complex showed that the DNA near the centre of the operator is slightly over-wound, so that the minor groove is compressed and that the operator DNA is slightly bent (Anderson et al., 1985; 1987; Aggarwal et al., 1988). In addition to these distortions of the DNA the highest resolution X-ray crystallography studies (Aggarwal et al., 1988) have revealed that there is significant distortion of the bases in the non-contacted region of the operator. Base-pairs 5-10 of the operator or 1 are significantly distorted by propeller twist, the relative rotation of the bases about their long axis, allowing additional hydrogen bonding as well as the normal Watson-Crick bonds. This forms bifurcated hydrogen bonds between some of the bases, even across the minor groove. A bifurcated hydrogen bond is one in which a
proton is shared between two acceptors. This gives significant importance to the centre 4 bases of the operator, as alteration of any of these would alter the formation of the bifurcated hydrogen bonds necessary to stabilise the distorted structure. Evidently this DNA conformation is required to align the two operator half-sites for optimum repressor contacts. If this is the case then an operator with a higher flexibility will be able to bind the repressor at a higher affinity than ones with a lower flexibility. Work by Koudelka et al. (1988) demonstrated that if an operator was specifically nicked on one strand in the centre of the operator making it more flexible then that operator bound the repressor with approximately 5 fold higher affinity than either the corresponding unnicked operator or DNA nicked 4 base-pairs outside the operator. By introducing a single amino acid substitution, Phe\textsuperscript{44} to Ala\textsuperscript{44}, in the repressor they generated a mutant repressor protein which has reduced sensitivity to the presence of G:C or C:G base pairs in the centre 2 positions of the operator in vitro. They proposed that this Ala\textsuperscript{44} protein was more flexible in the dimer interaction and that this flexibility compensated for the reduced flexibility in the operators containing G:C or C:G base-pairs in the centre 2 bases. It therefore seems likely that operators containing G or C base pairs in the centre of the operator have a reduced affinity for the repressor due to an increased rigidity imparted by the extra hydrogen bond between G and C bases in double-stranded DNA compared to A and T bases. The interactions with the repressor also bend the operator, causing it to appear as an arc with radius 65Å. The bending of the operator into the arc is not uniform, the DNA in the centre of the operator is relatively straight with the bending occurring 2-3 bases from the centre. The bent configuration is not especially extreme as DNA
bent around the histone octamer proscribes an arc of 43Å in the nucleosome core (Richmond et al., 1984).

By utilizing a ring-closure assay to determine the intrinsic twist of operator DNA, Koudelka and Carlson (1992) demonstrated that it is more likely to be the twisting of the operator DNA which influences the affinity of the repressor. The distribution between two topoisomers of circularised operators is an indication of the degree of twist. In the presence of the repressor two operators with differing affinities for the repressor have identical distributions between the two topoisomers. However, in the absence of the repressor the operator with the highest affinity shows a distribution closer to that in the presence of repressor than does the operator with weaker affinity. Thus, the degree to which the DNA is twisted by the repressor is independent of sequence, but, the intrinsic twist of the DNA is dependent on sequence and the ease with which the repressor can deform the DNA in the repressor/operator complex is therefore sequence dependent.

The affinity of the operator for the repressor is therefore determined both by the sequence-specific interactions between the conserved bases of the operator and the amino acids of the DNA-binding helix of the repressor, and the sequence-dependent flexibility of the non-contacted base-pairs of the operator.

1.2.2. Altering the Binding Specificity of the 434 Repressor

The 434 repressor has proved a useful tool for the investigation of the interactions responsible for DNA recognition and affinity. Before a full and detailed picture was available from the X-ray crystallography studies Wharton et al. (1984) showed that the sequence-specific DNA-binding could be altered by a changing the amino acid sequence of the putative DNA-binding helix in the helix-turn-helix motif. They created
a hybrid 434 repressor in which the amino acids within the DNA-binding helix of the 434 repressor were altered to those of the 434 cro protein. This hybrid protein had a binding specificity more similar to the cro protein than the repressor, as determined by DNA methylation interference experiments. It should be realised that the 434 repressor and cro protein already have similar DNA-binding specificities and that the changes observed were subtle.

A second, so called, "helix-swap" experiment was undertaken (Wharton and Ptashne, 1985) which exchanged amino acids from the repressor of the Salmonella phage P22 to the 434 repressor. At this stage the structure of the 434 repressor had been determined but not the amino acids responsible for sequence-specific DNA-binding. The structure of the P22 repressor was unknown. By identifying the amino acids in the α-3 helix of the 434 repressor and predicting which would project away from the main body of the protein and therefore towards the operator, the amino acids responsible for binding to the DNA were tentatively identified. The 434 repressor, like several other phage DNA-binding proteins utilizing the helix-turn-helix motif, including the P22 repressor, has a conserved Gly at the start of the helix responsible for specific DNA-binding (Pabo and Sauer, 1984). Using the conserved Gly in the P22 repressor the amino acids predicted to be responsible for defining DNA-binding specificity were identified by an analogous method to that used for the 434 repressor. A hybrid gene was then constructed where the codons within the region coding for the α-3 helix in the 434 cl gene were swapped for the codons of the amino acids in the P22 repressor predicted to project towards the DNA. This involved changing just five amino acids: six were identified, but one Gln^33 is common to the 434 and P22 repressors (Fig.1.5). The hybrid 434 protein with the altered amino acids (434P22) was tested for its ability to
Figure 1.5. The amino acid sequences of the DNA binding helices of (A) the 434 and (B) the P22 repressor. The amino acids predicted to project towards the DNA in the P22 repressor are in bold type, in the hybrid repressor these amino acids replace the corresponding amino acids in the 434 repressor.

distinguish between the 434 and P22 operators. It was demonstrated in vitro by methylation interference that the 434P22 repressor had identical DNA-binding specificity to the wild-type P22 repressor. This was also demonstrated in vivo by the prevention of super-infection by λ phages which have the immunity region of the P22 phage in cells expressing the 434P22 repressor. This repressor was not able to confer immunity to λ phages with the immunity region of the 434 phage. This demonstrates that 434P22 is a hybrid protein with all the characteristics of the wild-type 434 repressor except for DNA-binding specificity, which is that of the P22 repressor.

A third helix-swap experiment, this time between the λ repressor and the 434 repressor produced a hybrid 434 repressor which was unable to bind either the 434 or the λ operator in vivo or in vitro (Wharton, 1985). The λ repressor has a slightly different mechanism of conferring DNA-binding specificity involving an N-terminal arm in addition to the helix-turn-helix motif (Pabo et al., 1982) and the changes
within the α-3 helix of the 434 protein are insufficient to fully alter the DNA-binding specificity of the 434 repressor to that of the λ repressor.

Another mutant 434 repressor with altered DNA-binding specificity has also been characterised (Wharton and Ptashne, 1987). This however, involves the alteration of only one amino acid in the repressor and only one base-pair in each operator half-site. Eight mutant 434 repressors were designed which contained amino acid substitutions; Ala, Leu, Ser, Met, Arg, His, Lys or Pro in place of Gln28. These were tested for their ability to repress β-galactosidase expression from tester strains where the lac promoter contained a 434 operator immediately downstream of the -10 box. The operators used were mutant 434 operators with point mutations in each half-site in each of the 4 conserved base-pairs, to give each of the twelve possible operator sequences. Only one combination revealed any repression, that of an Ala28 repressor with a 1T operator; TCAATATATATTGA. This repressor half occupied its operator in vitro at 3×10⁻⁸M, compared with 1×10⁻⁸M for the wild-type repressor/synthetic operator (ACAATATATATTGT) combination. Neither repressor bound detectably to the other operator at concentrations 150-fold higher. The authors propose that the hydrogen bond between Gln28 and the adenine in the wild-type operator/repressor combination is replaced with a van der Waals interaction between the methyl side group of the Ala in the mutant protein and the 5-methyl group on the thymine ring in the mutant operator.

1.3. Heterodimeric DNA-binding Proteins

The DNA-binding specificity of a bacterial repressor could be changed by altering the DNA-binding specificity per se to create an
altered homodimer or two different DNA-binding specificity monomers could be combined to form a heterodimer. If this were possible then the heterodimer would bind to an operator that was different from either of the operators for each homodimer.

1.3.1. 434 Repressor Heterodimers

Using the 434 repressor, Hollis et al. (1988) demonstrated the possibility of forming a heterodimer between two monomers of differing DNA-binding specificity. Using the 434 repressor in which the solvent-exposed amino acids of α-3 helix had been altered to those of the P22 repressor (434^P22) together with a wild-type 434 repressor they demonstrated in vitro that a heterodimer could be formed. Their experimental procedure involved purifying both the 434^P22 repressor and the wild-type 434 repressor and combining them in the presence of operators synthesised as complementary oligonucleotides.

DNasel-protection experiments showed that an operator, which linked the two operator half-sites at their centres of symmetry (Fig 1.6) was protected in the presence of purified 434 and 434^P22 repressors suggesting the formation of a heterodimer. This operator was not protected by either repressor as a homodimer. If point mutations were introduced in the conserved bases of either half-site the ability of the heterodimer to bind the operator was abolished. Whilst this indicates that a heterodimer could exist between these two 434 repressor types, it does not demonstrate that they could form a functional heterodimeric repressor in vivo.

Figure 1.6. The hybrid operator of Hollis et al. (1988) which is bound by a heterodimer of 434 and 434^P22 repressors in vitro. The 434 half-site is underlined the P22 half-site is overlined. Bases derived from the consensus operators are in bold type

ATTTAAGTT ATCTTGT
1.3.2. Heterodimeric Eukaryotic Transcription Factors

Eukaryotic transcription factors have increasingly been shown to form heterodimers *in vivo* which extend the range of sequences to which an individual transcription factor can bind or to alter the affinity with which it binds to its recognition sequence.

The proteins fos and jun have been shown to be transactivators of gene expression from promoters which contain a binding site for the transcription factor AP1 (Chiu *et al*., 1988; Rauscher, *et al*., 1988; 1989). In combination fos and jun are able to activate gene expression to a greater extent than either protein individually (Chiu *et al*., 1988). This has been shown to be due to the formation of a heterodimer between the two proteins mediated by the leucine zipper motif (O'Shea *et al*., 1989; Ransone and Verna, 1990).

Other examples of mammalian transcription factors acting as heterodimers have come from the product of the gene MyoD, a member of a family of genes responsible for regulating skeletal muscle development (Davis *et al*., 1987). This protein contains a helix-loop-helix motif which unlike the helix-turn-helix motif of bacterial repressors is responsible for dimerisation. An adjacent region of basic amino acids is responsible for DNA-binding (Murre *et al*., 1989b). The MyoD protein forms heterodimers with a number of other DNA-binding proteins containing the helix-loop-helix motif, specifically, the non-cell type specific factors E12 and E47 (Murre *et al*., 1989a). E12 and MyoD form a heterodimer even in the absence of DNA, and *in vitro* the heterodimers form a complex with their binding site at higher affinity than the homodimers as determined by gel shift assay (Murre *et al*., 1989b). The proteins E12 and E47 are also able to form heterodimers with the *Drosophila* proteins daughterless and acetate-schute T3, two cell determination factors (Murre *et al*., 1989b). This
1-Introduction

suggests that a common mechanism may occur through many different species where DNA-binding is regulated by the formation of heterodimers utilizing the helix-loop-helix motif.

Another protein has been isolated which is able to form heterodimers with MyoD, E12 and E47 although this has quite a different effect on the transcriptional activation function of these proteins (Benzer et al., 1990). This protein, Id, has a marked similarity in amino acid sequence around the region of the helix-loop-helix domain to these proteins and is expressed in a wide variety of cell types, although at widely different levels. *In vitro* the formation of heterodimers between Id and MyoD, E12 and E47 results in protein complexes which are deficient in DNA-binding. *In vivo* over-expression of Id in muscle cell lines prevents expression of the MCK promoter, presumably by the formation of heterodimers with, and therefore titrating out, the enhancer-binding proteins acting at the MCK promoter.

Pongubala and Atchison (1991) also observed this effect with the developmentally regulated immunoglobulin Kappa 3' enhancer, again possibly indicating the formation of a transcription factor/Id heterodimer which is deficient in DNA-binding, thereby reducing transcription.

This phenomenon is not restricted to higher eukaryotes: the mating type locus in yeast (*MAT*) encodes several DNA-binding proteins which regulate the expression of genes in the three cell types of *Saccharomyces cerevisiae*: the a and α haploid cell types and the a/α diploid cell type (Herskowitz, 1989). One product of *MAT*, α2, functions in two ways: in α cells it represses the a-specific genes by binding to operators within the a-specific genes as a dimer (Johnson and Herskowitz, 1985) and in a/α diploid cells it acts with another *MAT*
gene $a_1$ to repress another set of genes, the haploid genes (Strathern et al., 1981, Goutte and Johnson, 1988). An $\alpha 2$ homodimer binds to its operator, consisting of two half-sites 25bp apart on opposite faces of the DNA (Sauer et al., 1988), with the central part of the operator occupied by a dimer of a second DNA-binding protein MCM1 (Keleher et al., 1988, Passmore et al., 1989, Keleher et al., 1989). The consensus sequence for the $\alpha 2$ binding site has been identified (Jarvis et al., 1988). Recently (Dranginis, 1990), it has been shown that the interaction between $\alpha 2$ and $a_1$ takes the form of a heterodimer, and a heterodimer of $\alpha 2$/$a_1$ has been shown to bind to a sequence different from the $\alpha 2$ homodimer (Miller et al., 1985) which does not involve any other proteins (Fig. 1.7).

The two proteins bind on the same face of the DNA helix and the centre of the consensus sequence is not conserved, suggesting that $a_1$ does not substitute for MCM1 and that $a_1$ and $\alpha 2$ are sufficient to bind the operator. This is a very good example of the binding specificity of DNA-binding proteins being altered by the formation of heterodimers.

If this could be applied to bacterial repressors then the range of operators to which they could bind would be increased, which for the regulation of gene expression at pre-existing operators or pseudo-operators would be greatly advantageous.

\[
\text{CATGTAATTACCNAATAAGGAATTTACATGN}T
\]

\[
\text{TCRTGTTNNWNANNTACATCA}
\]

Figure 1.7. The operators for the MAT DNA-binding proteins the $\alpha 2$ homodimer (upper) and the $\alpha 2$/$a_1$ heterodimer (lower). Ambiguity codes used are; $N =$ any base, $W= A$ or $T$ and $R= A$ or $G$
1.4. The Use of Bacterial Repressors to Regulate Gene Expression in Eukaryotes

If bacterial repressors were to bind within a eukaryotic promoter then transcription would be expected to be reduced if the repressor interfered with the binding of the eukaryotic transcription factors or if it interfered with the binding or passage of RNA polymerase. Bacterial repressors have indeed been shown to be able to repress gene expression in eukaryotic promoters when their operators have been inserted in suitable positions.

The $\textit{lexA}$ repressor from \textit{E. coli} has been shown to bring about repression of the expression of reporter gene constructs when $\textit{lexA}$ operators have been inserted into the promoters. In the \textit{Saccharomyces cerevisiae} gene $\textit{GAL1}$ the introduction of a $\textit{lexA}$ operator between the upstream activator sequence (UAS) and the TATA box enabled transcription to be blocked by the $\textit{LexA}$ repressor (Brent and Ptashne, 1984).

The $\textit{lexA}$ operator/repressor system has also been shown to function in mammalian cells. Smith \textit{et al.} (1988) demonstrated that the insertion of $\textit{lexA}$ operators into the thymidine kinase ($\textit{tk}$) promoter could repress expression of a chloramphenicol acetyl transferase (CAT) reporter gene upto 10-fold in mouse cells.

The $\textit{lac}$ repressor, which has been used in several studies of repression in eukaryotic cells, has advantages over some other repressors as it is inducible with lactose analogues such as isopropyl-$\beta$-D-thiogalactoside (IPTG). The $\textit{lac}$ operator/repressor system has now been shown to function in both mammalian and plant cells. Hu and Davidson (1987) showed that in mouse cells the insertion of a $\textit{lac}$ operator between the initiation codon and the transcription start site,
between the transcription start point and the TATA box or upstream of the TATA box could block transcription in the presence of the *lac* repressor and that this block was relieved in the presence of the inducer IPTG. This has also been reported by Brown *et al.* (1987) and Figge *et al.* (1988). A further development of this was shown by Deuschle *et al.* (1990) who demonstrated that the SV40 large T antigen gene could be repressed by the binding of a *lac* repressor to a *lac* operator inserted in the intron of the gene, thus demonstrating that a bacterial repressor is an effective block to transcriptional elongation by RNA pol II. Recently it has been shown that an inducible *lac* operator/repressor system can function efficiently in tobacco cells (Wilde *et al.*, 1992).

Another inducible repressor, the *tet* repressor, has also been shown to function in tobacco cells. It has been demonstrated that the *tet* repressor is able to repress the expression of a plant promoter efficiently and that the repression can be relieved in the presence of tetracycline (Gatz and Quail, 1988; Gatz *et al.*, 1991)

The *tet* repressor has also been shown to block RNA polymerase III transcription in yeast cells (Dingermann *et al.*, 1992).
1-Introduction

1.5. Aims of the Project

Whilst the interactions between the 434 repressor and its operator have been extensively studied in vitro, less information is available on these interactions in vivo.

We wish to be able to use the 434 repressor to regulate gene expression by binding to pre-existing sequences, which resemble the 434 operator but may have small sequence changes, a pseudo-operator, in the promoter, or maybe within the transcribed region, of a target gene.

The aims of this study are therefore: (1) To assess the ability of a heterodimer to repress gene expression in vivo. (2) To assess the effect the sequence of the central non-contacted bases of the operator have on the repression level in vivo. (3) To determine whether the 434 repressor is able function within a eukaryotic nucleus. The experimental approaches and the results obtained while pursuing these aims are described in the following chapters.
CHAPTER 2.

Repression by a Heterodimeric Repressor in *Escherichia coli.*
2.1. INTRODUCTION

By alteration of codons in the DNA coding for the DNA-binding helix of the 434 repressor Wharton and Ptashne (1985) created a mutant repressor which has the DNA-binding specificity of the P22 repressor. In vitro the wild-type 434 repressor and the 434\textsuperscript{P22} repressor are able to form a heterodimer which specifically binds to a hybrid operator consisting of one half-site from each of the 434 and P22 operators (Hollis et al., 1988). Results in this chapter demonstrate that such a heterodimer is able to bind to a hybrid operator and efficiently repress gene expression in E. coli.

Results presented in this chapter have recently been published (Webster et al., 1992).
2.2. RESULTS

2.2.1. Plasmid Constructions

2.2.1.1. Reporter Plasmids containing the lacZ gene.

In order to carry out repression studies it was necessary to have a plasmid in which operators to be studied could be cloned into a unique restriction site at a suitable location within a promoter controlling expression of a reporter gene. A low copy-number vector was chosen to prevent reduction of repression by a titration effect and to avoid overburdening the cell with highly expressed genes at high copy-number.

2.2.1.2. Construction of pADATc

This construction is shown diagrammatically in figure 2.1. A lac promoter which contains a SalI site engineered immediately 3' to the -10 box and the 5' end of the lacZ gene on a HindIII-EcoRI fragment from pRW283 (Wharton and Ptashne, 1987) and a HindIII-SalI fragment from pSKS107 (Casadaban et al., 1983) carrying the lac operon were ligated into pLG339P (Stocker et al., 1982), opened at its EcoRI and SalI sites, in a three-component ligation. The lac promoter and operon were cut from this plasmid on an NcoI fragment and recloned into the NcoI site of pLG339P, to give a plasmid with the lac operon and determinants for both kanamycin- and tetracycline-resistance. This created a 400bp duplication of the region between the EcoRI and NcoI sites of the pLG339P plasmid. To remove one of these the plasmid was cut with SalI, the lac operon removed and the vector religated. The vector DNA was digested with EcoRI and NcoI and the ends filled-in and religated. This retained the EcoRI site which was then removed by digestion with EcoRI, filling in the overhanging ends and religating. The lac operon was then reintroduced by ligating the appropriate SalI
Figure 2.1. Construction of pADΔTc

Repression by a Heterodimeric Repressor

Figure 2.1. Construction of pADΔTc
2 - Repression by a Heterodimeric Repressor

**EcoRI/NcoI**
fill in ends and religate

---

**Sall**, reintroduce lact operon
from pAD6

---

**BamHI/XmaIII**
introduce a small oligo to reclose vector
fragment into the SalII site in the vector to create pAD13 (pAD13 was a gift from D. Pioli, ICI Pharmaceuticals). Construction of pADΔTc (Figure 2.1) was necessary in order to remove a second SalII site in the tetracycline-resistance gene, thus leaving the SalII site in the lac promoter unique for the cloning of operators. This was achieved by deleting DNA within the tetracycline resistance gene between the XmaIII and the BamHI sites and replacing this region with a small 30bp fragment of non-functional DNA from annealed oligonucleotides with compatible ends.

2.2.1.3. pAD plasmids containing operators.

Operators under study were synthesised as complementary oligonucleotides with SalII-compatible ends and cloned into the SalII site in the lac promoter of pADΔTc. The oligos were constructed such that the SalII site was destroyed, thus allowing the removal of non-operator containing plasmids following ligation by digestion with SalII. All operator-containing constructs were checked by sequencing.

2.2.1.4. Plasmids expressing the 434 repressor.

For the expression of the 434 cl gene in E. coli a pUC-based vector was constructed, pPLRT1. The construction is represented diagrammatically in figure 2.2. The λpL promoter and the rrnB T1 terminator were inserted separately as annealed oligonucleotides into the polylinker of pUC19, between the HindIII and EcoRI sites: both sites were eliminated by the ligations. A small portion of the 5' end of the 434 cl gene is also included in the oligonucleotides for the λpL promoter. The remainder of the 434 cl coding region was then cut from pRP42-76 on an EcoRI/Sau3A fragment and cloned between the EcoRI and BglII sites between the 5' end of the 434 cl gene and terminator:
Figure 2.2. Construction of pPLRT1

HindIII rrnB T1 terminator

as annealed oligonucleotides

EcoRI/HindIII

λpL promoter as annealed
oligonucleotides

EcoRI/BglII

434 repressor gene from
pRP42-76
the BglII site was not retained by the ligation. The plasmid pRP42-76 is based on pRP42 (Lauer et al., 1981) with restriction sites for KpnI and XmaIII silently engineered into the 434 cl gene by site-directed mutagenesis to allow the replacement of the DNA coding for the α-3 helix of the protein (pRP42-76 was obtained from D. Pioli, ICI Pharmaceuticals). The KpnI site was engineered as described by Wharton and Ptashne (1985) and the XmaIII site was engineered as described by Wharton (1985). The λpL promoter allows high-level expression (Remaut et al., 1981) and the rrnB T1 terminator causes transcriptional termination in both directions (Brosius, 1984), preventing read through from the lac promoter and stopping transcription from the λpL promoter outside the 434 cl gene. The plasmid pPLRT2 was made by removal of the DNA between the KpnI and the XmaIII sites in pPLRT1 and its replacement with annealed oligonucleotides coding for the altered amino acids necessary to create the 434 cl gene-product with the binding-specificity of the P22 repressor (434P22) described by Wharton and Ptashne (1987).

2.2.1.5. Plasmids Expressing Two Repressor Genes.

This plasmid construction is represented diagrammatically in figure 2.3. The plasmid pPLRT2 was cut with SmaI and the tetracycline-resistance gene from pBR322 on a Ball/SspI fragment ligated in. The 434 repressor expression cassette from pPLRT1 was then ligated into the NcoI site 3’ to the tetracycline-resistance gene to give pPL2RTc (Fig 2.4). Both repressor genes were inserted in the same orientation. In this orientation, recombination between any part of the two repressor genes would result in the deletion of the DNA between them. Continued selection with tetracycline prevents any plasmid in which this has occurred from giving rise to a viable cell.
2 - Repression by a Heterodimeric Repressor

Figure 2.3. Construction of pPL2RTc

- 33 -
2.2.2. Western Analysis of 434 Repressor Expression

In order to determine the levels of 434 repressor expressed from the plasmids pPLRT1, pPLRT2 and pPL2RTc protein extracts were prepared from strains of E. coli carrying each plasmid and analysed by sodium dodecyl sulphate (SDS) polyacrylamide gel electrophoresis and Western blotting. The results (Fig 2.5) show that the 434 and 434\(^{P22}\) repressors, strains carrying pPLRT1 and pPLRT2 respectively, were expressed at approximately 0.5% of cell protein. Expression of a combination of the 434 and 434\(^{P22}\) repressors from pPL2RTc was also at approximately 0.5% of cell protein. However it is impossible to determine what proportion is made up of wild-type 434 repressor and what proportion is 434\(^{P22}\) repressor as the antiserum used does not distinguish between them.

![Western analysis of 434 and 434\(^{P22}\) expression from the plasmids pPLRT1, pPLRT2 and pPL2RTc. Lane a, 20\(\mu\)g of protein from E. coli. Lane b, 20\(\mu\)g protein from a strain carrying pPLRT1. Lane c, 20\(\mu\)g of protein from a strain carrying pPLRT2. Lane d, 20\(\mu\)g protein from a strain carrying pPL2RTc. Lane e, 200ng purified 434 repressor. Lane f, 100ng purified 434 repressor. Lane g, 50ng purified 434 repressor. All E. coli strains used are 6300A lacU169. Both the anti-434 antiserum and the purified 434 repressor were kind gifts from D. Pioli ICI Pharmaceuticals.](image-url)
2.2.3. Experimental Design

In order to test the feasibility of obtaining a functional heterodimeric repressor in vivo, it was necessary to co-express the wild-type 434 cl gene and a 434\textsuperscript{P22} cl gene in the same bacterial cells. This was achieved by the construction of pPL2RTc (see plasmid constructions). The activity of the heterodimer could then be detected by the ability to repress the expression of a lacZ reporter gene from a promoter containing a chimeric operator sequence with 434 and P22 half-sites. To compare the effects of a heterodimer with the results seen in vitro, the same operators were used as described by Hollis et al. (1988), as shown in Table 2.1.

2.2.4. Measurement of Repression in vivo.

*E. coli* 6300ΔlacU169 was cotransformed with a pAD-type plasmid containing the operator of interest and a pUC-based plasmid expressing the appropriate repressor and plated onto Luria Agar with appropriate antibiotics and 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal). This strain of *E. coli* lacks the entire lac operon and the chromosomal lacI gene, so that the lac repressor cannot interfere with lacZ expression and the lacZ gene-product can only be encoded by the pAD plasmid. Single blue colonies were picked and used to determine β-galactosidase activity.

Initially the effect of single repressors on the operators of 434 and P22 was investigated. This measurement included controls in which the effects of both 434 and 434\textsuperscript{P22} on β-galactosidase expression from pADΔTc were investigated. It was discovered that there was a decrease in lacZ expression with 434\textsuperscript{P22} repressor, even in the absence of an inserted operator (Figure 2.6). We presume that this is due to the
fortuitous presence of an operator-like site within the lacZ gene. Computer analysis of the DNA sequence of the lacZ gene reveals a P22 operator-like sequence at codons 797-802, suggesting the possibility that binding of the repressor at this site might bring about repression of β-galactosidase expression. Conversely both the 434 homodimer and the 434/434P22 heterodimeric repressor bring about slight increases in the expression of β-galactosidase, the reasons for which are unclear. In order to take these effects into account the data obtained from all repression studies were always normalised according to the repressor used in the investigation.
Figure 2.6. Effects of the 434, 434P22 and heterodimeric repressors on lacZ expression from a plasmid which contains no artificially introduced operators. Expression from the pAD plasmid in the absence of repressor is taken as 100%.
2.2.4.1. Repression by Heterodimeric Repressors.

The combination of a hybrid operator, HybOp1, and the heterodimer of the 434 and 434\textsuperscript{P22} repressors showed a repression level of about 2 fold (Figure 2.8). However, this operator contained the weak 434 half-site of 434Op1. The low level of repression could therefore be a reflection of the weaker binding of the 434 repressor to its half-site rather than the poor performance of heterodimers. Testing repression with this hybrid operator/heterodimer revealed little difference with the operator inserted in either orientation (Table 2.1; HybOp1 and HybOp2).

The improvement in the 434 operator described above was then applied to the 434 half-site of the hybrid operator and a new operator, HybOp3, was synthesised which included a GC pair at position 5 of the hybrid operator. The levels of repression seen with this hybrid operator and the heterodimer were improved dramatically compared to the first hybrid operators used (Table 2.1 and Figure 2.8). Repression increased by more than 6-fold to give approximately 95% reduction in \( \beta \)-galactosidase activity compared to the control. This result demonstrates that heterodimers are able to bring about repression, at a hybrid operator site, to a level similar to that achieved by either homodimer at its cognate operator.
Table 2.1. Repression achieved using the 434, P22 and hybrid operators with homo- and heterodimeric repressors. Repression is expressed as a percentage reduction in β-galactosidase activity with respect to that obtained with a control plasmid containing the operator under investigation adjusted for any effect of the repressor on expression from the parent plasmid, pADATc. The results are an average of at least 4 experiments. The sequence of the operators is the sequence inserted at the SalI site in pADATc, the complementary oligonucleotides used to insert operators carried a 4bp overhang at each end compatible with the overhang generated by the enzyme SalI. A space is included between the left and right sequences of the operator to show the extent of the two half sites. The S.D. (standard deviation) is calculated using formulae for the propagation of error, see Materials and Methods.

<table>
<thead>
<tr>
<th>Operator</th>
<th>Repressor</th>
<th>Sequence</th>
<th>%Repression</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>434Op1</td>
<td>434</td>
<td>ACAATAT ATATTGT</td>
<td>80</td>
<td>2.5</td>
</tr>
<tr>
<td>434Op2</td>
<td>434</td>
<td>ACAAGAT ATCTTG</td>
<td>95</td>
<td>0.8</td>
</tr>
<tr>
<td>434Op3</td>
<td>434</td>
<td>ACAAAAT ATTTTG</td>
<td>84</td>
<td>4.3</td>
</tr>
<tr>
<td>434Op4</td>
<td>434</td>
<td>ACAACAT ATGTTG</td>
<td>59</td>
<td>6.3</td>
</tr>
<tr>
<td>OR1</td>
<td>434</td>
<td>ACAAGAA AGTTTG</td>
<td>97</td>
<td>0.7</td>
</tr>
<tr>
<td>P22</td>
<td>434P22</td>
<td>ATTTAAGAT ATCTTAAT</td>
<td>94</td>
<td>0.8</td>
</tr>
<tr>
<td>HybOp1</td>
<td>Hetero</td>
<td>ACAATAT ATCTTAAT</td>
<td>60</td>
<td>5.8</td>
</tr>
<tr>
<td>HybOp2</td>
<td>Hetero</td>
<td>ATTTAAGAT ATATTG</td>
<td>57</td>
<td>4.3</td>
</tr>
<tr>
<td>HybOp3</td>
<td>Hetero</td>
<td>ACAAGAT ATCTTAAT</td>
<td>95</td>
<td>0.5</td>
</tr>
<tr>
<td>HybOp3</td>
<td>434</td>
<td>ACAAGAT ATCTTAAT</td>
<td>45</td>
<td>2.4</td>
</tr>
<tr>
<td>HybOp3</td>
<td>434P22</td>
<td>ACAAGAT ATCTTAAT</td>
<td>0</td>
<td>3.1</td>
</tr>
</tbody>
</table>
Figure 2.8. Repression of β-galactosidase activity on plasmids containing hybrid-operators by a heterodimeric repressor consisting of wild-type 434 repressor and 434 repressor with the binding specificity of the P22 repressor. Controls are also shown where hybrid operators are repressed by homodimeric 434 and 434P22 repressors. The sequences of the operator are given in Table 2.1.
2.3. DISCUSSION

The previous work on the 434 operator/repressor system using altered binding-specificity repressors (Wharton et al., 1984; Wharton and Ptashne, 1985; 1987), relied on the same synthetic 434 operator that was also used in the X-ray crystallography studies. During the study presented here it became clear that the synthetic 434 operator ACAATATATATTGT represents a relatively weak operator sequence and was not appropriate for this work. The wild-type operator OR1 is bound more tightly by the 434 repressor than is the synthetic operator. This is in part due to a central run of A's in OR1 which were found by Koudelka et al. (1987) to improve binding. In addition, interactions between the amino acid at position 33 and the fifth base of the operator may influence binding affinity. Based on these data the synthetic operator was altered to include bases present in the wild-type operator, OR1, at the inner-most contacted positions, 5 and 10. The operator 434Op2 includes the GC base-pair found at position 5 of OR1, together with complementary bases at position 10, and these changes improved the affinity of the repressor for the operator, as was found by Koudelka et al. (1988). No improvement is seen with 434Op3, incorporating the TA base pair found at position 10 of OR1, suggesting that the G of OR1 is responsible for the improved affinity. To ensure that the improvement seen with the GC base pairs in 434Op2 is due to specific interactions, a final 434 operator was constructed, 434Op4, in which the positions of the GC base-pairs were reversed. This operator demonstrated a reduced affinity for the repressor, suggesting that the improvement seen with 434Op2 is due to improved amino acid/base-pair contacts, rather than a change in the flexibility of the operator facilitating better binding. No
significant repression was detected with either homodimer in repression studies with a heterologous operator (data not shown).

Measurement of repression of $lacZ$ expression from pAD plasmids with the $434^{P22}$ repressor presents some complications. An unexpected reduction in $\beta$-galactosidase expression was obtained in the absence of inserted operators. Examination of the $lacZ$ gene reveals a sequence similar to the consensus P22 operator at codons 797-802, suggesting the possibility that binding of the repressor at this site might bring about repression of $\beta$-galactosidase expression. It has previously been established that repressor/operator interaction can repress expression of a gene by interference with transcription initiated from an upstream site (Deuschle et al., 1990). Conversely, both the 434 homodimer and the $434/434^{P22}$ heterodimeric repressor bring about slight increases in the expression of $\beta$-galactosidase, the reasons for which are unclear.

Initially the extent of repression by heterodimers was measured at 60% repression, compared to 80-94% from the homodimer repressor/operator combinations. However, applying the operator sequence improvements shown for the 434 operator by introducing a GC pair in the 434 half-site of the hybrid operator showed significantly improved repression. This resulted in a repression level of 95% for the hybrid operator HybOp3. However, inclusion of this GC pair in the hybrid operator also improved the binding of the 434 homodimer, as judged by repression data (Figure 2.8). It is difficult to estimate what proportion of the repression seen in the heterodimeric situation is due to 434 homodimers, as cells expressing both the 434 and $434^{P22}$ repressors will contain a mixed population of the two homodimers and the heterodimer. The level of repressor expressed from both pPLRT1 and pPLRT2 has been estimated at 0.5% of cell protein by
Western analysis: this approximates to 6500 repressor monomers per cell. The 434 and 434\textsuperscript{P22} repressors are indistinguishable by Western analysis making it impossible to estimate the levels of either repressor in strains carrying the plasmid pPL2RTc. However, the level of expression of 434 type repressors from this plasmid is 0.5\% of cell protein. It is clear however, that the major proportion of the repression is due to the heterodimer's binding at the hybrid operator, HybOp3. The level of heterodimer within the cell is expected to be saturating and precise expression levels of each monomer are not critical, since given the high levels of repressor fairly wide fluctuations would not be expected to affect repression. It has been shown that a heterodimer is able to bring about repression levels similar to those achievable with either homodimer, thereby establishing the ability of heterodimers to function efficiently \textit{in vivo}. 

- 46 -
CHAPTER 3.

The Effect of the Non-contacted Bases of the 434 Operator on the Levels of Repression In Vivo.
3.1. INTRODUCTION

Previous work has shown that, *in vitro*, the sequence of the central non-contacted bases of the 434 operator has a large effect on the affinity of the operator for the repressor (Koudelka *et al.*, 1987). The presence of G:C or C:G base-pairs in the centre 4 positions of the operator reduces the affinity of the operator for the repressor with reference to a synthetic operator, ACAATATATATTGT. This effect is most pronounced in the centre 2 base-pairs of the operator, where a reduction in affinity of up to 50-fold can occur if both base-pairs are G:C or C:G. This effect has been postulated to be due to changes in the flexibility of the central region of the operator, either through increased torsional rigidity imparting resistance to the overwinding (Koudelka and Carlson, 1992) or to the "bendability" of the operator (Koudelka, 1991).

X-ray crystallographic analysis of the operator/repressor complex has revealed that the operator DNA is overwound and bent when in complex with the repressor. Analysis at high resolution (Aggarwal *et al.*, 1988) shows that the minor groove of the DNA is compressed in the centre of the operator, compared to uniform B-DNA. This compression is not due to the bending of the DNA that is also associated with repressor binding, as the DNA in the centre of the operator is relatively straight. The operator DNA is bent 2 to 3 base-pairs out from the centre giving the operator the appearance of an arc. The ease with which these distortions can be made is sequence specific (Koudelka and Carlson, 1992) and therefore influences the affinity of the repressor.

The reduction in the affinity of the operator for the repressor can be overcome to some extent by altering the repressor's need to distort the DNA to such a great extent. If a mutant repressor was available
which overwound the central bases of the operator to a lesser extent than the wild-type repressor, then that repressor would have an increased tolerance to the presence of G:C or C:G base-pairs in the central region of the operator. Two repressors have been described which have these characteristics, both due to a mutation at Phe$^{44}$, one changed to Ala$^{44}$ (Koudelka et al., 1988) and one to Trp$^{44}$ (Koudelka and Carlson, 1992). However, these repressors have differing affinities for a reference operator, ACAATATATATTGT, in vitro, the Ala$^{44}$ mutation apparently not affecting the dissociation constant whereas the Trp$^{44}$ mutation lowers it by almost 7-fold.

In this chapter the effect of the sequence of the central 4 bases of the operator on the level of repression in vivo is investigated, together with the effect of a 434 repressor carrying the Ala$^{44}$ mutation on these levels of repression.
3.2. RESULTS

3.2.1. Plasmid Constructions

3.2.1.1. Repressor Plasmids

The plasmid pPLRT3 was constructed to express the 434 repressor containing the Ala\textsuperscript{44} mutation. This was achieved by cutting most of the 434 gene and the \textit{rrnB} T1 terminator from pPLRT1 on a \textit{KpnI}/\textit{SphI} fragment and cloning it into M13mp19 at the \textit{KpnI} and \textit{SphI} sites. The remainder of the vector was retained for the return of the mutated repressor gene. The codon for Phe\textsuperscript{44} (TTT) was then altered to the Ala codon (GCT) by site-directed mutagenesis using the method of Kunkel \textit{et al.} (1987). Recombinant plaques were tested for the presence of the mutation by sequencing. The 434 gene/\textit{rrnB} T1 terminator carrying the Ala\textsuperscript{44} mutation was then cut from the RF DNA of the recombinant M13 on a \textit{KpnI}/\textit{SphI} fragment and returned to the vector to give pPLRT3.

3.2.1.2. Reporter Plasmids Carrying Operators

Operators under study were synthesised as complementary oligonucleotides with \textit{SalI}-compatible ends and cloned into the \textit{SalI} site in the \textit{lac} promoter of pAD\textDelta Tc. The oligos were constructed such that the \textit{SalI} site was destroyed, thus allowing the removal of non-operator containing plasmids following ligation by digestion with \textit{SalI}. All operator-containing constructs were checked by sequencing.
3.2.2. Western Analysis of Strains Expressing Repressors

In order to compare the level of expression of the repressors from the plasmids pPLRT1 and pPLRT3 protein extracts from strains of *E. coli* 6300ΔlacU169 carrying each plasmid were analysed by SDS PAGE and Western blotting. The results in figure 3.1 show that each strain expresses similar levels of 434 repressor protein. The results in chapter 2 demonstrated that pPLRT1 was able to direct expression of the 434 repressor at 0.5% of cell protein: the results presented here demonstrate that pPLRT3 also expresses 434 repressor at approximately 0.5% of cell protein.

![Figure 3.1](image)

**Figure 3.1.** Western analysis of 434 repressor protein expressed from *E. coli* 6300ΔlacU169 carrying either pPLRT1, b, or pPLRT3, c. Lane a contains *E. coli* 6300ΔlacU169 with no introduced plasmids. 20µg of total protein was loaded in each lane.
3.2.3. Experimental Design

In order to test repression in vivo, plasmids expressing the 434 or 434$^{1344}$ repressor were introduced, together with a pAD plasmid containing the operator of interest, into E. coli 6300ΔlacU169 by electroporation, and the β-galactosidase activity determined for each strain.

3.2.4. Measurement of Repression In Vivo

Initially the effect of the wild-type 434 repressor was investigated to determine the effect of G and C bases in the centre 4 bases of 434Op1 and 434Op2. The results obtained for 434Op1 were generally in accordance with the findings of Koudelka et al. (1987), although the effect of the inclusion of G:C and C:G base-pairs in the centre 4 bases of the 434 operator is less pronounced in vivo. The inclusion of G or C bases at positions 7 and 8 greatly reduces the level of observed repression, from 80% to approximately 20%. This effect is less pronounced if G and C bases are included at positions 6 and 9, when repression is about 60-70% (Table 3.1, Fig 3.2). These data correspond to a relative reduction in repression levels of approximately 1.75-fold for 434Op1G6C9 and G9C6 and 4-fold for 434Op1G7C8 and G8C7. This compares to 5-fold and 50-fold respectively for reductions in the disassociation constant in vitro (Koudelka et al., 1987).

The situation observed with 434Op2 (Table 3.1, Fig 3.3) is similar to 434Op1, although the reduction in repression is less pronounced. Again the inclusion of G and C bases at positions 7 and 8 of the operator has the greatest effect on the repression level, reducing it to around 40%. The presence of G and C bases at positions 6 and 9 maintains a relatively high percentage level of repression, reducing...
repression to approximately 84%. The inclusion of single G and C bases within the centre 4 bases of 434Op2 was also investigated. At any of the 4 positions 6-9 the percentage repression level is reduced only weakly. However, examination of the change to the fold repression level reveals that the presence of G and C bases in the centre of the operator 434Op2 has a stronger negative effect on the affinity of the repressor compared to 434Op1. For 434Op2G6C9 and G9C6 there is a 3-fold reduction in repression and for 434Op2G8C7 there is an 11-fold reduction in repression. For the inclusion of single G and C bases into the operator the reduction in repression ranges from 0.7-fold to 5-fold. There are no data to compare the inclusion of G:C and C:G base-pairs in the centre of 434Op2 with the effect on repressor binding in vitro.
Table 3.1. The effect of G:C and C:G base-pairs in the centre 4 positions of the 434 operator on the levels of repression in vivo. The G or C base included is underlined in each operator. The sequence given is the sequence inserted at the SalI site in pADΔTc, complementary oligonucleotides used to insert the operator carry a 4 base overhang compatible with the overhang generated by the restriction endonuclease SalI. S.D. refers to the standard deviation of the percent repression calculated by formulae for the propagation of error, see Materials and Methods.
Figure 3.2. The Effect of G:C and C:G Base-pairs in the Centre 4 positions of 434Op1 on the Levels of Repression in vivo.
Figure 3.3. The Effect of G:C and C:G base-pairs in the Centre 4 bases of 434Op2 on the Levels of Repression in vivo.
The investigation of $434^{\text{Ala44}}$ in the repression of G and C base-containing operators revealed a marked difference from \textit{in vitro} results. Here Koudelka \textit{et al.} (1988), using a reference operator ($434^{\text{Op1}}$), showned that $434^{\text{Ala44}}$ had the same affinity as the wild-type repressor. When G and C bases were included in the centre two positions of the target operator $434^{\text{Ala44}}$ bound with a 10-fold higher affinity than the wild-type repressor. \textit{In vivo}, $434^{\text{Ala44}}$ represses constructs containing $434^{\text{Op1}}$ to a lesser extent than the wild-type $434$ repressor; repression is approximately 65% compared to 80%. This is also seen with $434^{\text{Op1G6C9}}$ and $434^{\text{Op1G9C6}}$, where the $434^{\text{Ala44}}$ repressor produces only 50% repression compared to 60-70% repression for the wild-type repressor. Repression in constructs containing $434^{\text{Op1G7C8}}$ and $434^{\text{Op1G8C7}}$ is however to a higher level than the wild-type repressor, approximately 33% compared to 18-28%.

The most significant difference, compared to the results obtained \textit{in vitro}, is seen with $434^{\text{Op2}}$. Koudelka \textit{et al.} (1988) reported the $434^{\text{Ala44}}$ bound this operator with a 4-fold higher affinity \textit{in vitro} than the wild-type repressor, corresponding to a $K_d$ of $3 \times 10^{-10} \text{M}$ for $\text{Ala44}$ compared to $1.3 \times 10^{-9} \text{M}$ for wild-type. \textit{In vivo} $434^{\text{Ala44}}$ represses constructs containing $434^{\text{Op2}}$ approximately 90% compared to the 95% seen for wild-type $434$. Hence \textit{in vivo}, $434^{\text{Ala44}}$ produces weaker repression than the wild-type $434$ repressor despite its higher affinity for the operator \textit{in vitro}.

Data obtained for operators based on $434^{\text{Op2}}$ containing G:C and C:G base-pairs shows an effect with $434^{\text{Ala44}}$ similar to that seen with operators based on $434^{\text{Op1}}$. The operators $434^{\text{Op2G6C9}}$ and $434^{\text{Op2G9C6}}$ show reduced repression with $434^{\text{Ala44}}$ but the operator $434^{\text{Op2G8C7}}$ shows a large jump in repression when the $434^{\text{Ala44}}$ repressor is used.
Interestingly the 434<sup>Ala<sub>44</sub></sup> repressor did not cause the small increase in expression from pADΔTc that was seen with the wild-type 434 repressor. This again points to its differing interactions with DNA.

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>% Repression, by:</th>
<th>fold change</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>434</td>
<td>434&lt;sup&gt;Ala&lt;sub&gt;44&lt;/sub&gt;&lt;/sup&gt;</td>
</tr>
<tr>
<td>434Op1</td>
<td>ACAATAT ATATTGT</td>
<td>80</td>
<td>65</td>
</tr>
<tr>
<td>434Op1G6C9</td>
<td>ACAATGT ACATTGT</td>
<td>73</td>
<td>56</td>
</tr>
<tr>
<td>434Op1G7C8</td>
<td>ACAATAG CTATTGT</td>
<td>18</td>
<td>33</td>
</tr>
<tr>
<td>434Op1G8C7</td>
<td>ACAATAC GTATTGT</td>
<td>28</td>
<td>34</td>
</tr>
<tr>
<td>434Op1G9C6</td>
<td>ACAATCT AGATTGT</td>
<td>61</td>
<td>54</td>
</tr>
<tr>
<td>434Op2</td>
<td>ACAAGAT ATCTTGT</td>
<td>95</td>
<td>91</td>
</tr>
<tr>
<td>434Op2G6C9</td>
<td>ACAAGGT ACCTTGT</td>
<td>84</td>
<td>78</td>
</tr>
<tr>
<td>434Op2G8C7</td>
<td>ACAAGAC GTCTTGT</td>
<td>43</td>
<td>75</td>
</tr>
<tr>
<td>434Op2G9C6</td>
<td>ACAAGCT AGCTTGT</td>
<td>85</td>
<td>78</td>
</tr>
</tbody>
</table>

Table 3.2. A comparison of repression by the wild-type 434 repressor and the 434<sup>Ala<sub>44</sub></sup> repressor at variants of the operators 434Op1 and 434Op2. The space in the sequence denotes the division between each half-site. Fold change refers to the ratio of the fold repression between the wild-type repressor and the Ala<sub>44</sub> repressor.
Figure 3.4. A Comparison of the Effect of the wild-type 434 repressor (pPLRT1) and the 434Ala44 repressor (pPLRT3) on the Levels of Repression at variants of the Operator 434Op1.
Figure 3.5. A comparison of the effect of the wild-type 434 repressor (pPLRT1) and the 434Ala44 repressor (pPLRT3) on the levels of repression at variants of the operator 434Op2.
3.3. DISCUSSION

The reason for examining the effect of the sequence of the centre 4 bases of the 434 operator on the level of repression in vivo is to determine what sequence would represent a suitable target for repression of genes by binding a repressor to a pre-existing operator sequence. The results have demonstrated a number of important effects that such sequences have on the levels of repression. It is immediately evident that operators containing G:C or C:G base-pairs in both the centre two bases of the operator represent very poor sites for repressor binding irrespective of the sequence at positions 5 and 10. G:C and C:G base-pairs at positions 6 and 9 can be tolerated by the 434 repressor but levels of repression are greatly improved if G:C and C:G base-pairs are present at positions 5 and 10 respectively. Single G:C or C:G base pairs in the centre 4 positions of the operator appear to to be tolerated by the 434 repressor.

It was initially hoped that the 434Ala44 repressor would increase tolerance to G:C and C:G base-pairs in the centre of the operator allowing even very unfavourable operators to be used. However, the 434Ala44 repressor has proved to have different effects in vivo compared to the results obtained in vitro, although the relative insensitivity of the Ala44 repressor to G:C base-pairs in the centre of the operator could be discerned. The reasons for these differences are not immediately obvious, however, several factors may play a part in its reduced ability to repress gene expression in vivo.

In vitro the effect being determined is the affinity of the repressor for the operator with no interference from other proteins or DNA structures. In vivo repression is not simply a measure of the repressor's ability to bind to DNA but is due to the degree to which the
repressor impedes the binding or passage of RNA polymerase. Thus it could be envisaged that although the 434^Ala44 repressor binds a 434 operator with as high an affinity as the wild-type repressor, it might offer less resistance to either the binding or passage of RNA polymerase. The mutation of the phenylalanine at position 44 to an alanine may have some important implications for the structure of the repressor outside the DNA-binding domain. In the position the 434 operator occupies in the reporter plasmids interaction of the repressor with RNA polymerase is critical for repression. If the position or structure of the face of the repressor which interacts with, or inhibits the binding of, RNA polymerase is altered then the crucial interactions with polymerase may be affected. Clearly at least some part of the structure of the 434^Ala44 repressor is altered due to the effect this mutation has on the affinity of the repressor for unfavourable operators in vitro. A basis for changes in other regions of the protein not detectable by the in vitro binding studies therefore clearly exists. Whether or not any other changes of structural have been caused could only be detected by structural determination.

On binding to the operator the DNA-binding domain of the repressor undergoes little alteration in structure (Aggarwal et al., 1988; Mondragon et al., 1989b). For the Ala44 repressor to accommodate an operator which is not overwound to such a large extent must require some deformation of the protein. Although this is hinted at by Koudelka et al. (1988) it is deemed an increase in the flexibility of the repressor. However, rotation of either half-site away from the repressor even a small distance could potentially require a large movement of some parts of the protein to retain the alignment of the hydrogen bonds. This movement of the protein would be predicted to disrupt the dimerisation and potentially weaken the interaction with the DNA.
However, this could be avoided by increasing the length of the hydrogen bonds to the DNA, which must be the outcome if the repressor is able to bind to the operator. This increase in the length of the hydrogen bonds would weaken the interaction with the DNA. Movement of the C-terminal of the protein would prevent dimerisation which would abolish DNA-binding.

Binding of the repressor to the operator encompasses many changes in free energy, the overall effect of which is a net gain, favouring the formation of the repressor-operator complex. One of these changes involves the energy required to deform the DNA by overwinding and bending. For the $434^{\text{Ala44}}$ repressor to bind to the operator $434\text{Op1}$ as tightly as the wild-type repressor \textit{in vitro}, as was found by Koudelka \textit{et al.} (1988), other energy changes must be involved to compensate for the reduced energy change required for the reduced deformation of the DNA. It is possible that the $434^{\text{Ala44}}$ repressor forms weaker or fewer hydrogen bonds on binding to the operator. Thus, the affinity for the repressor remains the same although the interaction through hydrogen bonds is weakened because this weaker interaction is offset by the reduction in the energy required to deform the structure of the operator DNA. This is the result obtained by Koudelka \textit{et al.} (1988) for $434\text{Op1}$ although for $434\text{Op2}$ their findings would suggest an overall increase in the free energy change possibly due to reduced alterations to the hydrogen bonding. \textit{In vivo} other components interact with this system altering the way the mutant repressor behaves.

It is possible that the reduction in energy required to deform the DNA is not offset by the reduction in the energy in the hydrogen bonds but by an equal distortion of the repressor. This could retain the strength of the hydrogen bonds. However, it is difficult to involve a
distortion of the protein to retain the alignment of the hydrogen bonds without distortion of the dimerisation domains of the protein which would severely impair binding to the operator.

In conclusion it appears that the interaction of the repressor with the operator is extremely complex and that alteration of any part of it could well result in other effects away from the intended change. Simple repression assays or in vitro binding assays do not provide enough information to determine what effect changes in the protein or operator sequence will have on the overall interaction of the repressor with the operator. These questions can only be answered by detailed structural analysis.
CHAPTER 4.

The 434 Operator/Repressor System as a Means of Regulating Eukaryotic Gene Expression
4.1. INTRODUCTION

The regulation of eukaryotic gene expression by bacterial repressors has been demonstrated in several different eukaryotic cells with a number of different bacterial repressors (Hu and Davidson, 1987; Brown et al., 1987; Figge et al., 1988; Gatz et al., 1991; Smith et al., 1988; Wilde et al., 1992). In yeast two operator/repressor systems have been shown to function. The LexA repressor was shown to repress expression from the GAL1 promoter when operators were inserted between the upstream activator sequences and the TATA box (Brent and Ptashne, 1984). More recently, transcription by RNA polymerase III was shown to be blocked by the tet repressor in yeast (Dingermann et al., 1992). Utilising the inducability of the tet repressor, tetracycline-reversible repression was demonstrated on a tRNA gene containing an amber suppressor anticodon.

The 434 operator/repressor system however, has not been shown to function in eukaryotes. In this study the effectiveness of the 434 repressor at regulating expression from the *Saccharomyces cerevisiae* phospho-glycerate kinase (*PGK*) promoter containing copies of 434Op2 at various positions is investigated.
4.2. RESULTS.

4.2.1. Plasmid Constructions

4.2.1.1. Location of Operators in the PGK promoter.

For repression of a yeast promoter using the 434 operator/repressor system the PGK promoter of *Saccharomyces cerevisiae* was chosen. Firstly, because it is able to direct high level expression of foreign genes in *Saccharomyces cerevisiae* (Derynck et al., 1983) which would enable ease of detection of the reporter gene-product even at repressed levels. Secondly, because a great deal of information is available about the location of regions within the PGK promoter responsible for the high levels of transcription associated with this gene.

The PGK promoter contains two functional TATA boxes at positions -152 and -113, although the one at position -152 is the preferred one (Rathjen and Mellor, 1990). It also contains a CT-rich region close to the transcription start point which is common to other yeast genes, although this has been shown by deletion analysis not to be required for efficient transcription (Rathjen and Mellor, 1990). A short sequence of 7 bases, ACAGATC, appears to direct the initiation site, and is sufficient for the discrete location of the initiation site when located the correct distance from the TATA box (Rathjen and Mellor, 1990). In this region two operators have been positioned, one immediately 5' to the TATA box at -152 and one by replacement of DNA within the CT-rich region close to the RNA initiation point. By replacement of the DNA this close to the initiation point it was hoped that the location of the initiation point would not be affected by the presence of the operator and by locating the operator upstream of the TATA box the
The spacing between the TATA box and the initiation point would not be altered.

The *PGK* promoter has been shown to be dependent on an upstream activator sequence (UAS) for its high activity (Ogden *et al.*, 1986). It has been demonstrated that only sequences 3' of position -538 are necessary for efficient transcription (Stanway *et al.*, 1989). This UAS has been studied in detail and the transcription factors which bind it characterised (Stanway *et al.*, 1987; 1989, Chambers *et al.*, 1988; 1989; 1990). A binding site for the transcription factor RAP1 has been detected and also a site for the transcription factor ABF1. In addition the UAS contains three copies of the pentamer sequence CTTCC which are important for high levels of transcription and have been suggested to be sites of interaction between the promoter and transcription factors but as yet these factors have not been identified.

The arrangement of the UAS is very important for the high levels of expression seen from this promoter therefore the 434 operators were introduced at sites least likely to disrupt the function of the promoter but in positions where the binding of the 434 repressor would be likely to disrupt the protein-DNA interactions between yeast transcription factors and the UAS. One operator was placed immediately 5' of the ABF1 binding site, a second immediately 3' of the 3' CTTCC box and a third introduced as a replacement of DNA between the ABF1 binding site and the RAP1 binding site such that the binding of a repressor would be likely to interfere with the binding of RAP1.

The sequence of the *PGK* promoter and the location of important regions is given in Fig 4.1.
**Figure 4.1** The *PGK* promoter of *Saccharomyces cerevisiae*, showing the location of important regions together with the positions and sequences of introduced restriction sites. The *BgIII* site is referred to in the text as the point of fusion between the *PGK* promoter and the *lacZ* gene.
4.2.1.2. Construction of Reporter plasmids.

This construction is represented diagrammatically in fig 4.2. The PGK promoter and terminator were cut from pCH137 (Hadfield et al., 1990) on a HindIII fragment and cloned into the HindIII site in the polylinker of M13mp19. Site-directed mutagenesis of Kunkel et al. (1987) was used to introduce a SmaI site at position -605, relative to the RNA initiation point, in the PGK promoter. Further site-directed mutagenesis was used to create unique restriction sites at various positions in the PGK promoter (Fig 4.1). The six different PGK promoter-terminator constructs were removed from M13 RF DNA as a 1kb SmaI/HindIII fragment and ligated into the polylinker of a variant of pUC19, pUC19EN, which contains a NarI site in place of the EcoRI site in the polylinker region to give the plasmids in table 4.1. The plasmids pYCW3, pYCW5, and pYCW6 were opened at the unique, introduced restriction sites and 434 operators inserted on annealed oligonucleotides, to give the plasmids pYCW8, pYCW10 and pYCW11 respectively. Plasmid pYCW4 was cut with EcoRV and XhoI and annealed oligonucleotides used to replace the DNA between the two sites, retaining existing sequences bound by yeast transcription factors but also introducing a 434 operator, to give plasmid pYCW9. Plasmid pYCW7 was cut with BglII and StuI and annealed oligonucleotides used to replace DNA between the two sites with original sequence containing a 14bp replacement down-stream of the mRNA initiation point to introduce the 434 operator, to give plasmid pYCW12.

Part of the LacZ gene was cut from pAH193-41A (van Gorcom et al., 1986) on a HindIII, EcoRI fragment and cloned between the HindIII and EcoRI sites in pBR328 (Soberon et al., 1980). This plasmid was cut with HindIII and SalI and the first 11 codons of the lacZ gene ligated in on annealed complementary oligonucleotides containing a BamHI site.
5' of the initiation codon to give the plasmid pCW7. Neither the SalI and nor the HindIII sites were retained in the ligation.

The plasmid pYCWPGK was cut at the unique HindIII site 3' of the PGK terminator and a HindIII fragment containing the URA3 gene from YIp30 inserted to give pYCW19. The last 17 codons of the lacZ gene were synthesised as complementary oligonucleotides with an EcoRI-compatible overhang at the 5' and a BglII-compatible overhang at the 3' end. This was ligated between the EcoRI and BglII sites in pYCW19 and checked by sequencing. This plasmid was opened at its NarI and EcoRI sites and the fragment containing the remainder of the lacZ gene, isolated from pCW7 on a EcoRI, NarI fragment, ligated in. This plasmid, pCW31, was cut at the unique BamHI site immediately 5' of the start codon of the lacZ gene and at the unique NarI site further upstream of the lacZ gene. The series of PGK promoters containing operators, isolated on BglII, NarI fragments from pYCWPGK, 8, 9, 10, 11 and 12 could then be ligated in to form a series of reporter constructs.
<table>
<thead>
<tr>
<th>Plasmid Name</th>
<th>Location of introduced unique restriction site</th>
<th>Restriction site</th>
</tr>
</thead>
<tbody>
<tr>
<td>pYCWPGK</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>pYCW3</td>
<td>5' of ABF1 binding site</td>
<td>XhoI</td>
</tr>
<tr>
<td>pYCW4</td>
<td>As above + 5' of RAP1 binding site</td>
<td>XhoI+EcoRV</td>
</tr>
<tr>
<td>pYCW5</td>
<td>3' of the last CTTCC box</td>
<td>SalI</td>
</tr>
<tr>
<td>pYCW6</td>
<td>5' to TATA box</td>
<td>HpaI</td>
</tr>
<tr>
<td>pYCW7</td>
<td>Downstream of the RNA start point</td>
<td>StuI</td>
</tr>
</tbody>
</table>

*Table 4.1.* The location and type of unique restriction sites introduced into the PGK promoter to allow the insertion of 434 operators.
Figure 4.2. Construction of PGK/β-galactosidase Reporter Plasmid

Site directed mutagenesis in M13 vectors

HindIII
URA3 gene from YIp30

BgIII/EcoRI
Annealed oligos for 3' LacZ
4.2.1.3. Construction of a Plasmid expressing the 434 Repressor

This construction is represented diagrammatically in fig 4.3. The 2µM origin and most of the TRP1 gene were removed from pYcDE-2 (Hadfield et al., 1986) by digestion with XbaI and HindIII followed by religating the remainder of the plasmid with annealed oligonucleotides used to create a small linker joining the XbaI and HindIII ends. The XbaI site was destroyed in the ligation and an XhoI site introduced. The resulting plasmid, pcDE1, was opened at the unique EcoRI cloning site between the ADH1 promoter and the CYC1 terminator, and a small linker inserted to introduce unique XbaI and BamHI sites, the EcoRI site was not retained by the ligation. This plasmid is pACXBl. The 434 repressor gene was cut from pPLRT1 on an XbaI/Sau3AI fragment and ligated into pACXB1 opened at its XbaI and BamHI sites to give pYCW34. Both pYCW34 and pYcDE-2 were cut with SalI and MluI and the 434 expression cassette ligated into the 2µm plasmid to give pYCW36.
4 - The 434 Operator/Repressor System in Yeast

Figure 4.3. Construction of pYCW36

XbaI/HindIII ligate with linker

EcoRI ligate with linker

MluI/SalI

XbaI/BamHI
434 from pPLRT1 on XbaI/Sau3AI

MluI/SalI
4.2.2. Western Analysis of 434 Expression in Yeast

The level of expression of the 434 repressor in strains carrying the plasmid pYCW36 was determined by western analysis. The results (Fig. 4.4) show that the 434 repressor is expressed at 0.25 to 0.5% of cell protein. This is less than the 1-2% of cellular protein produced by a similar plasmid expressing the *E. coli* chloramphenicol acetyl transferase (CAT) gene expressed from a similar plasmid (Hadfield et al., 1987).

![Figure 4.4](image.png)

**Figure 4.4** The expression of the 434 repressor in strains carrying the plasmid pYCW36. **Lane a**, 20μg total soluble protein from S150-2B carrying pYCW36. **Lane b**, 20μg protein from S150-2B, **Lane c**, 20ng purified 434 repressor. **Lane d**, 10ng purified 434 repressor. **Lane e**, 5ng purified 434 repressor. A weak band in lane b is the result of spill over from the loading of lane a.
4.2.3. Experimental Approach

In order to measure the repression achieved by the 434 repressor in yeast, strains carrying the reporter plasmids were transformed with pYCW36 expressing the 434 repressor and pYcDE-2 as a negative control. The difference in the levels of β-galactosidase expression from the reporter gene between the two strains would then give the level of repression. A control was also included which measured the repression of the reporter construct when no artificially introduced operators were present.

4.2.4. Measurement of Repression

4.2.4.1. Effect of operators on expression from the PGK promoter

Examination of the expression of the reporter gene in the constructs containing operators in the PGK promoter revealed little or no effect on the levels of β-galactosidase expression (Fig 4.5). This demonstrates that the choice of sites for the insertion of the operators was, as predicted, in regions where disruption of the promoter was not likely to occur.

4.2.4.2. Repression of Reporter Constructs

Repression of the 434 operator-containing constructs using the 434 repressor revealed variations in the level of repression depending on the position of the operator. The operators within or downstream of the UAS give similar levels of repression, at 41% for pYCW38 and 31% for pYCW39. Repression at an operator 5' of the UAS gives no detectable repression. An operator positioned close to the TATAA box gave the greatest repression in this study, reducing expression by 65%.
Data for an operator positioned downstream of the transcription start point showed very low repression at approximately 23% (Fig 4.6).
Figure 4.5. Expression of β-galactosidase from the PGK promoter containing operators at various locations. The results demonstrate that the operators have little effect on the level of expression. The results vary by ±1.1 fold with respect to the control, however, statistical testing reveals no significant difference between each set.
Figure 4.6. Repression of PGK promoters containing 434 operators by the 434 repressor. pYCWPBGK is the construct containing the wild-type PGK promoter. pYCW37 contains an operator 5' of the UAS. pYCW38 contains an operator 5' of the RAP1 binding site within the UAS. pYCW39 contains an operator 3' of the CTTCC boxes within the UAS. pYCW40 contains an operator 5' of the TATA box. pYCW41 contains an operator 3' of the transcription start point.
<table>
<thead>
<tr>
<th>Plasmid Name</th>
<th>Position of operator</th>
<th>% Repression</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>pYCWWT</td>
<td>no operator</td>
<td>0</td>
<td>3.0</td>
</tr>
<tr>
<td>pYCW37</td>
<td>5' of UAS</td>
<td>0</td>
<td>3.3</td>
</tr>
<tr>
<td>pYCW38</td>
<td>Within UAS</td>
<td>41</td>
<td>2.8</td>
</tr>
<tr>
<td>pYCW39</td>
<td>3' of UAS</td>
<td>31</td>
<td>2.2</td>
</tr>
<tr>
<td>pYCW40</td>
<td>5' of TATAA box</td>
<td>65</td>
<td>1.8</td>
</tr>
<tr>
<td>pYCW41</td>
<td>3' of transcription start point</td>
<td>22</td>
<td>9.8</td>
</tr>
</tbody>
</table>

**Table 4.2.** Data for the repression of the PGK promoter/β-galactosidase reporter fusion. S.D. refers to the standard deviation of the percent repression calculated using formulae for the propagation of error, see Materials and Methods. Using statistical tests repression results for pYCWWT and pYCW34 conform to the null hypothesis, that is that there is no significant difference between the control and repressed data. All other repression results show that the control and repressed data are from different populations to better than 95% certainty.
4.3. DISCUSSION

The data presented in this chapter clearly demonstrate that the 434 operator/repressor system is able to function in *Saccharomyces cerevisiae*.

The largest degree of repression is obtained when the 434 repressor binds close to the TATAA box element at position -152 in the PGK promoter. This repression is most likely due to interference of the 434 repressor with the transcription factors which bind to the TATAA box such as TFIID which is required for normal transcription initiation (Eisenmann *et al.*, 1989). Weaker repression is seen when the repressor binds in other areas of the promoter. This is particularly evident in the UAS, where binding to an operator at a site close to the binding site for the transcription factor RAPl represses expression by 41%. The binding site for the RAPl protein is essential for the high level expression from the PGK promoter (Stanway *et al.*, 1989) and preventing the efficient binding of this protein is likely to be the cause of repression. Similar repression of 31% is also achieved when the repressor binds to an operator situated downstream of the CTTCC boxes at the 3' end of the UAS. The cause of the repression in this area is difficult to pinpoint. The CTTCC boxes have been shown to be areas of protein-DNA interaction in the PGK promoter, hence, the binding of a repressor in this area would interfere with the binding of a yeast transcription factor. It is also possible that the 434 repressor, by bending the DNA at the operator alters the local conformation sufficiently to reduce the interaction between the proteins bound at the UAS and the transcription complex bound at the TATAA box and/or transcription start point. Characterisation of sub-elements of the UAS have shown that expression can be achieved if the RAPl binding site is present.
together with either the ABF1 binding site or at least one of the CTTCC boxes (Stanway et al., 1989). Single elements however, do not confer strong activation on a minimal promoter.

The weakest repression, of 22%, is seen when the repressor binds to an operator located downstream of the transcription start point. Here repression is presumably due to the interference with the passage of RNA polymerase. The low repression level indicates that the 434 repressor binds to the operator DNA too weakly to strongly interfere with the movement of the transcription complex.

Other published repression studies using bacterial repressors in yeast are limited to the tet repressor (Dingermann et al., 1992) and the lexA repressor (Brent and Ptashne, 1984).

Repression by the tet repressor in yeast has been quantified at greater than 50-fold repression. This was backed up by the effect the repressor has on the growth rate of cells expressing it using a complex reporter system employing a tRNA amber suppressor gene containing tet operators in the promoter, linked to the suppression of an amber mutation in the MET8 gene. Thus in the presence of the tet repressor repression of the tRNA amber suppressor gene prevents expression of the MET8 gene containing the amber stop codon within its reading frame resulting in cells auxotrophic for methionine. In the absence of the tet repressor or, when induction is performed in the presence of tetracycline, expression of the tRNA gene allows expression of the MET8 gene and hence restores prototrophy for methionine. The dramatic effect on the growth of the cells indicates that efficient repression is being achieved.

The repression due to the lexA repressor binding to operators in various positions in the GAL1 promoter was quantified and the results demonstrated up to 7.8-fold repression with the insertion of one
operator. This was improved to 10-fold with two operators inserted in tandem. The level of repression was strongly dependent on the location of the operator within the promoter. The best repression was obtained with the operator located 59bp downstream of the UAS. Weaker repression was observed closer (4bp) to the UAS and further (177bp) from the UAS. The repression data from Brent and Ptashne (1984) are complicated by the fact that much of the DNA between the UAS and TATAA box was deleted in the operator-containing reporter plasmids. Although this affects the level of expression of the reporter gene in only one instance it may alter the susceptibility of the promoter to repression. A direct comparison of the repression data from the *lexA* studies to those using the 434 repressor may not be valid. However, it is evident that the *lexA* repressor brings about a higher level of repression than the 434 repressor.

The reason why the 434 repressor does not repress gene expression in yeast as well as the *lexA* repressor nor the *tet* repressor is not immediately obvious. All three repressors have similar dissociation constants, approximately $1 \times 10^{-9} \text{M}$ for the *lexA* repressor (Brent and Ptashne, 1984), approximately $1 \times 10^{-9} \text{M}$ for the 434 repressor to the operator, ACAAGATATCTTG, used in this assay (Koudelka et al., 1988) and approximately $1 \times 10^{-9} \text{M}$ for the *tet* repressor (Takahashi et al., 1986). In the *lexA* study the repressor was expressed at 0.005-0.025% of cell protein whereas the 434 repressor was expressed at 0.25-0.5% of cell protein, at least 10-fold higher, therefore the level of the repressor is not likely to be influencing the results. The positional effects noted with the *lexA* repressor (Brent and Ptashne 1984) and with the 434 repressor in this study could be the cause of the weaker levels of repression achieved by the 434 repressor. This also seems unlikely, as at
least one of the operator positions used in this study, pYCW36, is comparable to those used by Ptashne and Brent (1984).

In mammalian cells Smith et al. (1988) used the *lexA* repressor to repress expression from the HSV tk promoter containing *lexA* operators. Here the best repression achieved by the binding of a repressor to a single operator was 61%. This operator was located 5bp upstream of the TATAA box in a very similar position to the operator in the construct pYCW37. Repression by the *lexA* repressor in mammalian cells varied from 16 to 61% with single operators, results which are very similar to the data obtained with the 434 repressor in this chapter, and 26-90% with two tandemly inserted operators.

In the genome of the coliphage 434 operators function in tandem with pairs of repressor dimers binding cooperatively to adjacent operator sequences. It may therefore be informative to attempt further work with the 434 repressor using tandem operators in the *PGK* promoter to attempt to improve repression. However, the main objective of assessing the ability of the 434 repressor to function in yeast has been achieved.
CHAPTER 5.

The Effect of the $434^{P22}$ Repressor on the Growth of *Saccharomyces cerevisiae*. 
5.1. INTRODUCTION

The initial aim of the work to be undertaken using the 434 repressor was to use both the wild-type and P22 binding variant to investigate repression in eukaryotes. Work presented in this chapter describes difficulties encountered in using the 434\textsuperscript{P22} repressor in yeast.

Bacterial repressors recognise fairly short sequences of DNA: as a consequence in prokaryotes, due to their small genome size, the chance occurrence of an operator within the control region of a gene is rare. This is also due to selective pressure against the random occurrence of operators in the host. In the larger genomes of eukaryotes operator sequences can occur in higher numbers purely by chance. In a rare number of instances these may coincide with important regions in the genome where the binding of a bacterial repressor could cause a severe problem to the host's normal gene expression. In repression studies where bacterial repressors are used to regulate gene expression in eukaryotes there is a risk that transformed cells will have their growth impaired by the chance repression of genes within their genome as well as repression of the reporter gene.

Repression studies to date have not reported such an effect on the host cells. In this chapter the inhibitory effect of the 434\textsuperscript{P22} repressor on the growth of yeast is reported.
5.2. RESULTS

5.2.1. Plasmid Constructions

To express the $434^{P22}$ repressor at high level, an expression plasmid was designed based on the yeast $2\mu m$ origin of replication. This construction is similar to the construction represented in figure 4.3 except that the $434^{P22}$ gene is substituted for the wild-type $434$ gene. The plasmid pYcDE-2 (Hadfield et al., 1986) was cut with XbaI and HindIII and the fragment containing the CYC1 terminator and ADH1 promoter ligated with a linker. The XbaI site was not retained in the ligation and an XhoI site introduced, to give pCDE1. This plasmid was cut with EcoRI and a small linker introduced which included an XbaI site and a BamHI site (the EcoRI site was not retained in the ligation) to give pACXB1. The $434^{P22}$ repressor gene was cut from pPLRT2 on an XbaI/Sau3AI fragment and ligated into pACXB1 opened at its XbaI and BamHI sites to give pYC17. Both pYC17 and pYcDE-2 were cut with SalI and MluI and the $434^{P22}$ expression cassette ligated into the $2\mu m$ plasmid to give pYC28.

A second expression vector was also constructed for the $434^{P22}$ repressor, by the route represented in fig 5.1. This was designed to integrate into the yeast genome maintaining single copy number. The promoter-repressor-terminator cassette in pYC17 was removed on a SalI, XhoI fragment and cloned into the unique SalI site in pYRG12 (Hadfield et al., 1987). Two clones were picked each having the repressor cassette in the opposite orientation: these were pYC20A and B. These plasmids express the $434^{P22}$ repressor from a truncated version of the ADH1 promoter, giving constitutive expression and carry the HIS3 gene. Linearisation of the plasmid, with
5 - The effect of 434^P22 on yeast

Figure 5.1. Construction of pYCW20
XhoI which cuts uniquely within the HIS3 gene, improves the frequency of homologous recombination into the HIS3 locus in the genome by up to 1000-fold (Orr-Weaver et al., 1981).

A third plasmid designed to express the 434P22 repressor was also designed, based on the centromere plasmid pBM150 (Johnston and Davis, 1984): construction of this plasmid is represented in fig 5.2. The 434P22 repressor gene was cut from pPLRT2 on an XbaI/Sau3AI fragment and cloned into the polylinker of pIC19H (Marsh et al., 1984) opened at its BamHI and XbaI sites to give pCW37B. This plasmid was then cut with BglII and SalI and the 434P22 repressor gene cloned into pBM150 opened at its SalI and BamHI sites. This places the 434P22 repressor gene under the control of the yeast GAL1 promoter. Therefore, during growth on glucose-based media the repressor gene would be silent, whereas, during growth on galactose in the absence of glucose the repressor gene would be expressed.
Figure 5.2. Construction of pYCW42

- 5 - The effect of 434P22 on yeast
5 - The effect of $^{434}{^P_{22}}$ on yeast

5.2.2. Experimental Approach

Each of the plasmids of interest was transformed into S150 2B by lithium acetate transformation and selection performed on semi-defined minimal media containing glucose, amino acids and vitamins. Other supplements to the media were added as required.

5.2.3. The Effect of $^{434}{^P_{22}}$ on Transformation Efficiency

Under the conditions described above no transformants were detected using the plasmids pYCW28 and pYCW20A or B. Examination of the plasmids revealed no defect in either of the auxotrophic selection genes, as determined by restriction mapping. Several other clones of pYCW28 and pYCW20 from the construction program were picked, but these too failed to transform. Control plasmids identical to the expression plasmids except lacking the $^{434}{^P_{22}}$ expression cassette transformed S150 2B, under identical conditions, at high frequency. This information implicates the $^{434}{^P_{22}}$ repressor as a lethal gene product in yeast.

To investigate this further the plasmid pYCW42 was constructed which allowed the inducible expression of the $^{434}{^P_{22}}$ repressor in yeast, construction of this plasmid is shown in fig. 5.2. *Saccharomyces cerevisiae* S150-2B was transformed with the vectors pYCW42 and pBM150 by lithium acetate transformation and the transformation mixes divided into two: one half was plated onto semi-defined minimal media containing glucose, leucine, tryptophan and histidine and the other half onto identical media containing galactose in place of the glucose. The plates were incubated at 30°C until colonies appeared. Both plasmids produced high efficiency transformation, as judged by
the number of colonies appearing on the glucose plates. The plasmid pBM150 also produced a large number of transformants on the galactose plates. However, the plasmid pYCW42 failed to produce large numbers of colonies on the galactose plates (table 5.1).

To further test this, colonies from the glucose plates were streaked onto galactose plates. Yeast containing the plasmid pBM150 were able to grow, whereas colonies containing the plasmid pYCW42 were unable to grow on the galactose media. Thus supporting the data that the 434P22 repressor is able to strongly inhibit the growth of *Saccharomyces cerevisiae.*

This supports the notion that the 434P22 repressor is a lethal gene product in yeast, as in cells grown on glucose the 434P22 repressor gene under the control of the GAL1 promoter would be silent whereas on galactose media the gene would be expressed.

Examination of yeast sequences in the EMBL data base for P22 operator-like sequences, reveals potential binding sites in a number of important genes (fig. 5.3). Three of these genes ADE2, ARG4 and LYS2 are involved in purine and amino acid synthesis pathways. If these genes were repressed and responsible for the apparent lethal nature of the 434P22 repressor, supplementation of the selection media with the three amino acids alanine, arginine and lysine would allow transformants to grow. A further potential binding site also exists in the GAPDH gene. Growth on a non-fermentable carbon source such as ethanol would overcome repression of this gene. However, these variations failed to relieve the effect of the 434P22 repressor. Many other genes contain potential repressor-binding sites and could be singly or jointly be responsible for this effect on transformation.
5 - The effect of 434² on yeast

<table>
<thead>
<tr>
<th>plasmid</th>
<th>Glucose</th>
<th>Galactose</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBM150</td>
<td>1120</td>
<td>1080</td>
</tr>
<tr>
<td>pYCW42</td>
<td>1030</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 5.1. The number of colonies appearing on glucose and galactose plates from transformations of *Saccharomyces cerevisiae* S150 2B with the plasmids pBM150 and pYCW42. The results are given as the number of colonies appearing on each plate to the nearest 10, except for pYCW42 where the actual number of colonies is given.
Figure 5.3. Potential binding sites within known yeast sequences for the P22 repressor. Sites were identified by using the UWGCG program "Find" (Devereux et al., 1984). Two patterns were included in the search which took into account the consensus sequence of the P22 repressor (Poteete et al., 1980) and the sequence requirements of the non-contacted bases of the 434 repressor, described in chapter 3.

Scade2  ck:  4310  len:  2,518  ! Saccharomyces cerevisiae
phosphoribosylamino-imidazole-carboxylase (ADE2) gene
"2"  
..... ANTAADSSWHHTNANT .....  
322: ATAAA ATTAAGAGATATAAAAT ATTAG
Scam4  ck:  574  len:  2,500  ! Saccharomyces cerevisiae AmdY gene
for putative yeast amidase
"1"  
..... ANTAADNWNHTTNANT .....  
1,294: CCAAG AATGAAAATCTTTGATT GTCAG
Scapsy  ck:  4718  len:  2,396  ! Yeast APS gene for cytoplasmic
aspartyl-tRNA synthetase (AspRS)
"1"  
..... ANTAADNWNHTTNANT .....  
1,494: GAAGA AATGAAAATATCTTTAAT GTCAT
Scarg4  ck:  4718  len:  2,396  ! Yeast (S.cerevisiae)
argininosuccinate lyase (ARG4) gene, complete coding sequence
"1"  
..... ANTAADNWNHTTNANT .....  
199: CAGGC AATGAAAATCTTTTACT CTTCC
Scars2  ck:  9258  len:  1,517  ! Yeast autonomously replicating
sequence encoding a tRNA (GIN)
"1"  
..... ANTAADNWNHTTNANT .....  
405: CCCAA AATCAATCAATTTTTTAAT GAAAT
Scatp21  ck:  8734  len:  1,932  ! Yeast (S.cerevisiae) nuclear ATP2
gene encoding mitochondrial F-1-ATPase beta-subunit, complete CDS
"1"  
..... ANTAADNWNHTTNANT .....  
1,817: TATTC ACTCAAAAATATTTTTAAT TCTTA
Scatr  ck:  7518  len:  3,292  ! S.cerevisiae ATR1 gene conferring
aminotriazole resistance, complete cds
"1"  
..... ANTAADNWNHTTNANT .....  
2,853: CAAAAC AATGAAAATACCTTTGATT CAATG
Scatr  ck:  7518  len:  3,292  ! S.cerevisiae ATR1 gene conferring
aminotriazole resistance, complete cds
"1"  
..... ANTAADNWNHTTNANT .....  
2,864: ATTAA AATGAAAATACCTTTTATT CTATG
Scapsy  ck:  4718  len:  2,396  ! Yeast APS gene for cytoplasmic
aspartyl-tRNA synthetase
"1"  
..... ANTAADNWNHTTNANT .....  
323: GAAGA AATGAAAATCTTTTCAAT AACTG
Sccapk  ck:  2534  len:  2,613  ! Saccharomyces cerevisiae BCY1 gene
encoding the regulatory subunit of cAMP-dependent protein kinase
"2" /rev  
..... ANTAADNWNHTTNANT .....  
2,420: ATCTG AATGAAAATCTTTTCAAT GGGCA
Sccdc7  ck:  6076  len:  2,100  ! Yeast (S.cerevisiae) cell cycle
gene (CDC7), complete CDS
"1"  
..... ANTAADNWNHTTNANT .....  
2,066: AAAAG ATGAAAATCTTTTCAAT AACTG
Sccup1  ck:  1792  len:  2,010  ! Yeast (S.cerevisiae) CUP1 locus
encoding copper chelatin and an UR5
"2" /rev  
..... ANTAADNWNHTTNANT .....  
426: TTGTTT ACTGAAAAATCTTTTCAAT TTATC
Sccusnd  ck:  5406  len:  1,037  ! Saccharomyces cerevisiae, Cu, Zn
superoxide dismutase gene
"1"  
..... ANTAADNWNHTTNANT .....  
390: AAACA AATGAAAATCTTTTCAAT AACTG
Sccd1  ck:  9677  len:  2,306  ! Yeast DIT1 gene involved in spore
wall maturation
"1"  
..... ANTAADNWNHTTNANT .....  
930: TTGAG AATGAAAATCTTTTCAAT AACTG
Scenoa  ck:  706  len:  2,013  ! Yeast (S. cerevisiae) enolase gene
(clone peno46) and flanks
"1"  
..... ANTAADNWNHTTNANT .....  
236: ATATT AATGAAAATCTTTTCAAT TTATC
The effect of 434P22 on yeast

5 - Gene

Scenod ck: 1248 len: 811 ! Yeast (S. cerevisiae) enolase

**1**

692: ATATT AATCCTAATTTTACT TCTTT

Scfas1 ck: 284 len: 7,545 ! Yeast FAS1 gene for fatty acid synthetase subunit beta

**2** /rev

4,405: AAGCT ATGGAAAGATTTTAAAT GTCTA

Scfas1x ck: 372 len: 7,545 ! S. cerevisiae fatty acid synthetase gene (FAS1), complete cds

**2** /rev

4,405: AAGCT ATGGAAAGATTTTAAAT GTCTA

Scfasb ck: 2993 len: 7,485 ! Yeast (S. cerevisiae) FAS1 gene encoding fatty acid synthase beta subunit, complete cds

**2** /rev

4,007: AAGCT ATGGAAAGATTTTAAAT GTCTA

Scfus3 ck: 283 len: 1,999 ! S. cerevisiae FUS3 protein kinase gene, complete cds

**1**

19: CTGGA ACTCAAAATTCTTTTACT CGAAA

Scgap1 ck: 8265 len: 1,261 ! Yeast gene (pgap49) for glyceraldehyde-3-phosphate dehydrogenase

**1**

1,149: GCTTA AGTGAATTTACTTTAAAT CTTGC

Scgly ck: 3986 len: 1,542 ! Yeast GCY gene (homologous to vertebrate eye lens protein)

**2**

725: TCAAA AATGAAGACATCTTGAGT GTGCC

Schn401 ck: 5212 len: 1,800 ! Yeast genes for histone H3 and H4 (copy-I genes from PMS191)

**1**

1,714: AAGCT ATTTAATTTTAAAT TGGGT

Scptck ck: 3472 len: 2,409 ! S. cerevisiae meiosis-inducing protein (IME1) gene, complete cds

**1**

974: TATGT ATTTAAGGTAAATTAATT ATCTT

Scllck: 4063 len: 4,168 ! Yeast (S. cerevisiae) tRNA ligase gene, complete cds

**1**

876: TTTGTT ATTTAAAAATTTAAAT ACTGT

Scll2ck: 2206 len: 4,980 ! S. cerevisiae alpha-aminoadipate reductase (LYS2) gene, complete cds

**2**

3,639: TTGAC ACTGAATACTACTTTAAT TTGTC

Scmit78 ck: 7067 len: 2,630 ! Yeast gene for 70kd mitochondrial outer membrane protein

**2** /rev

3,641: TTGTG ATTTAAAAATAATTAATT ACTGT

Scplck: 4006 len: 4,986 ! S. cerevisiae nucleoporin (NUPl) gene, complete cds

**1**

3: GA ATTTAATTTTAAAT TGGGT

Scly2ck: 2206 len: 4,980 ! S. cerevisiae alpha-aminoadipate reductase (LYS2) gene, complete cds

**2**

3,639: TTGAC ACTGAATACTACTTTAAT TTGTC

Scmit78 ck: 7067 len: 2,630 ! Yeast gene for 70kd mitochondrial outer membrane protein

**2** /rev

3,641: TTGTG ATTTAAAAATAATTAATT ACTGT

Scplck: 4006 len: 4,986 ! S. cerevisiae nucleoporin (NUPl) gene, complete cds
### The effect of 434P22 on yeast

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Type and Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCP8MA</td>
<td>S. cerevisiae SIN3 gene encoding a paired amphipathic helix protein</td>
</tr>
<tr>
<td>SCPDR4</td>
<td>Yeast PDR4 gene for pdr4 protein</td>
</tr>
<tr>
<td>SCPH081</td>
<td>Yeast regulatory gene PH081</td>
</tr>
<tr>
<td>SCPH035</td>
<td>Yeast genes PH05 and PH03 for repressible (PH05) and constitutive (PH05) acid phosphatase</td>
</tr>
<tr>
<td>SCPMSLA</td>
<td>S. cerevisiae PMS1 gene encoding DNA mismatch repair protein, complete cds</td>
</tr>
<tr>
<td>SCPOL2R</td>
<td>Yeast (S. cerevisiae) RBP2 gene encoding RNA polymerase II, 140 kd subunit</td>
</tr>
<tr>
<td>SCPRBL</td>
<td>Yeast (S. cerevisiae) protease B (PRP1) gene, complete cds</td>
</tr>
<tr>
<td>SCPRR9</td>
<td>Yeast prp9 gene for PRP9 protein involved in pre-mRNA splicing</td>
</tr>
<tr>
<td>SCRAD50</td>
<td>Yeast RAD50 gene for 153 kD protein</td>
</tr>
<tr>
<td>SCRADH</td>
<td>Yeast RADH gene for putative helicase</td>
</tr>
<tr>
<td>SCREV1</td>
<td>S. cerevisiae REV1 gene required for normal induction of mutations by physical and chemical agents, complete cds</td>
</tr>
<tr>
<td>SCRGR1</td>
<td>S. cerevisiae RGR1 protein gene, complete cds</td>
</tr>
<tr>
<td>SCRPL3</td>
<td>Yeast (S. cerevisiae) ribosomal protein L44 gene, complete cds</td>
</tr>
<tr>
<td>SCRPG4</td>
<td>Yeast (S. cerevisiae) ribosomal protein L45 gene, complete cds</td>
</tr>
<tr>
<td>SCRPG5</td>
<td>Yeast (S. cerevisiae) ribosomal protein L45 gene, complete cds</td>
</tr>
<tr>
<td>SCRP28A1</td>
<td>Yeast first copy gene for ribosomal protein rp 28</td>
</tr>
<tr>
<td>SCRP28A2</td>
<td>Yeast first copy genes for ribosomal proteins rp28 and S16A (sce-1)</td>
</tr>
</tbody>
</table>
The effect of 434P22 on yeast

2^2

235: AGATA ATTTAAACAAATTGAGT TGAAG
Scs10a ck: 4762 len: 3,458 Yeast second copy genes for ribosomal proteins 28 and 16A (Scs-2)

1

2,989: GTTTA ACTCAAACAAAATTGATTTAATAA
Scrds01 ck: 5164 len: 1,700 Yeast (S.cerevisiae) Gln-tRNA gene Ser-tRNA gene (5' end), and delta and sigme insertion elements

1

144: CCCTAC AATCAATCTTTTAAT GAAAT
Scrds4a ck: 5164 len: 1,700 Yeast (S.cerevisiae) Gln-tRNA gene Ser-tRNA gene (5' end), and delta and sigme insertion elements

1

144: CCGA AACCAATCTTTTAAT GAAAT
Scrds4a ck: 5164 len: 1,700 Yeast (S.cerevisiae) Gln-tRNA gene Ser-tRNA gene (5' end), and delta and sigme insertion elements

dehydratase (EC 4.2.1.13)

1

2,569: AGATA ATTTAAACAAATTGAGT TGAAG
Scrsc10a ck: 4762 len: 3,458 Yeast second copy genes for ribosomal proteins 28 and 16A (Scs-2)
5.3. DISCUSSION

The results presented in this chapter provided strong evidence that the $434^{P22}$ repressor protein has a binding site (or binding sites) in the yeast genome which coincide with a yeast gene or other important DNA sequence.

Analysis of the known yeast DNA sequences reveals a range of potential binding sites. Binding of a repressor at a number of these sites could potentially repress expression of the associated gene, resulting in an altered phenotype. However, in the cases of the $ADE2$, $ARG4$ and $LYS2$ genes supplementation of the media with adenine, arginine and lysine fails to reverse the phenotypic change. This suggests that it is not the binding of the repressor to the operator sites in these genes that is causing the observed effects. However, this does not rule out the possibility that the $434^{P22}$ repressor is binding at these sites and repressing gene expression. The repressor however, must be binding at another site or sites thus causing the effect on growth and the apparent effect on transformation.

Several of the genes which contain identified P22 operators are important to the correct function of yeast cells. However, if the $434^{P22}$ repressor binds at some or even all of these potential binding sites it would be a very complex task indeed to be able to express the $434^{P22}$ repressor in yeast without the observed effects on the growth of the cells. It should also be realised that the reported sequences of the genes in *Saccharomyces cerevisiae* represent only a fraction of the total genomic DNA and many binding sites may be present in, as yet, uncharacterised genes which could influence growth.

Repression studies using either the $434^{P22}$ homodimer or the $434/434^{P22}$ heterodimer have obviously been made impractical by these
findings. Due to the wide range of different genes potentially affected by the repressor, including genes whose products are involved in amino acid biosynthesis, both transcription and translation could be influenced. Therefore, any repression data from transient assay or from induced expression of the 434P22 repressor could be distorted by effects at operator sites outside the reporter genes promoter. These data do however, indirectly indicate that the 443P22 repressor is able to bring about efficient repression in *Saccharomyces cerevisiae.*
CHAPTER 6.

Selection of Altered Binding Specificity Repressors.
6.1. INTRODUCTION

All studies in which bacterial repressors are controlling gene expression in a heterologous system have involved the \textit{in vitro} insertion of operator sequences at appropriate positions in the target gene promoter. However, the range of applications for such a control system would be further extended if the \textit{in vitro} manipulation could be avoided. This could be potentially achieved by the use of repressors with an altered DNA-binding specificity. These altered specificity repressors (ASRs) recognise so called 'pseudo-operator' sequences.

At present there are no clear rules which predict how any particular amino acid substitution within the 434 \(\alpha_3\) helix will alter the DNA-sequence recognition, although a recent study using the \textit{lac} repressor has gone some way towards elucidating amino acid/base-pair interactions (Kleina & Miller, 1990). The strategy described in this chapter requires a mini-library of repressor genes to be created, in which the codons for certain amino acids important in protein-DNA recognition are randomised. A selection system is then used to identify those repressors able to bind the target pseudo-operator sequence.

The selection system used is based on the \textit{lac} operon and the suicide substrate, \(p\)-nitrophenyl 1-thio-\(\beta\)-\(D\)-galactopyranoside (TPNPG). The rationale behind the system is based on the ability of the 434 repressor to control \textit{lac} gene expression by binding to an operator cloned within the promoter of the \textit{lac} operon. In the absence of binding expression of \textit{lac} operon genes causes TPNPG to be actively transported into the cell by the \textit{lac} permease in an energy-utilising process. Since TPNPG is not a utilisable substrate, this transport is futile and, on a poor carbon source (eg succinate), the resultant energy drain prevents cell growth. Conversely, the binding of a repressor to its operator will
reduce the expression of lacY, thereby avoiding the futile transport of TPNPG, and allowing cell proliferation. In this way only cells in which the repressor and operator interact survive and grow in the presence of TPNPG.

To test the ability of TPNPG to select Lac− cells from a background of Lac+ cells, mixing experiments were performed with 6300Lac+ cells and its ΔlacU169 derivative. These strains were grown to mid-log phase in the presence of 1mM IPTG, to induce the lac operon. The cultures were mixed in various proportions before plating 100-150 cells on minimal succinate medium containing 1mM IPTG and 50µg/ml X-gal, both with and without TPNPG and incubation at 28°C. Initial experiments showed that 500µg/ml TPNPG was only able to retard growth of the Lac+ cells: however, the background growth was eliminated in the presence of 1mg/ml TPNPG. At this higher concentration the growth of LacY− cells was unaffected. It is worth noting that the recA1 derivative of the Lac− strain is unable to survive in the presence of TPNPG, the reason for which is unclear.

Using this system Wilde et al. (personal communication) have now isolated novel ASRs and demonstrated that a selection system that is capable of isolating functional ASRs from a background of random clones is able to function efficiently in E. coli.
6.2. RESULTS

6.2.1. Screening for 434 Pseudo-operator Sites

Using the University of Wisconsin Genetics Computer Group (UWGCG) program "Find" (Devereux et al., 1984) run on a VAX computer sequences could be screened for the presence of pseudo-operators. The patterns used were based on the data presented in chapters 1 and 2 for the sequence tolerances at the centre 6 bases of the 434 operator and used combinations of ambiguity symbols for the contacted bases of the operator (Table 6.1).

The PGK promoter of Saccharomyces cerevisiae contained a single pseudo-operator site found by this method. This was located downstream of the RNA initiation point with the 434 pseudo-operator located between the initiation point and the start codon.

The PGK promoter sequence was of particular interest because constructs were available where the PGK promoter was linked to the E. coli β-galactosidase gene, enabling easy assay for repression in clones expressing putative altered binding specificity repressors.

6.2.2. Plasmid Constructions

6.2.2.1. Plasmids Containing Operators

All operator-containing plasmids are based on pADΔTc, where the operator is synthesised as a self-complementary oligonucleotide with SalI compatible ends and inserted into the SalI site in pADΔTc. The operator used is a variation of the 434 operator (ACAAGATATCTTGT) containing changes at positions 2 and 3 with complementary changes at 12 and 13 (AGTAGATATCTACT). This is an idealised form of the pseudo-operator which is a palindrome and should simplify the selection procedure.

- 101 -
### Table 6.1

The patterns used in the Find program to locate pseudo-operators in sequences present in UWCGG format. The ambiguity symbols used are as follows, N, any base; D, A, G or T; H, A, C or T; W, A or T; S, G or C.

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>Name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT1</td>
<td>ACAADNWWNHTTGT</td>
<td>2G3C1</td>
<td>AGCADNWWNHTGCT</td>
</tr>
<tr>
<td>WT2</td>
<td>ACAADWSWWHTTGT</td>
<td>2G3C2</td>
<td>AGCADWSWWHTGCT</td>
</tr>
<tr>
<td>1T1</td>
<td>TCAADNWWNHTTGA</td>
<td>2G3G1</td>
<td>AAGADNWWNHTCCT</td>
</tr>
<tr>
<td>1T2</td>
<td>TCAADWSWWHTTGA</td>
<td>2G3G2</td>
<td>AAGADWSWWHTCCT</td>
</tr>
<tr>
<td>1C1</td>
<td>CCAADNWWNHTTGG</td>
<td>2T3T1</td>
<td>ATTAANWWNHTAAT</td>
</tr>
<tr>
<td>1C2</td>
<td>CCAADWSWWHTTGG</td>
<td>2T3T2</td>
<td>ATTAADWSWWHTAAT</td>
</tr>
<tr>
<td>1G1</td>
<td>GCAADNWWNHTTGC</td>
<td>2T3C1</td>
<td>ATCAADNWWNHTGAT</td>
</tr>
<tr>
<td>1G2</td>
<td>GCAADWSWWHTTGT</td>
<td>2T3C2</td>
<td>ATCAADWSWWHTGAT</td>
</tr>
<tr>
<td>2G1</td>
<td>AGAADNWWNHTTCT</td>
<td>2T3G1</td>
<td>ATGADNWWNHTCAT</td>
</tr>
<tr>
<td>2G2</td>
<td>AGAADWSWWHTTCT</td>
<td>2T3G2</td>
<td>ATGADWSWWHTCAT</td>
</tr>
<tr>
<td>2T1</td>
<td>AATADNWWNHTTAT</td>
<td>2A3T1</td>
<td>AATADNWWNHTATT</td>
</tr>
<tr>
<td>2T2</td>
<td>AATADWSWWHTTAT</td>
<td>2A3T2</td>
<td>AATADWSWWHTATT</td>
</tr>
<tr>
<td>2A1</td>
<td>AAAADNWWNHTTTT</td>
<td>2A3C1</td>
<td>AACADNWWNHTGTT</td>
</tr>
<tr>
<td>2A2</td>
<td>AAAADWSWWHTTTT</td>
<td>2A3C2</td>
<td>AACADWSWWHTGTT</td>
</tr>
<tr>
<td>3T1</td>
<td>ACTADNWWNHTAGT</td>
<td>2A3G1</td>
<td>AAGADNWWNHTCCT</td>
</tr>
<tr>
<td>3T2</td>
<td>ACTADWSWWHTAGT</td>
<td>2A3C2</td>
<td>AAGADWSWWHTCCT</td>
</tr>
<tr>
<td>3C1</td>
<td>ACCADNWWNHTGTT</td>
<td>4T1</td>
<td>ACATDNWWNHTATG</td>
</tr>
<tr>
<td>3C2</td>
<td>ACCADWSWWHTGTT</td>
<td>4T2</td>
<td>ACATDWSWWHTATG</td>
</tr>
<tr>
<td>3G1</td>
<td>ACGADNWWNHTCGT</td>
<td>4C1</td>
<td>ACACDNNWNNHTGTT</td>
</tr>
<tr>
<td>3G2</td>
<td>ACGADWSWWHTCGT</td>
<td>4C2</td>
<td>ACACDWSWWHTGTT</td>
</tr>
<tr>
<td>2G3T1</td>
<td>AGTADNWWNHTACT</td>
<td>4G1</td>
<td>ACAGDNWWNHCCTG</td>
</tr>
<tr>
<td>2G3T2</td>
<td>AGTADWSWWHTACT</td>
<td>4G2</td>
<td>ACAGDWSWWHCCTG</td>
</tr>
</tbody>
</table>

- 102 -
A

AATACATATT TGGTCTTTTC TAATTGCTAG TTTTCAAGT TCTTAGATGC
RNA start  434 Pseudo-operator

TTTCTTTTTC TCTTTTTTAC AGATCAACAA GGAAGTAATT ATCTACTTTTT
BgIII
TCAGATCTCC CATGCTCTTA

B

<table>
<thead>
<tr>
<th>Operator</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Find pattern 2G3T1</td>
<td>AGTADNW WNHTACT</td>
</tr>
<tr>
<td>PGK pseudo-operator</td>
<td>AGTAATT ATCTACT</td>
</tr>
</tbody>
</table>

Figure 6.1. A. The region of the PGK promoter containing a 434 pseudo-operator (AGTAAATTATCTACT) which may be bound by a 434 repressor with altered amino acids within the α-3 helix.

B. A table showing the Find pattern which located the pseudo-operator within the PGK promoter and the sequence of the pseudo-operator. The gap in the centre of each sequence denotes the division between each half-site.

6.2.2.2. Construction of the 434 expression vector

This plasmid construction is represented diagrammatically in figure 6.2. Annealed oligonucleotides encoding the wild-type trp promoter (Windass et al., 1985), the consensus Shine-Dalgarno (SD) sequence (AGGAGGT) and 33bp of the 5' of the 434 cl gene and annealed oligonucleotides encoding the rrnB T1 terminator (Brosius et al., 1981) were cloned between the EcoRI and HindIII sites of pUC19 to produce pTT1. The inserted DNA was checked by sequencing.

A ~600bp EcoRI-Sau3AI fragment carrying the 434 cl gene from pRP42-76 was ligated into EcoRI/BglII cleaved pTP1 to form pTRT1. This reforms the 434 cl gene and the stop codon is retained within the Sau3A/BglII junction.
Figure 6.2 Construction of the 434 mini-library

EcoRI/HindIII
Annealed oligos for the
trp promoter and rrrN T1 terminator

BglII/EcoRI
434 gene from pRP42-76 on EcoRI/Sau3AI

KpnI/XmalIII
Annealed oligos for Ball-Xmal-Scal linker

Continued over
Selection of Altered Binding Specificity Repressors

- **Ball**
  - Tet resistance gene from pBR322 on Ball/SspI

- **Ndel/EcoRI**
  - Annealed oligos for BclI-BglII-BclI linker

- **BamHI**
  - Partial digest

- **BglII**
  - Annealed oligos for the α-3 helix coding region containing random codons

- **KpnI/XmaII**
  - Annealed oligos for the α-3 helix coding region containing random codons

- **pTRTRAN** (from over)
  - pTRTRAN containing random codons

- **pTRTTc**
  - pTRTTc containing random codons

- **pCW26**
  - pCW26 containing random codons
In order to simplify further constructions the plasmid pTRTI was digested with \textit{KpnI/XmaIII} and a \textit{KpnI-BalI-XmnI-Scal-XmaIII} linker inserted to produce pTRTRAN. The 1.6kb \textit{SspI-BalI} fragment encoding the tetracycline resistance gene from pBR322 was cloned into the \textit{BalI} site of pTRTRAN to yield pTRTTc. When pTRTTc is digested with \textit{KpnI} and \textit{XmaIII} the 3.5kb fragment required for inserting annealed oligonucleotides to alter the DNA binding helix of the 434 repressor is easily separated from the 1.6kb fragment. Unfortunately, it was found that the \textit{lac} sequences derived from pUC19 recombined with reporter plasmids derived from pAD\Delta Tc under the stress imposed on cells by the selection system. To eliminate this problem, the \textit{lac} sequences within pTRTTc were removed in the following way. A \textit{BclI-BglII-BclI} linker was ligated into the 2.1kb \textit{EcoRI-ndel} fragment from pBR322 to yield pCW26. The 2.4kb fragment from a \textit{BamHI} partial digest of pTRTTc was cloned into the \textit{BglII} site of pCW26 to yield pTRTTc\Delta lac.

The DNA coding for the \textit{\alpha-3} helix of the 434 repressor was synthesised as an oligonucleotide including the \textit{KpnI} and \textit{XmaIII} sites plus an extra 10 bp at each end. The codon for Gln29 was randomised by inclusion of all 4 bases in the synthesis step at the 3 positions in codon 29. A complementary oligonucleotide was synthesised which annealed to the first oligonucleotide up to the random codon. The remainder of the double stranded DNA was synthesised by filling in the overhanging end using the Klenow fragment of \textit{E. coli} DNA polymerase and all four deoxy nucleoside triphosphates. The DNA was then cut with \textit{XmaIII} and \textit{KpnI} and ligated into the plasmid pTRTTc\Delta lac in place of the tetracycline resistance gene. The ligation mix was transformed into DH5\textalpha{} and the randomness at codon 29 checked by double stranded DNA sequencing.
The minilibrary randomised at codon 29 cloned in pTRTΔlac was a gift from R. Wilde.

6.2.3. Screening the N29 library with operator plasmids on TPNPG and Analysis of Transformants

The N29 library together with the operator plasmid or pADΔTc (operator-minus) were introduced into E. coli 6300ΔlacU169 by electroporation, plated onto minimal media (with and without TPNPG) and then incubated at 28°C for 4 days. Approximately 20000 transformants were plated onto each TPNPG selection plate.

Following several separate experiments no colonies were detected on either the TPNPG plate for the 29-library with the operator containing plasmid or the 29-library with pADΔTc.
6.3. DISCUSSION

Wharton and Ptashne (1987) have previously shown that wild-type repressor will tightly bind the operator ACAATATATATTGT, but will not recognise the mutant operator TCAATATATATTGA. Model building based on X-ray crystallography suggested that this specificity resulted from the ability to form hydrogen-bonds between groups in the side chains of Gln^{28} and the bases at the first position in the operator. Furthermore, the replacement of Gln^{28} with various amino-acids demonstrated that this mutant operator could only be bound by an Ala^{28} mutant repressor (Wharton and Ptashne, 1987). Based on these data a test system for the TPNPG selection was devised by Wilde et al., (Personal Communication). Their results demonstrated that a TPNPG selection system is able to differentiate between Lac^+ and Lac^- strains. They also have demonstrated that from libraries randomised at positions 28 and 29 wild-type repressors are uniquely isolated from the selection when the operator 434Op2 (ACAAGATATCTTGT) is used. Also, an Ala^{28} repressor could be uniquely isolated from a library randomised at position 28 with a 1T operator (TCAAGATATCTTGA). Novel specificity 434 repressors have also been isolated from a library randomised at position 28 when a 1G operator (GCAAGATATCTTGC) is used, although the repressors isolated have yet to be fully characterised (R. Wilde, Personal Communication). This is the first demonstration that the TPNPG selection can be used to isolate repressors with an altered DNA-recognition specificity.

The reason why the AGTAGATATCTACT operator failed to give survivors under TPNPG selection can be perhaps resolved by looking in detail at the interaction between the repressor and the operator at positions 2 and 3 of the operator. Although both base-pairs
Figure 6.3. An expanded view of the interactions between Gln²⁹ of the 434 repressor and the base-pairs at positions 2 and 3 in the 434 operator. Gln²⁹ projects towards the operator DNA and the ε-NH₂ and the ε-O making hydrogen bonds to the side groups of the guanine in base-pair 2 of the operator. Van der Waals interactions are made between the methyl group of the thymine (represented by a semicircle) at position 3 in the operator and a hydrophobic pocket consisting of the side arm of Gln²⁹ and Thr²⁷, the side arm of Glu³² may also be involved. Hydrogen bonds are represented by fine dotted lines.

interact with Gln²⁹ other amino acids are also involved. The hydrophobic pocket created by the side arm of Gln²⁹ and Thr²⁷ which accommodates the methyl group of the thymine in base-pair 3 of the operator would now be orientated towards the Nitrogen at position 7 in adenine, although at an increased distance. The side arm of Gln²⁹ also makes hydrogen bonds with the guanine ring in base-pair 2 of the operator. In the altered operator this base-pair is reversed, no longer
offering the groups of guanine to the amino acid at position 29. In altering the binding specificity of the 434 repressor to that of the P22 repressor, Wharton and Ptashne (1987) altered Thr^{27} and Glu^{32} both of which are potentially involved in determining specificity for the base-pair at position 3 in the 434 operator. This is slightly confusing as this base-pair is retained in the P22 operator and in the 434^{P22} repressor similar contacts would be expected to be made as compared to the wild-type 434 repressor. However, specificity of the repressor for the equivalent of position 2 in the P22 operator is not required in the 434^{P22} repressor. This strongly implies that all three amino acids Thr^{27}, Gln^{29} and Glu^{32} need to be altered to obtain an altered binding specificity repressor. At present the ability of the TPNPG selection system to select 434 repressor genes with random codons in three positions is not proven.

Further refinement of the TPNPG selection system to include 3 randomised codons may improve the prospects of isolating an altered binding specificity repressor which binds to the operator AGTAGATATCTACT. However, even under these conditions it may prove impossible to isolate such a repressor.
Discussion and Conclusion.
7.1. Discussion

At the start of the project there were clearly defined targets for
the use of the 434 repressor. These were, 1) to assess the function of
heterodimeric repressors in *E. coli*: 2) to express the 434 repressor and
its variants in eukaryotic cells and to assess their ability to repress gene
expression: The final target fell into two sub-targets: a) to repress gene
expression either through the introduction of operators into reporter
constructs or b) to repress gene expression through repressor binding to
pre-existing operator sites.

The data presented in this thesis have shown that a heterodimer
of the 434 and 434P22 repressors is able to repress gene expression in *E.
coli* at a level very similar to that achieved with either homodimer,
demonstrating that the heterodimer is an efficient repressor *in vivo*.
Thus, the strength of the heterodimer monomer/monomer
interaction and the interaction with each half-site of the altered
operator is not affected by the presence of altered DNA-binding
determinants in one of the repressor monomers. This information
allows the heterodimer to be considered, alongside the 434 homodimer
and the 434P22 homodimer, for use in repressing gene expression in
eukaryotes.

As part of the desire to use the 434 repressor and its variants to
repress gene expression in eukaryotes, it is desirable to understand the
influence of all parts of the operator on the affinity of the repressor. To
be able to target repressors to pre-existing operator sequences requires
that the influence of the non-contacted bases on the strength of the
protein-DNA interaction is understood. This was investigated *in vitro*
by Koudelka *et al.* (1987; 1988). Their findings showed that the 434
repressor bound with higher affinity to operator sequences containing
A:T base pairs in the central non-contacted region of the operator. If these A:T base-pairs were altered to G:C base-pairs the affinity of the operator for the repressor could be lowered as much as 50-fold. This reduction in affinity could be relieved to some extent if a mutant repressor was used in which it was proposed that the mutant amino acid increased the flexibility of the repressor to counter the increase in the torsional rigidity of the operator. Results presented in chapter 3 show that in vivo the effect of G:C base-pairs on the efficiency of repression is similar to the effect on the affinity of the operator to the repressor in vitro. However, the effect of the Ala⁴⁴ mutation on the levels of repression in vivo does not reflect the effect on the affinity of the operator for the repressor in vitro. The reasons for this discrepancy are not fully clear, although it has been established that the 434Ala⁴⁴ repressor is not suitable for the targeting of repression to operator sites which contain G- and C-rich central regions. Indeed, it is evident that such operator sites could not be used as a target for high-level repression by currently known variants of the 434 repressor. Further work on the in vivo interactions between the 434 repressor and its operator, together with their influence on the block to the passage of RNA polymerase and possibly the structure of the 434Ala⁴⁴ repressor may clarify the reason for these differences.

Results presented in this thesis have also shown that the 434 operator/repressor system is able to function in Saccharomyces cerevisiae. The repression data demonstrated that the repression levels are strongly dependent on the position of the inserted operator. The results obtained are comparable to those obtained by Smith et al. (1988) working with the lexA repressor in mammalian cells. However, the levels of repression are considerably weaker than those seen by Brent and Ptashne (1984), using the lexA repressor.
In general the levels of repression seen with the 434 operator/repressor system in yeast are very much weaker than other bacterial operator/repressor systems used to regulate gene expression in eukaryotes. There are a number of possible reasons for this relative inefficiency. The promoter used for the repression studies in this thesis was the yeast PGK promoter. The PGK gene is very highly expressed, the protein product accounting for up to 5% of cell protein (Holland and Holland, 1978). Thus the 434 repressor may not be able to demonstrate high levels of repression against such strong expression. This possibility could easily be investigated by the construction of a reporter construct driven from a weaker promoter. If the levels of repression are dependent on the strength of the promoter, then higher levels of repression would be seen. This, however, may not be the case and the 434 repressor may simply be unable to produce high levels of repression in the context of eukaryotic chromatin.

Another reason for the ineffectiveness of the 434 repressor in yeast may be due to the strength of the interaction between the 434 repressor and its operator. This is similar to the lexA and tet repressors on binding to their operators, all three of which have a $K_D$ of $10^{-9}$M (Brent and Ptashne, 1984, Koudelka et al., 1988, Takahashi et al., 1986). The 434 repressor produces similar repression to the lexA repressor when the lexA repressor is used in mammalian cells (Smith et al., 1989), but lower then the lexA repressor in yeast and the tet repressor in yeast (Dingermann et al., 1992) and in plants (Gatz and Quail, 1988; Gatz et al., 1991).

Some other bacterial repressors bind to their operators with a much higher affinity, particularly the lac repressor, where the $K_D$ has been measured at $10^{-13}$M (Riggs et al., 1970). The lac repressor has been used in a number of studies for the regulation of gene expression in
Discussion and Conclusion

eukaryotes and high levels of repression achieved (Hu and Davidson 1987, Wilde et al., 1992). It is possible that the higher levels of repression seen with the lac repressor, compared to the 434 repressor, is due to the higher affinity with which it binds to its operator. However in view of the fact that the tet and lexA repressors are both capable of high levels of repression in eukaryotes other considerations may also be involved.

Although the gene encoding the 434 repressor could readily be introduced, it proved impossible to transfec an active gene for the 434P22 repressor into yeasts. It is tempting to speculate that the 434P22 repressor is repressing gene expression at a higher level than the wild-type 434 repressor, in view of this severe effect on the growth of host yeast. However, this effect may be due to repression at more than one gene or at genes which are expressed at a considerably lower level than expression from the PGK promoter, therefore making direct comparison impossible. Study of the ability of the 434P22 repressor to repress gene expression in eukaryotes, could only be done in yeast with great difficulty and might more readily be carried out using mammalian, plant cell culture or transgenic plants as the host system.

The use of heterodimeric 434 repressors is also limited, at present, by the effects the 434P22 repressor manifests in yeast. Until other variants of the 434 repressor become available which have a binding affinity similar to that of the wild-type 434 repressor, efficient heterodimers may not be available to test in the yeast reporter system described in this thesis. The continuing work of Wilde et al. on the development of a selection system for altered binding-specificity repressors may produce other 434 repressor variants which could potentially allow the formation of an efficient heterodimer consisting of two non-wild-type 434 repressors. This work will greatly enhance the
range of targets the 434 repressor could be used to repress in any transgenic system. However, as work in this thesis has shown, the number of possible DNA-binding variants of the 434 repressor is likely to be limited.

7.2. Conclusions

The potential of the 434 repressor to be used as a back-bone for the generation of DNA-binding proteins with which to regulate the expression of genes is beginning to be revealed. The use of heterodimers increases the chances of finding pseudo-operators in positions where the binding of a repressor will influence the expression of an adjacent gene. The demonstration that the 434 repressor is able to function in eukaryotes furthers the aim of being able to repress expression of a given eukaryotic gene. This aim is likely to be fulfilled by further research into the generation of specific, altered-specificity repressors.
CHAPTER 8.

Materials and Methods
8. MATERIALS AND METHODS

General cloning methods were performed according to Sambrook et al. (1989), and all enzymes purchased commercially were used according to the manufacturers instructions.

8.1. Transformation of *E. coli* cells.

8.1.1. Preparation of Competent Cells

A single colony was picked from an agar plate and grown overnight in 10mls ψ broth (ψB is: 0.5%(w/v) Bacto Yeast Extract, 2.0%(w/v) Bacto Tryptone and 0.5%(w/v) MgSO₄). A 1ml aliquot was diluted 1:100 in ψ broth prewarmed to 37°C and the culture grown to an optical density at 550nm of 0.5. It was chilled on ice for 15 min and the cells collected as a pellet by centrifugation at 4°C, 3,500 rpm for 5 min. The pellet was carefully drained of ψB, the cells resuspended in 40ml of TbfI (TbfI is: 30mM Potassium acetate, 100mM RbCl, 10mM CaCl₂, 50mM MnCl₂ and 15%(v/v) glycerol at pH 5.8 and filter sterile) and incubated on ice for 5 min. The cells were collected by a second centrifugation as above and the pellet resuspended in 4mls TbfII (TbfII is: 10mM MOPS pH 6.5, 75mM CaCl₂, 10mM RbCl and 15%(v/v) glycerol, filter sterile). Following a 15 min incubation on ice the cells were flash frozen in 200µl aliquots in microfuge tubes using liquid nitrogen. They could be stored in this state at -70°C until required.

8.1.2. Introduction of DNA.

An aliquot of competent cells was thawed on ice for 15 min. 50-100ng of DNA from a ligation, or 5-10ng of closed circular DNA, was chilled on ice while the cells were thawing. The DNA was added to the cells, mixed by gentle flicking and incubated on ice for 1 hour. The
DNA cell mix was heated to 42°C for 2-3 min to heat shock the cells and returned to ice for 10-30 min.

8.1.3. Treatment of Transformed Cells Carrying a Plasmid.

1ml of 2xYT (16g/l Bacto-tryptone, 10g/l Bacto-yeast extract, 5g/l NaCl) was added and the cells incubated at 37°C with shaking for 1 hour. The cells were collected by centrifugation at full speed for 30 secs in a micro centrifuge the 2xYT discarded and the pellet resuspended in 100μl fresh 2xYT. Dilutions were prepared at 1:10, 1:10² and 1:10³ and the cells plated out on LA (10g/l Bacto-Tryptone, 5g/l Bacto-Yeast Extract, 10g/l NaCl, 15g/l agar) including appropriate antibiotics and chromogenic selective agents if required. Transformed cells could be seen as discrete colonies on one or more of the plates following an overnight incubation at 37°C.

<table>
<thead>
<tr>
<th>Stocks:</th>
<th>Final Concentration:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ampicillin 100mg/ml in 50% ethanol.</td>
<td>100μg/ml</td>
</tr>
<tr>
<td>Kanamycin 50mg/ml in H₂O.</td>
<td>50μg/ml</td>
</tr>
<tr>
<td>Chloramphenicol 50mg/ml in 50% ethanol.</td>
<td>50μg/ml</td>
</tr>
<tr>
<td>Tetracycline 7.5mg/ml in 50% ethanol.</td>
<td>7.5μg/ml</td>
</tr>
<tr>
<td>X-gal 40mg/ml in Dimethyl formamide</td>
<td>40μg/ml</td>
</tr>
<tr>
<td>IPTG 20mg/ml in H₂O</td>
<td>4μg/ml</td>
</tr>
</tbody>
</table>

8.1.4. Treatment of Transformed Cells Containing M13 Vectors.

To three 4ml aliquots of molten Top agar (Bacto-Tryptone 10g/l, Bacto-Yeast Extract 5g/l, NaCl 10g/l Agar 7g/l) at 42°C were added 20μl of an overnight culture of E. coli XL1-blue (plus 20μl X-gal solution and 4μl IPTG solution, if required). The transformed cells were added to one aliquot and dilutions at 10⁻¹, 10⁻² and 10⁻³ prepared in the remaining aliquots. The aliquots were then poured onto preset LB plates containing 7.5μg/ml tetracycline and allowed to set before the
plates were inverted and incubated at 37°C overnight. Transformed cells give rise to plaques in the bacterial growth in the top agar.

8.2. Mini Preparation of Plasmid DNA from *E. coli*.

A 5ml overnight culture was grown from a single colony or a glycerol stock in 2xYT under appropriate antibiotic selection. A 1.6ml aliquot was spun down in a microcentrifuge tube, the pellet resuspended in 100μl sterile water and incubated on ice for 5min. 200μl of alkaline SDS (1% SDS, 0.2M NaOH) were added, mixed and the cells incubated on ice a further 5min. Following this 150μl of potassium acetate solution (5M acetate, 3M potassium) were added, the solution gently mixed and incubated on ice a final 5min. Precipitated matter was collected by centrifugation and discarded, the supernatant was retained. RNase A was added to 100μg and the solution incubated at 37°C for 1 hour. Precipitated protein was removed by phenol extraction. The plasmid DNA was recovered by ethanol precipitation.

8.3. Midi-preparation of Plasmid from *E. coli*.

Midi-preps of plasmid DNA from *E. coli* were prepared for the transformation of yeast. These were prepared using the column purification methods of Qiagen.

The strain of *E. coli* containing the plasmid of interest was grown overnight at 37°C with vigorous agitation in a 2l conical flask containing 250ml 2xYT with appropriate antibody selection. The following morning the culture was transfered to a 250ml centrifuge tube and the cells harvested by centrifugation at 5000rpm for 5min in a Sorval GSA rotor in a Sorval RC-5B centrifuge. The culture media was discarded and the cells resuspended in 10ml solution P1 (50mM Tris
8 - Materials and Methods

HCl pH 8.0, 10mM EDTA pH 8.0, 100μg/ml RNase A). To the cell suspension was added 10ml solution P2 (0.2N NaOH, 1% SDS), the solution incubated for 5min at room temperature then 10ml solution P3 (2.55M potassium acetate, pH 4.8) added. The solution was mixed briefly and centrifuged at 12000rpm in a 35ml centrifuge tube in a Sorval SS-35 rotor in a Sorval RC-5B centrifuge at 4°C for 20min.

A Qiagen-tip 500 column was equilibrated by the passage of 10ml solution QBT (50mM MOPS pH 7.0, 750mM NaCl, 15% (v/v) ethanol, 0.15% (v/v) Triton X-100) through the column by gravity flow. The supernatant from the 35ml centrifuge tube was applied to the column and allowed to empty through the equilibrated column by gravity flow. The column was washed with 30ml solution QC buffer (50mM MOPS pH 7.0, 1M NaCl, 15% (v/v) ethanol) and the column allowed to empty by gravity flow. The DNA was eluted from the column with 15ml buffer QF (50mM MOPS pH 8.2, 1.25M NaCl, 15% (v/v) ethanol) which was allowed to pass through the column by gravity flow and the elutate collected. The elutate was mixed with 0.7 volumes of isopropanol in a 30ml Corex tube and the DNA collected as a pellet by centrifugation in a Sorval LB-4 swingout rotor at 12000rpm and 4°C for 30min. The supernatant was discarded and the DNA pellet air dried for approximately 1hr. The DNA pellet was then resuspended in 500μl sterile deionised water and the concentration determined by measurement of the OD_{260}.

8.4. Preparation of M13 phage DNA

A culture of E. coli XL1-Blue was grown overnight in 2xYT with 7.5μg/ml tetracycline. The following morning 50μl was diluted into
5ml of fresh media and 100μl of a primary phage stock added. The culture was grown at 37°C with vigorous agitation for 5-6 hours.

8.4.1. For RF DNA.

The culture was then treated as for a mini prep of plasmid DNA (see above). Except that the supernatant from the first centrifugation step was retained as a high titre phage stock.

8.4.2. For single stranded DNA.

1.5mls of the culture were spun in a microfuge tube at 13,000 rpm for 5 minutes and the supernatant retained. The supernatant was transferred to a fresh microfuge tube and to it was added 200μl of a solution of 2.5M NaCl and 20% PEG 6000. The tube was left to stand at room temperature for 15 min and the phage particles collected by centrifugation at 13,000 rpm for 10 minutes. The phage pellet was carefully drained and the supernatant discarded. The pellet was resuspended in 100ml of TE (pH 8.0) and an equal volume of TE equilibrated phenol added. The mixture was vortexed briefly and the two phases separated by centrifugation. A second extraction was performed with phenol/chloroform/IAA as above and finally a chloroform extraction. The phage DNA was Precipitated by the addition of 2.5 volumes of ice cold absolute ethanol and 0.1 volumes of 3M sodium acetate. The DNA was collected as a pellet by centrifugation at 13,000 rpm in a microfuge for 15 minutes.

8.5. Sequencing of DNA.

Two methods were employed for this, each depending on the source DNA, ie double-stranded plasmid DNA or single-stranded
phage DNA. Each method is taken from the Pharmacia T7 sequencing kit booklet.

8.5.1. DNA Preparation for Plasmid Sequencing

For double stranded plasmid DNA a mini prep was prepared as above, the pelleted DNA resuspended in 10 µl 0.4 M NaOH solution and in order to denature the DNA it was incubated at room temperature for 10 minutes. The DNA was precipitated by the addition of 3 µl 3 M sodium acetate pH 5.4, 7 µl of sterile deionised water and 60 µl ice cold absolute ethanol. Following incubation on dry ice for 10 min the DNA was pelleted by centrifugation for 15 min at full speed in a microfuge.

8.5.2. DNA Preparation for Single-Stranded Phage Sequencing.

For sequencing of single stranded phage DNA the method of preparation of the phage DNA is as described in 8.4.2, except that the growth period of the infected culture was increased to 16 hours. The pellet of phage DNA could be used directly in the sequencing protocol bellow.

8.5.3. Sequencing Reactions.

Both protocols were identical from now on. The pellet of DNA was briefly dried and resuspended in 10 µl distilled water. The primer for the sequencing reaction at 5 ng/µl in 2 µl was added along with 2 µl of annealing buffer (280 mM Tris HCl pH 7.5, 100 mM MgCl₂, 350 mM NaCl). This was incubated at 37°C for 20 min then at room temperature for at least 10 min. The reaction can be stored in this state indefinitely at -20°C. To continue the sequencing reaction 1 µl of a 35S αATP at 370 MBq/ml was added, 2 µl T7 DNA polymerase (Pharmacia) at 1.5 units/µl diluted from a stock by the addition of 1x enzyme dilution buffer (this was supplied as a 10x solution with the enzyme), and 3 µl of label mix.
8 - Materials and Methods

(Label mix is: 2µM dGTP, 2µM dCTP, 2µM dTTP). The reaction was incubated at room temperature for 5 min. To 2.5 µl aliquots of each of the 4 nucleotide termination mixes (150 µM dGTP, 150 µM dCTP, 150 µM dATP, 150 µM dTTP, 10 mM MgCl₂, 40 mM Tris HCl pH 7.5, 50 mM NaCl with for G, 15 µM ddGTP, for A 15 µM ddATP, for T 15 µM ddTTP, and for C 15 µM ddCTP), prewarmed to 37°C, 4.5 µl of the above reaction was added and all were incubated at 37°C for 5 min. After this time 5 µl of stop buffer (98% Formamide, 10 mM EDTA pH 8.0, 0.025% Bromophenol Blue, 0.025% Xylene Cyanol) were added. The reactions could be stored in this state at -20°C or loaded onto an acrylamide sequencing gel. Before electrophoresis it was necessary to denature the DNA in the reaction mixture by heating in a boiling water bath for 2 min.

Sequencing gels (6% acrylamide, 50% urea 1x TBE) were formed between two glass plates with 0.4 mm spacers and run in 1x TBE.

8.6. Purification of Restriction Fragments from Agarose

This was carried out by a variation of the method of Langridge et al. (1979).

Equal volumes of butan-1-ol and distilled water were mixed and equilibrated by shaking, the layers allowed to settle then separated. A 1%(w/v) solution of CTAB (hexadecyltrimethylammonium bromide) was made in the butanol fraction. The two phases were again mixed and equilibrated. When the aqueous and butanol phases had settled out they were separated and bottled separately. These solutions may be stored at room temperature indefinitely.

The required sized band of DNA electrophoresed in LMP agarose, using TBE buffer (TBE is: 90 mM Tris-borate 2 mM EDTA), was
cut from the gel in a minimum of agarose. The volume was doubled by adding an equal volume of the CTAB equilibrated water and the agarose was melted by heating to 65°C for 10 min. Once melted the agarose must not be allowed to re-set. For volumes less than 1ml, 300μl of CTAB-equilibrated butanol were added, the tube shaken vigorously, but not vortexed, and the two phases separated by a brief centrifugation. The butanol phase was retained and a second extraction performed on the aqueous phase. A final extraction using 200μl CTAB equilibrated butanol was performed and the three butanol phases pooled. The aqueous phase was discarded. The DNA was recovered from the butanol by extraction with 200μl of 200mM NaCl twice. The CTAB was precipitated from the two NaCl fractions by adding an equal volume of CHCl₃ and incubating on ice for 30 min shaking occasionally. The sample was centrifuged for 30 secs and the top aqueous layer retained. The DNA was collected by ethanol precipitation with the addition of 1μl 20mg/ml glycogen. The glycogen did not appear to affect subsequent ligations or labelling reactions.

8.7. Radiolabelling of DNA

8.7.1. Restriction Fragments of DNA.

Restriction fragments of DNA purified from agarose by the method above or fragments derived from PCR reactions were labelled by the following method. 10ng of DNA in 10μl of sterile deionised water was heated to 100°C in a boiling water bath for 5 min then the following added sequentially; 3μl oligolabelling buffer (1M HEPES pH 6.6, 200mM Tris pH 8.0, 25mM MgCl₂, 100μM dGTP, 100μM dATP, 100μM dTTP, 5mM 2-mercaptoethanol, 27 OD₂₆₀ units/ml hexanucleotide (Pharmacia)) 0.6μl 10mg/ml BSA, 1μl ³²P α dCTP at
370MBq/ml (Amersham), 0.6μl 1U/μl Klenow fragment DNA polymerase (BRL). The reaction was incubated at room temperature for at least 6hrs and 85μl stop solution (20mM NaCl, 20mM Tris HCl pH 7.5, 2mM EDTA, 0.25% SDS) added. Prior to adding to hybridisation solution the DNA was denatured by heating to 100°C in a boiling water bath for 5 min.

8.7.2. End-Labelling of Oligonucleotides.

This was performed as described in Sambrook et al. (1989) for phosphorylating the 5' end of oligonucleotides except that 0.37MBq of γ³²P ATP (Amersham) were substituted for cold ATP.

8.8. Transformation of Yeast.

8.8.1. Preparation of competent yeast.

An overnight culture of *Saccharomyces cerevisiae* was diluted to 2.5x10⁶ cells/ml in YPD (1% w/v yeast extract, 2% w/v peptone and 2% w/v glucose) and grown until there were 10⁷ cells/ml. The cells were precipitated by centrifugation at 3.5k for 5min in a Sorval RT6000B centrifuge and the supernatant discarded. The cells were washed in TE (10mM Tris HCl pH 7.5, 0.1mM EDTA) several times and resuspended in 1/10 the original culture volume lithium acetate solution (Lithium acetate solution is: 0.1M lithium acetate in TE pH 7.5). They were incubated at 30°C for 1 hour with light agitation and divided into 300μl aliquots.

8.8.2. Introduction of DNA.

DNA was added to the 300μl aliquots at 1μg for closed circular plasmids and 10μg for linearised plasmids. 700μl of a 50%(w/v) solution of PEG4000 in sterile deionised water was added and incubated at 30°C for 30 minutes without agitation. The cells were heat shocked at
42°C for 5 minutes. An aliquot of 100μl was plated directly onto selective media and the remainder harvested by centrifugation in a microfuge at 4.5k for 1min. The PEG solution was removed and the cells resuspended in 100μl TE (pH 7.5) and plated onto selective media. Transformants appear as individual colonies after 3-7 days at 30°C.

8.9. Isolation of Genomic DNA from *Saccharomyces cerevisiae*.

A 20ml overnight culture of *Saccharomyces cerevisiae* was transferred to a centrifuge tube and the cells pelleted at 3500rpm for 5min and room temperature in a Sorval RT 6000B centrifuge. The broth was discarded, the pellet of cells resuspended in 1ml 1M sorbitol, 25mM EDTA (pH 8.0), 8mg/ml DTT and incubated for 15min at 30°C. The cells were again collected by centrifugation, the supernatant discarded and the pellet resuspended in 1ml, 1.2M sorbitol, 100mM sodium citrate, 10mM EDTA (pH 8.0) and 10μl 10mg/ml zymolyase added. The suspension was incubated at 30°C for 30min. The cells were collected by centrifugation, washed twice in 1.2M sorbitol, resuspended in 0.5ml 3% sarkosyl, 500mM Tris HCl (pH 7.6) 200mM EDTA and 5μl 10mg/ml proteinase K added. The suspension was incubated at 55°C for 1hr. The volume was made up to 5ml with TE (10mM Tris HCl, 1mM EDTA pH 8.0) and extracted with an equal volume of phenol:chloroform:isoamylalcohol (25:24:1) until the interface was clean. Nucleic acids were collected by ethanol precipitation and the pellet resuspended in 500μl TE (pH 8.0) and 10μl 10mg/ml RNase A added. Following incubation at 37°C for 30min a further phenol extraction was performed followed by a chloroform extraction. The DNA was ethanol-precipitated, resuspended in 500μl sterile deionised
water and the yield determined by measuring the OD$_{260}$ compared to the OD$_{280}$.

**8.10. Southern Blotting.**

Genomic DNA was isolated from yeast as described above and 10µg digested with an appropriate restriction endonuclease. The digested DNA was run on a 0.8% agarose (BRL ultra pure) gel in TBE buffer. The DNA was transferred to a charged nylon membrane (Amersham, Hybon N) by the method of Southern (1975) using the variation of Chomczynski and Qasba (1984). DNA was transferred to the membrane by capillary transfer using 0.5M NaOH, 1.5M NaCl, the membrane was neutralised in 1M Tris HCl pH 7.0, 1.5M NaCl and allowed to dry at room temperature. The DNA was fixed to the membrane by exposure to UV light on a transilluminator for 45 seconds.

**8.11. Hybridisation of probes to DNA immobilised on nylon membranes**

The membrane with the attached DNA was placed in prehybridisation solution at 65°C with constant weak agitation for at least 4 hours. The denatured, radiolabelled, DNA probe was added and hybridisation continued at the same parameters overnight. The membrane was removed from the hybridisation solution and washed at room temperature in 6x SSC, 0.1% SDS for several minutes to remove unbound probe. Nonspecifically bound probe was removed by washing at 65°C in 6x SSC, 0.1% SDS for 15min. If required a second wash was performed for 15min at 65°C in 1x SSC, 0.1% SDS. Specific
DNA or RNA species to which the probe has hybridised can be detected by exposure to X-ray film (Kodak type XAR).


Both for *E. coli* and *Saccharomyces cerevisiae* the method used for assay of β-galactosidase was a variation on the method of Miller (1972).

8.12.1. For *E. coli*.

An overnight culture of *E. coli* 6300ΔlacU169 containing the plasmids of interest was diluted 1:20 into 5mls of minimal media (35mM Na$_2$HPO$_4$, 10mM KH$_2$PO$_4$, 4mM NaCl, 9mM NH$_4$Cl, 2mM MgSO$_4$, 0.1mM CaCl$_2$, 2% Glucose) with antibiotic selection, and grown at 37°C with vigorous agitation. Duplicate 360µl samples were taken at 1, 3 and 5 hours from inoculation. One sample was used to determine the OD$_{600}$, the other the β-galactosidase activity.

Each sample was mixed with 90µl 5xZ buffer (300mM Na$_2$HPO$_4$, 200mM NaH$_2$PO$_4$, 50mM KCl, 5mM MgSO$_4$, 250mM 2-mercaptoethanol, filter sterile), 25µl CHCl$_3$ and 4.5µl 0.1% SDS, vortexed for 10 seconds and left on ice until all the samples have been taken over the time course of the experiment. 200µl aliquots from each sample time were transferred to individual wells of a 96-well microtitre plate, avoiding the CHCl$_3$, and dilutions prepared at 10$^{-1}$ and 10$^{-2}$ in 1xZ buffer. In a separate microtitre plate 264µl 1xZ buffer was pipetted into each well (300µl for the blank wells) and 36µl of each sample transferred from the previous plate to individual wells and mixed by pipetting.

When all samples have been transferred 60µl of 4mg/ml ONPG in sterile deionised water was added to each well and the plate
incubated at room temperature. The OD\text{550} was determined for each well then the OD\text{420} determined for each well at 5, 10 and 15 min.

The results were treated as follows; each reading was normalised to take into account any cell debris by subtracting a treatment of the OD\text{550} from the OD\text{420}:

\[
\Delta \text{OD} = \text{OD}_{420} - (1.3 \times \text{OD}_{550})
\]

The \(\Delta \text{OD}\) for readings at 5, 10 and 15 minutes were plotted against time and the rate determined. Each rate was plotted against the OD\text{600} of the culture at the time the sample was taken and the final rate determined. Results were expressed as units/ml/min/OD\text{600}.

Calculations were performed on an Amstrad PC running a home written program (Appendix 1).

8.12.2. For \textit{Saccharomyces cerevisiae}.

The strains to be assayed were grown overnight in quadruplicate in semidefined minimal selective media (6.7\% Yeast Nitrogen Base without amino acids, 2\% glucose, leucine 40\(\mu\)g/ml, histidine 40\(\mu\)g/ml) at 30\(^\circ\)C with vigorous shaking. The OD\text{600} of each culture was determined immediately prior to determining the \(\beta\)-galactosidase activity. A sample of 440\(\mu\)l was taken and mixed with 110\(\mu\)l 5\(\times\) Z buffer, a sample of this was diluted 1/10 in 1\(\times\) Z buffer. To each was added 5\(\mu\)l 0.1\% SDS and 25\(\mu\)l chloroform and the sample vortexed for 10 sec. To start the reaction 90\(\mu\)l of 4mg/ml ONPG was added, the reactions incubated at 30\(^\circ\)C for 30 min then the reaction stopped by the addition of 250\(\mu\)l 1M Na\textsubscript{2}CO\textsubscript{3}. Cells were cleared from the solution by a 30 sec spin in a microfuge at full speed and the OD\text{420} determined for each sample and the 1/10 dilution.
Results were treated as follows; the OD$_{420}$ was corrected for any dilutions, multiplied by 1000 and divided by 30 min and the OD$_{600}$ of the culture. This gives the results in β-galactosidase units/OD$_{600}$.

8.13. Preparation of Protein From *E. coli* and *Saccharomyces cerevisiae*.

Cells were harvested from a 1.5ml aliquot of an overnight culture by a 30 second centrifugation in a microcentrifuge. The cells were pelleted, the supernatant discarded and the cells resuspended in 150µl cracking buffer (62.5mM Tris.HCl pH 7.5, 3% SDS, 10% Glycerol, 5% β-mercapto-ethanol). It was incubated at 65°C for 5 min then at 100°C for 10 min. Prior to heating yeast were vortexed for 30 sec with 0.5mm glass beads. Debris was precipitated by centrifugation at full speed for 15 min in a microcentrifuge. The supernatant was transferred to a fresh tube and 1:100 dilutions prepared for determination of the protein concentration by the Bradford method using a 1:100 dilution of the cracking buffer as a blank.


Protein standards were made up in 0.01x cracking buffer with a 0.01x cracking buffer blank then mixed with Bradford reagent (Bradford reagent is: 0.1%(w/v) Coomassie Brilliant Blue, 8.5%(v/v) Ortho phosphoric acid) in a ratio of 1:20 (sample:reagent). The samples were incubated at room temperature for two minutes and the O.D.$_{600}$ determined within 1 hour.
8.15. Western Analysis of Proteins From *Escherichia coli* and *Saccharomyces cerevisiae*

8.15.1. Polyacrylamide SDS Gel Electrophoresis

Protein extracts of known concentration in cracking buffer were aliquoted to the required quantity of protein and an equal volume of loading buffer added (100mM Tris HCl pH 6.8, 200mM DTT, 4% SDS, 0.2% bromophenol blue, 20% glycerol). A 10% acrylamide SDS resolving gel (10% Acrylamide, 0.075% SDS, Tris HCl pH 8.8) was cast between two glass plates to 3/4 of their height. The gel/air interface was covered with a layer of butanol-saturated water during polymerisation to ensure a smooth interface was maintained. When the gel had completely set the butanol was washed away with deionised water and excess water removed with filter paper. A stacking gel (5% Acrylamide, 0.075% SDS, 125mM Tris HCl pH 6.8) was added on top of the resolving gel and a suitable comb inserted before the gel polymerised. Once completely set the gel was placed in the tank and running buffer added (25mM Tris, 0.1% SDS, 250mM Glycine). The comb was carefully removed and protein samples in loading buffer applied following heating to 100°C for 5min. Current was applied and the gel was run until the Bromophenol blue reached the bottom.

8.15.2. Western Blotting

The Polyacrylamide gel containing the protein to be transferred was equilibrated in transfer buffer (25mM Tris, 50mM glycine, 25% (v/v) Methanol) for 5 min on a piece of 3MM paper. A piece of nitrocellulose paper was cut to the size of the gel, wetted in transfer buffer and placed over the gel avoiding any bubbles. A second piece of 3MM paper was placed over the nitrocellulose and the assembly clamped between two fibre pads. The gel was placed in a transfer tank
with the nitrocellulose between the gel and the anode. Current was applied and transfer carried out at 4°C. Protein transferred to the nitrocellulose was visualised by reversible staining with Ponceau S stain (500mM Ponceau S in 5% w/v Trichloro acetic acid) for 5min at room temp. Excess dye was removed by washing in deionised water. The stain was removed from the protein attached to the membrane by washing in Tris-Saline-Tween (0.1%) (10mM Tris HCl pH 7.5, 0.9% NaCl w/v, 0.1% (v/v) Tween 20).

8.15.3. Detection of antigens

The nitrocellulose filter was incubated in Tris Saline Marvel (10mM Tris HCl pH 7.5 0.9% w/v NaCl, 5% Marvel(Cadburys)) for between 1 and 18 hours with mild agitation at room temperature. It was washed for 2x 5min in Tris Saline Tween (0.1%) and incubated in Tris Saline Tween (0.02%) BSA (10mM Tris HCl pH 7.5, 0.9% w/v NaCl, 0.02% v/v Tween 20, 0.1% w/v BSA) with the primary antiserum at 1:20000 dilution for 1hr. The anti-434 repressor antiserum was a kind gift from ICI Pharmaceuticals. The filter was washed in Tris Saline Tween (0.1%) for 2x 5min and incubated in Tris Saline Tween (0.02%) containing an antirabbit IgG peroxidase conjugate (Sigma) at room temperature for 30min. The filter was again washed for 2x 5min in Tris Saline Tween (0.1%) at room temperature before the antigen/antibody/anti-IgG-peroxidase conjugate was visualised by incubation at room temperature in development solution (10mM Tris HCl pH 7.5, 0.9% w/v NaCl, 200μg/ml 3,4,3',4'-tetra-amino biphenyl hydrochloride, 0.005% Hydrogen peroxide) for 1 to 5min.
8. Site Directed Mutagenesis.

8.16. Preparation of template.

This was carried out by a variation of the method of Kunkel et al. (1987). An overnight culture of E. coli RZ1032 (Kunkel et al., 1987) was grown in 2.5ml 2x YT (2xYT is: 10g/l bactotryptone, 15g/l yeast extract, 10g/l NaCl, pH 7.5) with 7.5μg/ml tetracycline. The following morning phage was added as either 0.5ml of a primary phage stock or 200μl of a high titre phage stock. The culture was left at room temperature for 30 min then diluted into 250ml of 2x YT with 7.5μg/ml tetracycline and grown for 6 hours at 37°C with vigorous agitation. After this time the cells were removed by centrifugation at 6,000 rpm for 20 min in a Sorval centrifuge. The supernatant was placed in a fresh centrifuge tube and 7.5g of NaCl added along with 10g PEG (polyethyleneglycol) 6000. The solution was stirred at room temperature for 30 min. Phage particles which were precipitated by this may be collected by centrifugation at 6,000 rpm for 20 minutes in a sorval centrifuge. The pellet of phage particles, drained of all supernatant, was resuspended in 10ml TE (pH 8.0) by vigorous vortexing. The protein coats were removed by extraction with an equal volume of phenol saturated with TE, phenol/chloroform/isoamylalcohol and finally chloroform. The DNA was precipitated by the addition of 2.5 volumes of ice cold absolute ethanol and 0.1 volumes of 3M sodium acetate and collected by centrifugation at 10,000 rpm for 30 minutes in a sorval centrifuge. The pellet was drained and briefly dried before being resuspended in 0.5ml distilled water. The concentration was determined by reading the OD$_{260}$ and the concentration adjusted to 1pmol/ml. The DNA so prepared contains U residues in place of T residues at a small number of positions, this
makes it unstable in strains of *E. coli* which do not carry the *ung* mutation.

### 8.16.2. Preparation of the Mutagenic Oligonucleotide

For the introduction of restriction sites as a replacement of several bases in the sequence of the target DNA oligonucleotides were designed with a 15 base clamp either side of the mutation. Following extensive cleaning of the oligos by washing in ice cold 70% ethanol they were resuspended in distilled water. Each was kinased in a 50μl reaction containing 50pmol of the oligo.

### 8.16.3. Mutagenesis Reaction.

To anneal the oligo to the template 1pmol of the template and 10pmol of the kinased oligo were mixed in 20ml and 1.2ml of 20x SSC added. They were heated in a boiling water bath for 5 minutes and allowed to cool to room temperature very slowly. The complementary strand of the DNA was synthesised using the mutagenic oligonucleotide as a primer in a 100ml reaction containing each of the following, 500mM of each dNTP, 1mM ATP, 10mM MgCl₂, 100mM Tris pH 7.5, 2mM DTT, 20 units of T4 DNA polymerase and 20 units of T4 ligase. The reaction was incubated on ice for 5min, at room temperature for 5 minutes and at 37°C for 2 hours, the reaction was terminated by the addition of 5ml 0.5M EDTA.

### 8.17. Statistical Calculations

#### 8.17.1. Linear Regression

The linear regression used to calculate degree of fit of the straight lines involved in the calculation of β-galactosidase expression from *E. coli* is based on the equation:
### Materials and Methods

\[
r = \frac{n \sum xy - \sum x \sum y}{\sqrt{\left[ n \sum x^2 - (\sum x)^2 \right] \left[ n \sum y^2 - (\sum y)^2 \right]}}
\]

The value of "r" varies between -1 and +1 according to the slope of the line. A value of -1 describes a set of data with perfect fit to the straight line equation with a negative value of m. A value of +1 describes a set of data with perfect fit to the straight line equation with a positive value of m.

The straight line equation is:

\[
y = mx + c
\]

where y is the value of the y coordinate, x is the value of the x coordinate, c is a constant and m is the slope of the line.

#### 8.17.2 Standard Deviation of Repression

The standard deviation for percentage repression was calculated by the propagation of random error as follows: The average and the standard deviation of the control data and repressed data were calculated. Assuming the data fits to a normal distribution, which was reasonable for this kind of data, then the standard deviation of the percentage repression can be calculated by the formula for the propagation of standard error by multiplication and division, using the equation below.

\[
S_y/y = \sqrt{[(S_a/a)^2 + (S_b/b)^2]}
\]
where $S_y$ is the standard deviation of the treated result which results from the multiplication or division of $a$, the average of data set $a$ with a standard deviation of $S_a$, and $b$, the average of data set $b$ with a standard deviation of $S_b$, $y$ is the result of the multiplication or division of $a$ and $b$.

8.17.3. The Students $t$-Test.

The students $t$-test is used for a comparison of two small data sets, with a normal distribution, to evaluate if there is any significant difference between them. Thus for repression data the $t$-test answers if the difference between the control and the repressed data is significant. To calculate this the following equation is used:

$$t_{calc} = \frac{x - y}{\sqrt{\frac{S_x^2 + S_y^2}{2}} \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}}$$

where $x$ is the average of data set $x$ and $y$ is the average of data set $y$, $S_x$ is the standard deviation of $x$ and $S_y$ is the standard deviation of $y$, $n_x$ is the number in set $x$ and $n_y$ is the number in set $y$.

t_{calc} is compared to tabulated values of $t$ at the degrees of freedom $n_x + n_y - 2$ and the chosen degree of certainty usually 95%.

If $t_{calc}$ is greater than or equal $t_{tab}$ in a one tailed $t$-test then the hypothesis that $x>y$ at the chosen degree of certainty is proven. If $t_{calc}$ is less than $t_{tab}$ then there is no significant difference between the two data sets.

In all calculations performed using the students $t$-test the data set with the largest average was taken to be $x$ and the one with the smallest data set taken to be $y$. 

- 135 -
8.18. DESCRIPTION OF PLASMIDS

8.18.1. E. coli plasmids

pPLRT1. A plasmid expressing the 434 cl gene under the control of the \( \lambda_{PL} \) promoter and \( rrnB \) T1 terminator.
pPLRT2. A plasmid expressing the 434\(^{P22} \) repressor under the control of the \( \lambda_{PL} \) promoter and \( rrnB \) T1 terminator.
pPLRT3. A plasmid expressing the 434\(^{Ala44} \) repressor under the control of the \( \lambda_{PL} \) promoter and \( rrnB \) T1 terminator.

8.18.2. Yeast plasmids

pBM150 (Johnston and Davis, 1984). A yeast centromere plasmid carrying the \( CEN4 \) and \( ARS1 \) replicons for yeast and the \( URA3 \) gene for auxotrophic selection. It also carries the \( GAL1 \) and \( GAL10 \) promoters. The plasmid can be replicated in \( E. coli \) and carries the ampicilin resistance gene for selection.
pYcDE-2 (Hadfield et al., 1986). A yeast 2\( \mu \)m plasmid containing the \( TRP1 \) gene for auxotrophic selection and a promoter-terminator cassette consisting of the \( SphI \) fragment of the \( ADH1 \) gene conferring constitutive expression (Beier and Young 1982) and the \( CYC1 \) terminator separated by a unique EcoRI cloning site. The plasmid can be replicated in \( E. coli \) and carries the ampicilin resistance gene for selection.
pCH100 (Hadfield et al., 1986). A yeast 2\( \mu \)m plasmid based on pYcDE-2 expressing the \( E. coli \) chloramphenicol acetyl transferase gene from pBR328 under the control of the \( ADH1/CYC1 \), promoter/terminator cassette.
pCH137 (Hadfeild et al., 1990). A plasmid based on pUC19 containing a clone of the \( PGK \) promoter and terminator from \( Saccharomyces \).
cerevisiae, separated by a unique BgIII cloning site, at the HindIII site of pUC19.

pYRG12 (Hadfield et al., 1986). A plasmid containing the HIS3 gene of Saccharomyces cerevisiae on a BamHI fragment cloned at the BamHI site in pUC19.

pYCW20. A plasmid for the expression of the 434P22 repressor in yeast at single copy integrated into the HIS3 locus. The plasmid was cut with XhoI before transformation.

pYCW28. A plasmid for the expression of the 434P22 repressor in yeast at multicopy on a 2μm plasmid. The plasmid also carries the TRP1 gene for auxotrophic selection in strains carrying a null mutation in the TRP1 gene.

pYCW36. A plasmid for the expression of the 434 repressor in yeast at multicopy on a 2μm plasmid. The plasmid carries the URA3 gene for auxotrophic selection in strains prototrophic due to a null mutation in the URA3 locus.

Reporter Plasmids pYCWWT, pYCW37, pYCW38, pYCW39, pYCW40, pYCW41. These plasmids carry the E. coli β-galactosidase gene under the control of the Saccharomyces cerevisiae PGK promoter and terminator. The PGK promoter has been altered such that it contains 434 operators at various specific locations. These plasmids are for single copy integrated at the URA3 locus and are cut with NcoI prior to transformation to increase the frequency of homologous recombination (Orr-Weaver et al., 1981).

pYCW42. A plasmid designed to express the 434P22 under the control of the Saccharomyces cerevisiae GAL1 promoter. This is based on the yeast centromere plasmid pBM150 which carries the CEN4 and ARS1 replicons for yeast and the URA3 gene for auxotrophic selection. The
plasmid can be replicated in *E. coli* and carries the ampicilin resistance gene for selection.

**8.19. STRAINS**

**8.19.1. Escherichia. coli:**

**DH5α** (Hanahan, 1983); F-, *endA1, hsd(rK-, mK+), supE44, thi-1, λ-, recA1, gyrA96, relA1, Δ(argF-lacZYA)U169, φ80dlacZΔM15.

**XII-blue** (Bullock *et al.*, 1987); *supE44 hsdR17 recA1 endA1 gyrA46 thi relA1 lac*<sup>-</sup>. F′ [proAB<sup>+</sup> lacIq lacZΔM15 Tn10]

**RZ1032** (Kunkel *et al.*, 1987); HfrKL16 PO/45 [lysA(61-62)], *dut1, ung1, thi1, relA1, supE44, Zbd-279::Tn10

**6300ΔlacU169;** *E. coli* 6300 is a F<sup>-</sup>, λ<sup>-</sup>, prototrophic strain obtained from Dr B. Bachmann at the *E. coli* genetic stock centre. The ΔlacU169 mutation was introduced by P1-transduction using linkage to *proC*, from a ΔlacU169, *proC::Tn5* donor strain NK 6991, obtained from Dr. N. Kleckner. The Tn5 transposon was eliminated from a lac<sup>-</sup> transductant by selection for a spontaneous proC<sup>+</sup> derivative. *E. coli* 6300ΔlacU169 was a kind gift from D. Pioli (ICI Pharmaceuticals).

**8.19.2. Yeast:**

*Saccharomyces cerevisiae* S150-2B; *MATα, ura3-52, trp1-289, his3-Δ1, leu2-3, leu2-112.*
CHAPTER 9.

References.
References


9 - References


References


- 144 -


APPENDIX
The DNA sequence and deduced amino acid sequence of the 434 Cl gene encoding the 434 repressor.

ATGAGTATTTCTCCAGGGTTAAAAGCAAAAGATTGAGCTTTGACCTTAACCAGGCTGAA
TACCTAATAGAAAGTGCTCCATTTTTGTTTCTTAAAGCTGAACTTGAGCTTGGACT
MetSerIleSerSerArgValLysSerSerGlnLeuGlyLeuAsnGlnAlaGlu 20

CTTGCTCAAAGTTGGTACCTACCCACGACATCTATAGACGCTCGAAACGGTTAAA
GAACGAGTTTTCCACCATGTCGAGCTTGAGTTGGGGTTTCAAC
LeuAlaGlnLysValGlyThrThrGlnGlnSerIleGluGluAsnGlyLysThr 40

AAAGCGGCCGCGTTTTTATTACGAAACTTTGGCTAGCTTTGGGCTAAGCTGACGT
TTGGCGCGCGAAAAATGGTCTTGAACGCAGTCGAGAACCGCATTCACAACTGAC
LysArgProArgPheLeuProGluLeuAlaSerAlaLeuGlyValSerValAspTrpLeu 60

CTCAATGGCACCTCTGATTGCTAACATTTGGGGCTGAGCCCAAAGGGAAA
GAGTTACCCTGAGACTAAGCTTACAATCTAACAACCCGGACTCGGAGCTTTCC
LeuAsnGlyThrSerAspSerAsnValArgAlaGluHisValGluProGlyLys 80

TATCCATTGATTTGTTAGCTGGTGTTCTGTTGCTGAAGCTGTAGCTCCTAGAT
ATAGGTAACATAATCTGGACAAAAGCTCAGCACCACACTTCGAGAATTGGAGATTGCA
TyrProLeuIleSerValArgAlaGlySerTrpCysGluAlaCysGluProTyrAsp 100

ATCGAGGCACTGATGATGGTACGAGCTAGTCTAGGACATGGATTCTGG
TAGTTCTGTAACACTTACACATACGTCACTGCAATTGAAATACGTTACCTAAGACC
IleLysAspIleAspGluTrpTyrAspSerAspValAsnLeuLeuGlyAsnGlyPheTrp 120

CTGAAGGTTGAAGGTGATTCCATGACCTCACCTGTAGGTCAAAGCATCCCTGAAAGTCTC
GACTTCACCTCACAAGCTGGAGTGGACCAGATCAGTTTCTGGTACCACTTTGGAGACT
LeuLysValGluGlyAspSerMetThrSerProValGlyGlnSerIleProGlyGlyHis 140

ATGGGTCTTCTAGTGAACGTGGACCGAGCTGAAATGGAGACCTTTGTTGAGCCTCGAAGT
TACCACAATCATCTAGACCTCGTCCGTCGTCATCTTTTCTCGAAACACATCGTGTGAC
MetValLeuValAspThrArgGluProValAsnGlySerLeuValValAlaLysLeu 160

ACTGACCGGACAGCAACATCCAAAGAATGCTGACTGAGGCTAGAATCTCCTG
TGACCTCGGCTTCTGGTTAAGTTCTTCTGGGACCAGATCTACCCGGACGTCTTCTAGGAC
ThrAspAlaAsnGluAlaThrPheLysLeuValIleAspGlyGlnLysTyrLeu 180
AAAGGCCTGAATCTTCAAGGCCCTATGACTCCTATCAACGGGAACACTGCAAGATTATCGGTTTTCCGGACTTAGGAAGTACCGGATACTGAGGATAGTTGCCCTTGACGTTCTAATAGCCA

LysGlyLeuAsnProSerTrpProMetThrProIleAsnGlyAsnCysLysIleIleGly 200

GTTGTCGTGGAAGCGAGGGTAAAATTCGTATGA

CAACAGCACCTCGCTCCCATTTTAAGCATACT

ValValValGluAlaArgValLysPheVal 210

The numbers following the amino acid sequence refer to the position in the sequence of the amino acid at the end of the line. It should be realised that the mature protein lacks the methionine at the start of the protein due to post-translational modification.
Program for the Processing of Data from a Microtitre Plate Reader for β-Galactosidase Assays.

5 CLEAR
10 CLS
20 INPUT "ENTER DATE ", WHEN$  
30 LPRINT WHEN$
40 INPUT "ENTER ANY COMMENTS FOR THIS DATA SET, <RETURN>. FOR NOTHING ", TEXT$  
50 LPRINT TEXT$
60 OPEN "COM1:1200,N,8,L,CSO,DSO,BIN" AS 1 LEN=28
70 DIM IN$(12)
80 DIM ABSORB$(12,8)
90 INPUT "ENTER NUMBER OF PLATES TO BE READ; ", NP%
100 DIM PLATE(NP%)
110 FOR A%=1 TO NP%
120 PRINT "ENTER THE NUMBER OF COLUMNS TO BE USED ON PLATE ";A%;
130 INPUT; PLATE(A%)
140 PRINT
150 NEXT
160 LET READINGS%=0
170 FOR X%=1 TO NR%
180 LET READINGS%=READINGS%+PLATE(X%)
190 NEXT
200 LET SAMPLES%=READINGS%/3
210 INPUT "ENTER THE NUMBER OF TIME POINT READINGS PER PLATE; ", TP%
220 DIM VALUE(TP%,8,READINGS%)
230 LET ZZ%=TP%+1
240 DIM ARRAY(ZZ%,8,READINGS%)
250 CLS
260 LET T%=TP%+1
270 LET PR$="550"
280 GOTO 330
290 LET PR$="420"
300 FOR T%=1 TO TP%
310 CLS
320 LET SSS%=0
330 FOR A%=1 TO NP%
340 GOSUB 1410
350 CLS
360 FOR R%=1 TO 8
370 LET FACT%=SSS%
380 FOR C%=2 TO (PLATE(A%)+1)
390 LET COLM%=C%+FACT%-1
400 LET ARRAY(T%,R%,COLM%)=VAL(ABSORB$(C%,R%))
410 NEXT
420 NEXT
430 LET SSS%=SSS%+PLATE(A%)
440 NEXT
450 IF PR%="550" THEN GOSUB 290
460 NEXT
470 REM THIS CALCULATES THE ACTUAL ABSORBANCE READINGS
480 FOR R%=1 TO 8
490 FOR C%=1 TO SAMPLE%
500 IF C%>1 THEN LET E%=C%+(C%-1)*2 ELSE LET E%=C%
510 FOR P%=1 TO TP%
520 LET D%=0
530 LET NO%=1
540 LET D%=D%+E%
550 LET DC=ARRAY(P%,R%,D%)
560 IF DC<=1 THEN GOTO 600
570 LET D%=D%+1
580 LET NO%=NO%+1
590 GOTO 550
600 LET OD420=ARRAY(P%,R%,D%)
610 LET OD550=ARRAY(ZZ%,R%,D%)
620 IF NO%=1 THEN LET XXX=.1
630 IF NO%=2 THEN LET XXX=.01
640 IF NO%=3 THEN LET XXX=.001
650 LET VALUE(P%,R%,C%)=((OD550*1.3)*1000)/(.36*XXX)
660 NEXT
670 LET NO%=1
680 NEXT
690 LET NO%=1
DIM TIME(TP%)
INPUT "ADVANCE PRINTER TO THE TOP OF A NEW SHEET. PRESS <RETURN> TO CONTINUE", WELLS
FOR T%=1 TO TP%
PRINT "ENTER TIME OF READING £"; T%;
INPUT TIME(T%)
PRINT
NEXT
DIM LR(2,8,READINGS%)
LPRINT "BETA-GALACTOSIDASE ASSAYS FROM ";WHENS
FOR C%=1 TO SAMPLES%
FOR R%=1 TO 8
LET SX=0
LET SY=0
LET SXS=0
LET SYS=0
LET SXY=0
FOR T%=1 TO TP%
LET SX=SX+TIME(T%)
LET SY=SY+VALUE(T%,R%,C%)
LET SXS=SXS+TIME(T%)^2
LET SYS=SYS+VALUE(T%,R%,C%)^2
LET SXY=SXY+(TIME(T%)*VALUE(T%,R%,C%))
NEXT
LET LR(2,R%,C%)=((TP%*SXY)-(SX*SY))/(((TP%*SXS)-(SY^2))^0.5)
LET DX=TIME(TP%)-TIME(1)
LET Y1=(((SX*SY)/TP%)-SXY)/((SX^2/TP%)-SXS)*TIME(1)+
((SY/TP%)-(((SX*SY)/TP%)-SXY)/((SX^2/TP%)-SXS))*(SX/TP%))
LET Y2=(((SX*SY)/TP%)-SXY)/((SX^2/TP%)-SXS)*TIME(TP%)+
((SY/TP%)-(((SX*SY)/TP%)-SXY)/((SX^2/TP%)-SXS))*(SX/TP%))
LET DY=Y2-Y1
LET LR(1,R%,C%)=DY/DX
PRINT "ROW; "; R%; " COLUMN; " ;C%; " r=";LR(2,R%,C%)," RATE=";LR(1,R%,C%)
LPRINT "ROW; "; R%; " COLUMN; " ;C%; " r=";LR(2,R%,C%)," RATE=";LR(1,R%,C%)
NEXT
1030 LPRINT
1040 NEXT
1050 CLS
1060 INPUT "ENTER THE NUMBER OF SAMPLE PER FLASK"; FLASK%
1070 DIM RL(FLASK%) 
1080 FOR A%=1 TO 8
1090 LET SX=0
1100 LET SY=0
1110 LET SXS=0
1120 LET SYS=0
1130 LET SXY=0
1140 FOR B%=1 TO FLASK
1150 PRINT "WHICH COLUMN FOR SAMPLE "; B%;
1160 INPUT COL%
1170 PRINT "ENTER OD600 FOR COLUMN "; COL%; " ROW "; A%;
1180 INPUT RL(B%) 
1190 LET SX=SX+RL(B%) 
1200 LET SY=SY+RL(1,A%,COL%) 
1210 LET SXS=SXS+RL(B%)^2 
1220 LET SYS=SYS+RL(1,A%,COL%)^2 
1230 LET SXY=SXY+(RL(B%)*(1,A%,COL%)) 
1240 NEXT
1250 LET CC=((FLASK%*SXY)-(SX*SY))/(((FLASK%*SXS)-(SX^2)) *((FLASK%SYS)-(SY^2)))^.5 
1260 LET DX=RL(FLASK%)-RL(1) 
1270 LET Y1=((((SX*SY)/FLASK%)-SXY)/((SX^2/FLASK%)-SXS)) *RL(1)+((SY/FLASK%)-(((SX*SY)/FLASK%)-SXY)/((SX^2/FLASK%)-SXS)) 
1280 LET Y2=((((SX*SY)/FLASK%)-SXY)/((SX^2/FLASK%)-SXS)) *RL(FLASK%)+((SY/FLASK%)-(((SX*SY)/FLASK%)-SXY)/((SX^2/FLASK%)-SXS))*(SX/FLASK%)) 
1290 LET DY=Y2-Y1 
1300 LET RATE=DY/DX 
1310 PRINT "FOR FLASK ";A%+"SET," RATE=";RATE; "
1310 UNITS/ML/MIN/OD600. ", "r="; CC 
1310 LPRINT "FOR FLASK ";A%+"SET," RATE=";RATE; "
1310 UNITS/ML/MIN/OD600. ", "r="; CC 
1330 NEXT
CLS
PRINT "IS THERE ANY MORE DATA TO CALCULATE (Y/N)?";
INPUT ANSWERS
IF ANSWERS="Y" THEN LET SET=SET+8
IF ANSWERS="y" THEN LET SET=SET+8
IF SET>=8 THEN GOTO 1080
END
IF PRS="420" THEN GOTO 1440
PRINT "OD";PRS;" PLATE NO. ";A% :PRINT
LPRINT "OD";PRS;" PLATE NO. ";A% :GOTO 1460
PRINT "OD";PRS;" PLATE NO. ";A%;" TIME POINT NO. ";T% :PRINT
LPRINT "OD";PRS;" PLATE NO. ";A%;" TIME POINT NO. ";T% :PRINT
LOCATE 11,24:PRINT "READY TO ACCEPT DATA FROM MULTISCAN"
I%=1:INPUT I,INS(I%):IF LEN(INS(I%))<>75 THEN 1460 ELSE 1480
FOR I%=2 TO 8
INPUT I, INS(I%)
NEXT
LOCATE 2,1
FOR I%=1 TO 8
PRINT INS(I%):LPRINT INS(I%)
NEXT
CLS
LOCATE 11,24:PRINT " ** DATA RECEIVED **"
FOR I%=1 TO 8
FOR J%=1 TO 12
PART=((J%-1)*6)+4
ABSORS$ (J%,I%)=RIGHTS (ABSORS$ (J%,I%),5)
NEXT
NEXT
RETURN
END