Purification and Characterisation of the Repressor Proteins from the *Streptomyces* Bacteriophage ØC31

A thesis submitted in accordance with the requirements for the degree of Doctor of Philosophy

August 1993

by

Charles Edward Owen

Institute of Genetics
University of Glasgow

ProQuest Number: 11007694

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

ProQuest 11007694

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.
This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

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

GLASGOW
UNIVERSITY
LIBRARY

Thesis
9593
Copy 1

Abstract

ØC31 is a bacteriophage of *Streptomyces* species in which it undergoes a lysogenic life-cycle. The phage possesses a 41.5kb genome and cohesive ends. The lysogenic phase of development is regulated by the repressor gene (*c*) which lies in the centre of the phage genome (Lomovskaya *et al*, 1972). The *c* gene was cloned on a 3.4kb *Sph*I-G fragment; it was sequenced and mapped at the transcriptional level (Sinclair and Bibb, 1988; Sinclair and Bibb, 1989). Further low resolution S1 mapping identified that transcription of the repressor gene consisted of 3 nested N-terminally different, C-terminally identical, in-frame mRNA's (Smith and Owen, appendix). Primer extension analysis mapped the 5' ends of the transcripts; upstream of each was found a good consensus -10 promoter sequence and identically conserved ribosome binding sites situated 7-10bp upstream of the ATG initiation codon. The three transcripts each possessed an ORF predictedto produce proteins of 74, 54 and 42kDa (Smith and Owen, appendix).

The repressor gene was expressed on a high copy number plasmid in *E.coli* and proteins unique to the *c* gene were analysed. Initial work showed that in *E.coli,* a promoter located upstream of the *c* gene, *c*P1 was functional, but that promoters internal to the *c* gene, *c*P2 and *c*P3 did not. Three overexpressed proteins of 110, 70 and 50kDa were observed on SDS-PAGE gels of crude lysates of *E.coli* carrying the repressor gene. The introduction of frameshift mutations at the unique *Nco*I restriction site revealed that the 74kDa repressor actually migrated on SDS-PAGE with the mobility of a protein of 110kDa, and that the observed 70 and 50kDa repressor proteins were actually the predicted 54 and 42kDa proteins.

The 74,54 and 42kDa repressor proteins were purified away from *E.coli* proteins. This enabled N-terminal protein sequencing of the 54 and 42kDa proteins which confirmed the identity of the predicted ORF assigned for each protein. A plasmid, pMS221, encoding only the 42kDa protein was used to isolate the 42kDa protein in isolation. Ammonium sulphate precipitation followed by anion-exchange and heparin-agarose chromatography resulted in 95% pure protein. Gel filtration and chemical crosslinking identified that the pure 42kDa repressor most probably exists as a tetramer.

DNA binding studies using the bandshift assay demonstrated that the 42kDa repressor binds to all three putative promoters in the repressor gene itself with

varying degrees of strength. It appears to bind the *c*P2 promoter with the greatest affinity followed by the *c*P1 promoter and finally *c*P3. The *c*P1 and *c*P2 promoters contain a 20bp repeated sequence and is a putative operator sequence. DNaseI footprinting of the 42kDa protein binding to the *c*P1 promoter was attempted. The 42kDa repressor also binds to a site within the early region of the phage. This site possesses two putative operator sites interspaced by a terminator-like sequence. Each half of this site, containing only a single 20mer motif was shown to bind the 42kDa protein with relatively low affinity. However, when the entire site was used DNA-binding was found to be very strong and comparable to that of the *c*P2 promoter.

The work presented in this thesis is my own unless otherwise acknowledged. No part of this thesis has been previously submitted for examination leading to the award of a degree.

Charles E Owen
August 1993

This thesis is dedicated to my family, particularly my Parents and Grandparents
for their support and encouragement.

Acknowledgments

Like many theses, this has been the product of 3 years of laughter, tears and sheer bloodymindedness. I would like to thank Maggie Smith for attentively supervising the project and Iain Hunter for allowing me to "abuse" his lab. and providing sound advice.

I am indebted to all members of the lab. and the department, past and present, particularly Paul, Phil, Frances and Richard for their humour and company, making life in Glasgow an enjoyable time. Special thanks to Sohail, Martin, Colin and Marshall for allowing me to bug them and the media ladies for 'debugging' my glassware.

Table of Contents

Section

Title i
Abstract ii
Declaration iv
Dedication v
Acknowledgements vi
Table of Contents vii
Abbreviations xi

Chapter 1- Introduction

1.1 Scope........ 1
1.2. DNA-Protein Interactions........ 1
1.2.1 Structural Conformation of DNA........ 1
1.2.2 Helix-Turn-Helix Motif........ 2
1.3 λ Bacteriophage - A model for gene control........ 4
1.3.1. Action of cII and cIII Regulatory Proteins........ 4
1.3.2. λ Repressor and Cro........ 6
1.3.3. Role of *E.coli* Encoded Proteins in Regulating λ Development........ 8
1.4. Interesting Features in Other Bacteriophages........ 9
1.4.1. P22 Phage and the Anti-repressor........ 9
1.4.2. Bacteriophage P1........ 11
1.5 Streptomyces- Hosts of ØC31........ 12
1.6 Regulation of Gene Expression in *Streptomyces*........ 13
1.6.1 RNA Polymerase Heterogeneity in *Streptomyces coelicolor*........ 15
1.7 The Actinophage ØC31........ 17
1.7.1. Organization of the ØC31 Genome........ 20
1.7.2 The ØC31 Repressor Locus 21
1.8 Aim of Experimental Work 28

Chapter 2- Methods and Materials

2.1 Bacterial Strains........ 29
2.2 Plasmids........ 29
2.3 Synthetic Oligonucleotides........ 29
2.4 Chemicals........ 31
2.5 Buffer Solutions........ 31
2.6 Microbiological Methods........ 33
2.6.1 Culture Media........ 33
2.6.2 Sterilization........ 33
2.6.3 Selection Supplements........ 33
2.6.4 Bacterial Growth Conditions........ 33
2.6.5 *E.coli* Strain Preservation........ 34
2.6.6 Introduction of Plasmid DNA into *E.coli*........ 34
2.7 Manipulation of DNA........ 34
2.7.1 Large Scale Plasmid Isolation (mod.Birnboim and Doly1976)........ 34
2.7.2 Small-scale Plasmid Isolation........ 35
2.7.3 Precipitation of DNA........ 35
2.7.4 Digestion of DNA with Restriction Endonucleases........ 35

2.7.8.3 Native Polyacrylamide Electrophoresis of DNA 37
2.7.8.4 Denaturing Polyacrylamide Sequencing Gel Electrophoresis............ 38
2.7.9 Photography of Ethidium Stained Gels 38
2.7.10 Recovery of DNA from Agarose Gels 39
2.7.10.1 Isolation of Small DNA Fragments from LMP Agarose................ 39
2.7.10.2 Isolation of DNA from Agarose Gels Using Electroelution............ 39
2.7.11 Elution of DNA from Polyacrylamide Gels 40
2.7.12 Quantification of Nucleic Acid 40
2.7.13 Plasmid DNA Sequencing 40
2.7.14 Oligonucleotides 41
2.7.14.1 Oligonucleotide Synthesis 41
2.7.14.2 Deprotection of Oligonucleotides 41
2.8. Preparation of Radio-labelled DNA Probes 41
2.8.1. 5'-End-labelling of Oligonucleotides 41
2.8.2. 5' End-labelling of DNA Fragments 42
2.8.2 Sephadex G50 Column Chromatography 42
2.8.3 Polymerase Chain Reaction to Produce End-Labelled Probes from the Repressor Gene 42
2.8.3.1 PCR Protocol for Repressor Promoter Fragments
2.8.3.2 Oligonucleotides Used for PCR End-Labelling 43
2.8.4. Preparation of Early Gene DNA Fragments by PCR............................ 43
2.8.4.1 PCR Protocol for Preparation of Early Gene Promoter Fragments 43
2.8.4.2 Oligonucleotides used for Production of PCR Products of the Early Gene Region 44
2.8.5 Clean Up and Analysis of PCR Products 44
2.8.6 Autoradiography 45
2.9 Methods Used in the Analysis and Purification of the Repressor Proteins 45
2.9.1 Chromatography Media 45
2.9.2 Protein Separation using Liquid Chromatography 45
2.9.3 Quantification of Protein 45
2.9.4 Electrophoresis of Protein Gels 45
2.9.4.1 SDS-Polyacrylamide Gel Electrophoresis 45
2.9.4.2 Native Continuous Polyacrylamide Gel Electrophoresis.................... 45
2.9.5 Visualization of Protein 46
2.9.5.1 Coomassie Staining 46
2.9.5.2 Silver Staining 46
2.9.6 Drying Polyacrylamide Gels 46
2.9.7 Molecular Weight Markers 46
2.9.8 Cross-linking of Protein Subunits Using Dimethylsuberimidate 46
2.10. Production of Anti-Repressor Anti-serum 47
2.11. Western Blotting 47
2.11.1 Transfer of Proteins to a Membrane Support 47
2.11.2 Incubation of Western Blot with Antibody 48
2.12. N-terminal Protein Sequence Analysis 48
2.12.1 Preparation of the Protein Sample for Sequencing 48
2.12.1.1 Electrophoresis of the Protein Sample 49
2.12.1.2. Electroblotting of Protein Samples onto Pro-BlottTM Membrane 49
2.13 Purification of the Repressor Proteins 50
2.13.1 Construction and Maintenance of Chromatographic Columns 50
2.13.2 Buffers Used During Purification of the Repressor Proteins 50
2.13.3 Purification of the Putative 74, 54 and 42kDa Repressors 50
2.13.4 Purification of the 42kDa Repressor Protein 51
2.14 Storage of the Purified Protein 52
2.15 Assays for DNA Binding Activity 52
2.15.1 Band-shift Assay 52
2.15.2 DNaseI Footprint Analysis of DNA Protein Complexes 53

2.13.1 Construction and Maintenance of Chromatographic Columns 50
2.13.2 Buffers Used During Purification of the Repressor Proteins........ 50
2.13.3 Purification of the Putative 74, 54 and 42kDa Repressors 50
2.13.4 Purification of the 42kDa Repressor Protein 51
2.14 Storage of the Purified Protein 52
2.15 Assays for DNA Binding Activity........ 52
2.15.1 Band-shift Assay 52
2.15.2 DNaseI Footprint Analysis of DNA Protein Complexes........ 53
2.15.3 Maxam and Gilbert Chemical Sequence Analysis 53

Chapter 3- Overexpression of the Repressor Proteins from ØC31in *E.coli*

3.1 Introduction........ 54
3.2 Overexpression of the Repressor Proteins from ØC31 in *E coli*........ 54
3.3 Introduction of Frameshift Mutations into the Repressor Gene to Identify Repressor Proteins........ 54
3.4 Expression of Repressor Proteins From Plasmids Containing Alternative Subclones of the *c* Gene........ 57
3.6 Discussion and Conclusions........ 60

Chapter 4- Development of a Purification Strategy for the ØC31 Repressor Proteins

4.1 Introduction........ 66
4.2 Theory of Separation Methods Used in the Purification of the ØC31 Repressor Proteins........ 66
4.2.1 Salt Fractionation........ 66
4.2.2 Ion-Exchange Chromatography 67
4.2.3 Affinity Chromatography........ 67
4.3 The Purification of the 74, 54 and 42kDa ØC31 Repressor Proteins 68
4.3.1 Procedure........ 68
4.3.2 Salt Fractionation of the ØC31 Repressor Proteins........ 68
4.3.3 Separation of the ØC31 Repressor Proteins From Total Protein using Heparin-agarose Chromatography........ 68
4.4 NH_2 Terminal Amino Acid Sequence Analysis........ 73
4.5 Conclusions and Discussion........ 73

Chapter 5- Purification and Biochemical Characterization of the ØC31 42kDa Repressor

5.1 Introduction........ 78
5.2. Purification Method for Isolating the 42kDa Repressor Protein........ 78
5.2.1 Purification of the 42kDa Repressor By Anion-Exchange Chromatography........ 78
5.2.2 Fractionation of the 42kDa Repressor Using Heparin-Agarose Affinity Chromatography........ 82
5.3. Purification Summary for the 42kDa Repressor Protein........ 87
5.4 Biochemical Characterisation of Pure 42kDa Repressor........ 87
5.4.1 Superose 6 Determination of the Oligomeric Molecular Weight of the Monomeric 42kDa Repressor Protein 92
5.5. Chemical Cross-linking of the 42kDa Repressor Using Dimethylsuberimidate........ 95
5.6. Discussion and Conclusions 95

Chapter 6- Characterization of DNA Binding by the ØC31 42kDa Repressor Protein

6.1 Introduction.......... 99
6.2. *c*P1 Promoter.......... 99
6.2.1. Binding of the 42kDa Repressor Protein to the *c*P1 Promoter.......... 102
6.3. *c*P2 Promoter.......... 102
6.3.1. Binding of the 42kDa Repressor to the *c*P2 Promoter Fragment.......... 102
6.4. *c*P3 Promoter.......... 103
6.4.1. Binding of the 42kDa Repressor to the *c*P3 Promoter Fragment.......... 103
6.5 Calculation of the Dissociation Constant for the 42kDa Repressor With Promoter Fragments Isolated From the Repressor Gene.......... 107
6.6. Specificity of Binding by the 42kDa Repressor Protein to *c*P1 and *c*P2.......... 107
6.6.1. Specificity of Binding of $cP1^{H3}$ with the 42kDa Repressor.......... 107
6.6.2. Specificty of Binding of the 42kDa Repressor to the $cP1^{Nhe}$ Promoter Fragment.......... 107
6.7. DNaseI Footprinting of the cP1Nhe Fragment With the 42kDa Repressor Protein.......... 108
6.8. The 42kDa Repressor Binds to a Site in the Early Region.......... 108
6.8.1. The 42kDa Repressor Binds to the CIFC Early Region.......... 108
6.8.2. Binding of the 42kDa Repressor with Each Half of the CIFC Early DNA Fragment.......... 109
6.9. Calculation of the Dissociation Constant for the CIFC, CIGC and CHFC DNA Fragments.......... 109
6.10. Competition for the 42kDa Repressor Binding to the CIFC Early Region.......... 114
6.11. Conclusions and Discussion.......... 114

Chapter 7- Conclusions and Further Work 122

Chapter 8- Bibliography 126

Appendix

Abbreviations

A	- Absorbance (1cm path length)
Amp.	- Ampicillin
bp	- base pairs
BSA	- Bovine serum albumin
CIP	- Calf Intestinal Alkaline Phosphatase
CIR	- Conserved Inverted-Repeat
cpm.	- counts per minute
DMSO	- Dimethylsulphoxide
DNA	- Deoxyribonucleic acid
ds	- Double-stranded
DTT	- Dithiothreitol
EtBr	- Ethidium bromide
IPTG	- Isopropyl-β-D-thiogalactopyranoside
kb	- kilobase
kDa	- kilodalton
lac	- Lactose
M.Wt	- Molecular weight
nt	- Nucleotide (s)
ORF	- Open reading frame
PAGE	- Polyacrylamide gel electrophoresis
PCR	- Polymerase chain reaction
PEG	- Polyethyleneglycol
PMSF	- Phenylmethylsulfonyl fluoride
RBS	- Ribosome binding site
RNA	- Ribonucleic acid
SDS	- Sodium dodecyl sulphate
XGal	- 5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside

Chapter 1

Introduction

1.1 Scope

Whilst the depth of knowledge of the regulation of gene expression in bacteriophage λ is now quite detailed (Section 1.3.0), much less is known concerning the molecular biology of bacteriophages of the Gram (+ve) bacteria, particularly the actinomycetes. Actinophage ØC31 is the most studied of the phages which infect *Streptomyces* species, and is targeted for study in a similar way to studies of phage λ in *E.coli.* This thesis describes some work directed at understanding gene regulation in ØC31, in particular the interaction of the lytic repressor with its DNA target sites.

The paradigms for studies on gene regulation are the lactose utilization genes and the lysis versus lysogeny decision in phage λ (Jacob and Monod, 1961; Ptashne, 1986). Ptashne and co-workers have developed a detailed understanding of the molecular mechanisms of gene regulation from their studies on λ. These studies have been important to our understanding of other regulatory systems in both prokaryotes and eukaryotes. In this 'Introduction' I will discuss our current understanding of gene regulation in *Streptomyces* and ØC31 and place this knowledge in context of well-studied regulatory systems.

1.2. DNA-Protein Interactions

The molecular basis of sequence-specific DNA binding by a number of proteins has been studied. Some of these will be considered below. For more comprehensive reviews of DNA binding by proteins, structural motifs and the implications of DNA:protein interactions, a number of excellent reviews should be consulted including: Schleif, 1988; Brennan and Matthews, 1989a; Brennan and Matthews, 1989b; Pabo and Sauer, 1984; Harrison and Aggrawal, 1991; Magasanik, 1989; Struhl, 1989; Berg, 1990; Harrison, 1991 .

1.2.1 Structural Conformation of DNA

Most DNA in the cell exists in the B form, in which the base pairs are displaced from the helical axis. This asymmetry creates a narrow minor groove and a wider major groove.

DNA should be considered a dynamic molecule capable of adopting different helical conformations, not all of which are regular. For a protein to bind DNA, the DNA may have to alter conformation to optimize the number of favourable contacts, so promoting the formation of essential charge interactions. Each unit of the phosphate-sugar-base structure of DNA possesses six bonds about which rotation is possible. The rotations are; helical twist from one nucleotide to the next; tilt of an individual base pair around an axis within the plane of the hydrogen bonds and perpendicular to the hydrogen bonds between the bases; propellor twist of one base with respect to another, and buckle of the bases out of one plane after removal of the propellor pitch. All these rotations vary from one base pair to the next along the DNA (Diekman, 1988). A DNA binding protein, in theory at least, may be able to bind specifically to a target site based solely on the conformation of the DNA and without regard for the primary sequence, so-called 'indirect read-out'. If the specificity of binding is directly the result of the base sequence, this information is termed 'direct read-out'. The majority of prokaryotic DNA-binding proteins recognize binding sites on DNA via 'direct read-out'. The first protein structure solved with this capability was the helix-turn-helix motif found in the λ Cro protein (Anderson *et al,* 1981). This structure has since been found in other prokaryotic DNA-binding proteins and is possibly the most commonly used protein:DNA recognition motif within the prokaryotes. A structure similar to the helix-turn-helix motif is predicted to exist within the ØC31 repressor proteins (Sinclair and Bibb, 1988).

1.2.2 Helix-Turn-Helix Motif

The structures of the first DNA binding motif elucidated conclusively were the helix-turn-helix structure of the λ phage Cro protein(Anderson *et al,* 1981) and the *E.coli* catabolite gene activator protein (CAP) (McKay and Steitz, 1981). The helix-turn -helix motif (HTH) consists of two helices between which is a tight ß-turn, resulting in an inter-helical angle of nearly 90°. The helices are spaced 34Å apart (one turn of duplex B form DNA approximates to 34Å), which come together as a dimeric pair of helices. The HTH motif from each dimer makes contact with the same face of DNA within two successive major grooves of two-fold symmetrical DNA (Anderson *et al,* 1981; Ohlendorf *et al,* 1982; Sauer *et al,* 1982; Steitz *et al,* 1982) . An alpha helix has a diameter of approximately 12Å, whilst B-form DNA is about 12Å wide and between 6-8Å deep. Thus, an alpha helix can fit almost perfectly within the major groove of B DNA (Pabo and Sauer, 1984).

Conserved amino acid sequences have been identified in both helices to form what is now called the helix-turn-helix motif (Figure 1.1). The HTH motif conventionally consists of 20 amino acids; the first 7 forming the first helix and residues 12-20 forming

Figure 1.1. Comparison of Helix-Turn-Helix Motifs (Brandon and Tooze, 1990)

|<-----Helix-2--------->|<--Turn->|<---------Helix-3---------->|

Residue:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
434 Rep.	Q	A	D	L	A	Q	K	V	G	T	T	Q	Q	S	I	D	Q	L	D	N
434 Cro	Q	T	D	L	A	T	K	A	G	V	K	Q	Q	S	I	Q	L	I	D	A
Lam. Rep	Q	D	S	V	A	N	K	M	G	M	G	Q	S	G	V	G	A	L	Y	N
Lam. Cro	Q	T	K	T	A	K	E	L	G	V	Y	Q	S	A	I	N	K	A	I	H
Lam. cII	T	E	K	T	A	E	A	V	G	V	D	K	S	Q	I	S	R	W	K	R
CAP	R	Q	E	I	G	Q	I	V	G	C	S	R	E	T	V	G	R	I	L	K
Trp Rep.	Q	R	D	L	K	N	D	L	G	A	G	I	A	T	I	T	R	G	S	N
Lac Rep.	L	Y	E	V	A	R	L	A	G	V	S	Y	Q	T	V	S	R	V	V	N
Cyt Rep.	I	K	D	V	A	L	K	A	K	V	S	T	A	T	V	S	R	A	L	M
Lys Rep.	L	T	E	A	A	H	L	L	H	T	S	Q	P	T	V	S	R	E	L	A
ØC31 Rep.	A	D	K	T	A	E	A	L	K	V	I	R	A	Q	V	E	K	A	G	K
Consensus:	X	X	Ac	Hy	A	X	X	Hy	G	Hy	X	Ba	X	T	Hy	X	Ba	X	X	X

Key

Ac- Acidic

Ba- Basic

Hy- Hydrophobic

X- No Consensus

the second. Residues 4, 8, 10, 16 and 18 are usually hydrophobic and define the interior of the two-helix elbow, but require hydrophobic residues from elsewhere to complete the hydrophobic core. Residue 5 is usually a glycine or alanine, although other residues have been found at this position. Residues with large side chains at position 5, such as the lysine residue in the *trp* repressor perturb the geometry of the HTH as they interfere with the peptide backbone at the N-terminus of the second helix. This has been demonstrated crystallographically for the *trp* repressor (Schevitz *et al,* 1985). Residue 9 in the HTH is frequently a glycine residue and the geometry of the turn reportedly relies on the presence of a glycine at this position. However, a number of putative HTH motifs have been reported with alternative residues at this position, including replacement by serine, cysteine, glutamic acid, histidine and lysine (Pabo and Sauer, 1984; Laughon and Scott, 1984; Hochschild, Irwin and Ptashne, 1983; Stragier and Patte, 1983; Sanger *et al,* 1982; Ohtsubo *et al,* 1981) as does the λ B58 protein (Sanger *et al,* 1982) and the *Shigella* isoIS1 *ins*A protein (Ohtsubo *et al,* 1981).

1.3 λ Bacteriophage - A model for gene control

Regulation of lytic growth by bacteriophage λ provides a model for gene control governed by a number of DNA-binding proteins and other proteins which influence gene expression. The probability of entering the lysogenic state or the lytic mode is ultimately governed by environmental stimuli.

The colon is a natural habitat of *E.coli* providing a rich source of available nutrients. Such an environment is likely to stimulate lytic growth and the rapid enlargement of the phage population. On less energy-rich growth substrates, the lysogenic mode of growth provides the phage with a state in which it can maintain itself indefinitely (Kourilsky, 1973). The frequency of lysogeny is also influenced by the multiplicity of infection; low multiplicity of infections lead to sub-optimal levels of cII and cIII proteins, which are required to initiate the lysogenization process (Kourilsky, 1974; Herskowitz and Hagan, 1980). The decision between lytic development and lysogeny is determined by a set of three proteins; namely Cro, CI, and CII (Ptashne, 1986).

1.3.1. Action of cII and cIII Regulatory Proteins

The level of the cII protein in the infected cell plays a pivotal role in determining the lytic-lysogenic developmental decision. Protein CII promotes lysogeny by stimulating the synthesis of the λ phage repressor (CI) and the integrase protein for the integration of the phage DNA and successful lysogenization. The CII protein is unstable and must be continuously synthesized for transcription of *c*I from P_{RE} (Promoter for Repressor

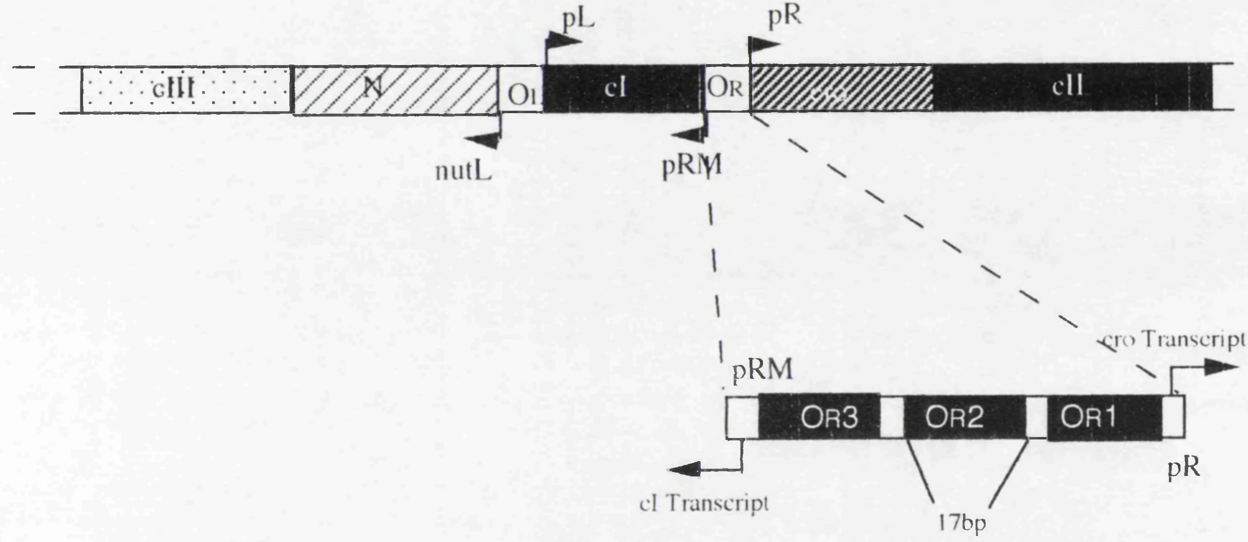


Figure 1.2. Organization of Gene Expression in Coliphage λ

Establishment) to occur. The *E.coli* protein HflA acts as a protease on cII to reduce the level of cII activity. To combat this, the phage produces cIII protein which protects cII from proteolytic attack by HflA, and so increases the half-life of the cII protein (Hoyt *et al,* 1982). Binding of cII to P_{RE} and P_I facilitates RNA polymerase binding and enables transcription from the P_{RE} and P_I promoters to allow lysogeny via the synthesis of the cI repressor and the integrase protein (Wulff and Rosenberg, 1983; Herskowitz and Hagen, 1980). Thus when the concentration of cII protein is high, cI and integrase are in excess and lysogeny occurs. If cII concentrations are low, then Cro, Q and Xis are synthesized in relative excess and the lytic pathway is chosen. The λ Q gene product is negatively regulated by the activation of promoter Pa_Q (Wulff and Rosenberg, 1983; Herskowitz and Hage, 1980). This promoter is considered to produce an anti-sense RNA which is responsible for reducing the expression of the Q gene product (Hoopes and McClure, 1985; Ho, Wulff and Rosenberg, 1986). Q protein is required for the expression of all the 'late' genes.

The target site used by cII is unusual consisting of a 25bp region centred upon the sequence $TTGCN_6TTGT$. This site is highly degenerate with only the 4bp repeat being common to all three promoters, even though cII binds to each promoter with almost identical affinity (Ho and Rosenberg, 1988).

The cII protein itself is a tetramer of 44kDa with a pI of 8.8 (Ho *et al,*1982). Near the N-terminus of the protein is a helix-turn-helix motif which is the DNA binding domain of the protein. This has been confirmed by mutational analysis, in which amino acid substitution was used to disrupt the motif (Ho *et al,* 1986). The cII protein is processed by removal of the first two residues of the protein, f-Methionine and Valine, leaving an Arginine as the first amino acid to form the fully mature protein (Ho *et al,* 1986).

1.3.2. λ Repressor and Cro

λ repressor and Cro share the same DNA recognition sites, but their effects are mutually antagonistic. λ repressor controls the entry and the maintenance of the lysogenic state of the prophage, whereas Cro is a cI antagonist and facilitates entry into the lytic cycle. The decision of whichever cycle is followed is influenced by the physiological state of the host and the level of cII and cIII proteins.

The repressor is essential for lysogenization. By binding to a series of two sets of operators, O_L and O_R, the repressor blocks transcription from two promoters p_L and p_R, which are responsible for the expression of the phage 'early' genes including *cro, O, P, Q* and *N,* whose products are required for the progression to the lytic cycle. In

addition, repressor bound at O_R stimulates expression from P_{RM} which is responsible for expression of *c*I during the prophage state of the life cycle. During the establishment of lysogeny, repressor synthesis occurs via the P_{RE} promoter, which is positively regulated by cII and cIII (Reichardt and Kaiser, 1971; Reichardt, 1975; Gussin *et al,* 1983). The repressor blocks synthesis from P_L and P_R, hence blocking synthesis of cII and cIII (Figure 1.2). Thus, expression of cI repressor switches primarily to the P_{RM} promoter (Gussin *et al,* 1983). The transition from expression from P_{RE} to P_{RM} is dependent on the occupancy of binding sites at the O_R repressor binding sites. The repressor binds to O_R1 with the highest affinity, followed by co-operative binding to O_R2 and eventually at relatively high concentrations to O_R3 (Meyer *et al,* 1980). Only when the repressor is bound to O_R3 is its own synthesis repressed (Ackers *et al,* 1982; Gussin *et al,* 1983).

Cro promotes lytic development by inhibiting the synthesis of the repressor. This is a result of binding of the Cro repressor to the operators, O_L and O_R (Figure 1.2). In contrast to the repressor, Cro binds most strongly to O_R3, thus inhibiting expression from P_{RM} and so reducing repressor synthesis (Johnson *et al,* 1978). At higher concentrations Cro binds to O_R2 and O_R1 to prevent expression from P_R promoter (Takeda, 1979; Gussin *et al,* 1983).

λ repressor is a dimer of 52kDa composed of two monomers of 236 amino acids. The repressor gene has been cloned and sequenced (Sauer, 1978) and the protein has been purified to homogeneity (Chadwick *et al,* 1970; Wu *et al,* 1971). Proteolytic cleavage studies using papain with the purified cI repressor have demonstrated that the DNA binding region is located within the first 92 amino acids of the N-terminus. The residues between amino acids 92-236 are responsible for oligomerization of the protein.

The N-terminal 92 amino acid fragment has been crystallized free (Pabo and Lewis, 1982) and bound to DNA (Jordan *et al,* 1985; Jordan and Pabo, 1988). In the free state the repressor was demonstrated to form dimers, but with only very weak association. The N-terminal domain was demonstrated to consist of 5 alpha-helices, the first four of which contribute to a hydrophobic core. The hydrophobic core has been shown to be essential for stabilization of the protein architecture (Lim and Sauer, 1989). The fifth helix is believed to participate in monomer-monomer interactions and has been shown to make contacts with the sister helix 5 in the dimeric structure. The crystal structure of cI N-terminal domain to O_L1 demonstrates that the DNA bends slightly with repressor bound. The repressor contacts the DNA by making specific interactions via the helix-turn-helix motif i.e. helices 2 & 3 (described in detail in section 1.2.4) which is stabilized by the first 5 amino acids of the N-terminus. The N-terminal arm wraps

around the DNA in the major groove to help stabilize the bound complex (Jordan and Pabo, 1988).

The *cro* gene has been cloned and sequenced (Hsiang *et al,* 1977; Roberts *et al,* 1977), and the Cro protein has been purified to near homogeneity (Folkmanis *et al,* 1976). It is very small consisting of only 66 amino acids and forms a dimer of 18kDa in solution. The crystal structure of the unbound Cro repressor has been elucidated and shown to consist of three alpha-helices and three strands of anti-parallel ß-sheet. Residues in the C-terminus are believed to be important for oligomerization to the dimeric state (Anderson *et al,* 1979; Anderson *et al,* 1980, Harrison and Aggrawal, 1991). The biological signifigance of the interactions of Cro and cI repressor with the same set of overlapping operators, demonstrate the economical nature of gene regulation within λ phage. Other regulatory systems have evolved for other phages and some of these novel features will be discussed in section 1.4.

In addition to their actions on the phage genome, cI and Cro have also been shown to complement host-encoded genes (Dai and Ishiguro, 1991). The *lyt* D gene of *E.coli* is involved in peptidoglycan hydrolysis and cell lysis. A *lyt* D mutant was shown to be complemented by either *cI* or *cro.* Lyt D is believed to be a DNA-binding protein which shares similar recognition sites to Cro and cI repressor. These observations pose interesting questions on the interaction of λ-encoded proteins with the host.

1.3.3. Role of *E.coli* Encoded Proteins in Regulating λ Development

Viruses are intracellular parasites and depend on host transcription and translational machinery to replicate. Host proteins such as integration host factor (IHF) from *E.coli* play a significant role in the development of λ phage. IHF was first identified as an activator of λ site-specific recombination, but has since been shown to play a role in other λ processes, in addition to its function in *E.coli* (Miller *et al,* 1979; Kikuchi *et al,* 1985). These roles include phage packaging (Miller and Feiss, 1988), cII gene expression (Hoyt *et al,* 1982; Mahajna *et al,* 1986) and stimulation of the λ P_L promoter which controls expression of 14 genes including cIII, N and those with recombination functions (Giladi, Gottesman and Oppenheim, 1990). In addition, IHF has been found to cause repression of the P_R promoter, most probably by steric-hinderance of the binding of RNA polymerase (Kur , Hasan and Szybalski, 1989). The role of IHF is to act as a regulatory protein. It binds to a well-conserved DNA sequence (Craig and Nash, 1984) and causes bending of DNA (Prentki *et al,* 1987; Thompson and Landy, 1988). Repression is most probably caused either by IHF occluding the RNA polymerase binding site, or by enhancement of the binding of a repressor protein (Jose-Gama, Toussaint and Higgins, 1992; Alazard, Betermier and Chandler, 1992). In

addition to λ, IHF also influences the expression of *E.coli* Mu phage genes, including the promoter for the lytic repressor and a separate promoter for the lytic genes (Jose-Gama, Toussaint and Higgins, 1992. The Mu repressor, in addition to IHF, also requires another host-encoded protein, H-NS (Falconi *et al,* 1991) and a particular DNA topology for active repressor binding (Jose-Gama, Toussaint and Higgins, 1992). These data suggest that phages have become adapted to their host, to such an extent that host-encoded proteins are siphoned away for use in phage expression. The interaction of *E.coli* phages with their host is unlikely to be an isolated example and most probably occurs in some form or other, in most phage:host interactions, and also in virus:cell interactions.

1.4. Interesting Features in Other Bacteriophages

A number of interesting features have evolved to regulate gene expression in phages other than λ. Two examples are presented below which have a particular significance to our findings with ØC31.

1.4.1. P22 Phage and the Anti-repressor

P22 is a bacteriophage from *Salmonella typhimurium* which is remarkably similar in genomic organization and a fairly high degree of DNA sequence homology to λ (Skalka and Hanson, 1972). The right-hand half of the phage differs somewhat to λ and encodes the 'late' genes required for the lytic cycle (King, 1973). Within the non-homologous region are encoded a number of proteins which play an important role in influencing the developmental fate of P22. This region (*immI*) is effective by interacting with the P22 c2 repressor, which is functionally homologous to the λ cI repressor (Susskind and Youderian, 1983).

The P22c2 repressor is negatively regulated by an anti-repressor protein, specified by the *ant* gene within the *imm*I region (Levine, 1972). The anti-repressor is responsible for the derepression of the lytic genes and was confirmed by the isolation of mutants (*ant*) which could not be induced by a temperature shift into a virulent state in a $c2^{ts}$ lysogen (Levine *et al,* 1975).

The anti-repressor gene forms part of an operon with expression emanating from the promoter P_{ANT} (Figure 1.3) The other product of the operon is an auto-regulatory protein called Arc (anti-repressor control) which is responsible for the negative regulation of its own synthesis and consequently that of the anti-repressor (Youderian *et al,* 1982). The Arc protein belongs to a recently discovered class of DNA binding proteins which interact with DNA via a ß-ribbon structural motif, as does another P22

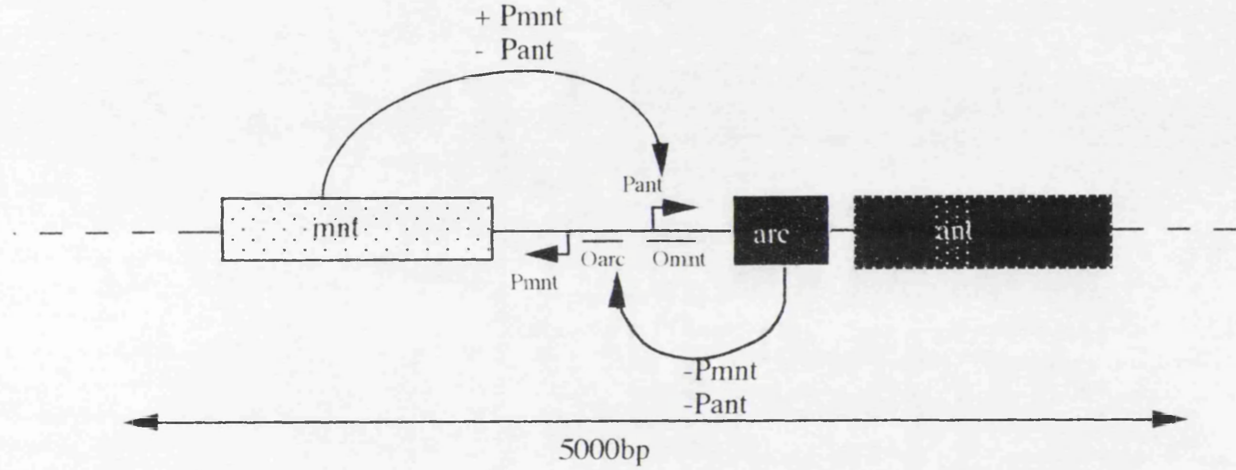


Figure 1.3. Control of Gene Expression in Bacteriophage P22 by Mnt, Ant and Arc

regulatory protein, Mnt. These two proteins are 40% identical when the amino acid sequence of each is compared (Sauer *et al,* 1983).

The Mnt protein, which is responsible for the **Maintenance** of lysogeny, acts by binding to an operator within the P_{ANT} region and repressing the expression of both Arc and Ant (Youderian *et al,* 1982). It is within this region that Arc repressor binds, and the recognition sites for both proteins are similar (Figure 1.3). The mode of DNA binding and the recognition sites are sufficiently similar that a hybrid protein containing the 9 N-terminal amino acids of Arc fused to Mnt, results in Mnt losing specificity for its own operator, and binding the Arc operator (Knight and Sauer, 1989). Recent data has concentrated on the mode of DNA binding by these two proteins and an NMR study for Arc has been published (Breg *et al,* 1990).

1.4.2. Bacteriophage P1

Bacteriophage P1 was originally isolated by G. Bertani in 1951 from *E.coli* (Bertani, 1951). It is an unusual phage in that it is a temperate phage which does not integrate into the host genome but remains as a plasmid of approximately 90kb in the cytoplasm (Yarmolinsky and Sternberg, 1988). Lysogeny is maintained by a complex immunity system of the *imm*I, immC and *imm*T regions which are widely separated from each other on the phage genome (Yarmolinsky and Sternberg, 1988). The C1 repressor which is a product of the *imm*C locus acts at least 14 sites, for which a consensus binding site of 17bp has been identfied (Velleman *et al,* 1987). The binding site is asymmetric and possesses directionality. In addition to C1 repressor, another small protein, Bof, has been implicated in the regulation of lysogeny. Bof complexes with the operator-C1 complex to down-regulate further gene expression (Velleman *et al,* 1990). Bof does not have the ability to bind DNA on its own (Velleman and Schuster, 1991). C1 repressor auto-regulates its own synthesis and also the synthesis of Coi protein (Heinzel *et al,* 1990). Coi protein functions by inactivating the C1 repressor and inhibiting it from binding to its operators i.e. functions as an anti-repressor. This is achieved by Coi oligomerizing with C1 to form an inactive C1:Coi complex (Heinzel *et al,* 1990). This reaction has been found to be reversible leaving both C1 and Coi in an active form (Heinzel *et al,* 1990). C1 and Coi are believed to be the crucial step in the decision between lysis and lysogeny. It is envisaged that a race between the synthesis of C1 and Coi ensues, if C1 wins and shuts-off synthesis of Coi then lysogenic development follows and if Coi wins then the lytic cycle predominates (Heinzel and Schuster, 1991).

1.5 *Streptomyces*- Hosts of ØC31

Studies of *Streptomyces* originally arose as a result of both their commercial importance as antibiotic producers and as a consequence of interest in their complex developmental cycle. They are Gram-positive, filamentous, aerobic, chemoheterotrophic, spore-forming eubacteria. Their cell wall is characterized by the prescence of L-L-diaminopimelic acid and glycine (in contrast to other members of the actinomyces). The DNA base composition of *Streptomyces* species has a very high d(G-C) content of between 69-73% on a molar basis, contrasting with that of the *Streptomyces* phage ØC31 which has a d(G-C) content of 63%. Pulsed-field gel electrophoresis has shown the genome size of *Streptomyces* species to be between 7 and 9 megabases (Woese, 1987; Hopwood and Kieser, 1990; Kieser, 1991; Kieser *et al.* 1992). Streptomycetes possess a complex life-cycle involving a penetrating substrate mycelium which gives rise to aerial spore-bearing hyphae.

Following an environmental stimulus such as nutrient deprivation or possibly a signalling mechanism, the spore germinates, and substrate mycelium spreads by apical extension and branching. The network of vegetative, branching hyphae secrete a number of hydrolytic enzymes. These enzymes play an important role in the recycling of biomass, degrading cellulose, lignin and xylan (McCarthy, 1987). At the substrate-air interface, the substrate mycelium differentiates to form hydrophobic, aerial hyphae (Chater and Merrick, 1984). The formation of the aerial mycelium is typically associated with lysis of the substrate mycelium. The onset of morphological differentiation often coincides with the biosynthesis of antibiotics (Chater, 1984), these are released into the substratum. The aerial hyphae septate and differentiate, producing long chains of spores. These represent the reproductive phase of the *Streptomyces* life-cycle and require dispersal by water, wind currents and mechanical transportation by mobile organisms (Chater, 1984).

The morphology, growth and formation of metabolites from *Streptomyces* bear a remarkable resemblance to the eukaryotic soil fungi and it may be more worthwhile to compare *Streptomyces* to these fungi than to enteric bacteria such as *Escherichia coli*. It has been demonstrated that some of streptomycete enzymes bear a remarkable similarity to their fungal counterparts. The enzyme 3-dehydroquinase from *Streptomyces coelicolor*, is highly homologous at the amino acid level and in function with the catabolic enzyme from *Neurospora crassa* and may result from either inter-species gene-transfer or convergent evolution (White *et al.*, 1990).

The genus *Streptomyces* have generated a lot of academic and commercial interest. The *Streptomyces* have a life-cycle which includes temporal, spatial and physiological

differentiation. Secondary metabolites (metabolites which are not essential for growth) produced by streptomycetes include many antibiotics, anti-cancer agents, immunosuppressants, anti-helminthics, and innumerable other compounds and proteins which have found a commercial use. The first commercial fermentation was the production of streptomycin from *Streptomyces griseus* in 1954 following the initial discovery of the anti-microbial action of streptomycin from *S.griseus* (Schatz, Bugie and Waksman, 1943), shortly followed by the commercial fermentation of chlorotetracycline from *S. aureofaciens* by Cyanamid (F. Burke, personal communication). These discoveries were later complemented by the genetic study of the genus pioneered by D.A. Hopwood in the 1950's.

Streptomycetes produce over 70% of all known antibiotics (Berdy,1980), and over 80% of the commercially-viable antibiotics (Baumberg and Hunter, 1989). *Streptomyces* are also used as the source of a number of industrially-important enzymes, including glucose isomerase, amylase, proteases, xylanases, lipases and a number of important restriction endonucleases (Baumberg, Hunter and Rhodes, 1989). *Streptomyces coelicolor* A3(2) has found favour as a strain for academic studies; it is the genetically best-characterized streptomycete species and efficient DNA transfer, protoplast formation and regeneration, and DNA isolation techniques were developed in the early stages of study. *S.coelicolor* has proved notoriously difficult to study at the physiological and biochemical level, due to difficulty in obtaining reproducible growth performance in liquid culture. This is partly the result of filamentous growth and low growth rates, combined with a tendency to form pelleted mycelial growth in liquid culture and has proved to be a bottle-neck to a coherent genetical and biochemical study of the organism. A major shift of emphasis, away from *S.coelicolor* to a more suitable *Streptomyces* species is one option which may provide a better, more coherent picture.

1.6. Regulation of Gene Expression in *Streptomyces*

Genetic regulation seems to exist at two distinct levels in *Streptomyces* (i) pleiotropic regulators- a number of proteins and factors have been identified which regulate gene activity at a global level. These include the regulation of morphological development, differentiation and responses to physiological stresses from the environment via cell messenger systems. (ii) local regulation- the regulation of transcription of a single gene/operon.

Pleiotropic control of gene expression currently known in streptomycetes is relatively limited. A number of candidate global regulators have been identified and include ppGpp (Ochi, 1986), A-factor (Horinouchi and Beppu, 1988), and*bld*A tRNA (Lawlor *et al,* 1987). In other bacterial species such as *E.coli*, ppGpp is known to alter the activity of RNA polymerase (Cashel et al, 1987). Studies in *S.coelicolor* have

demonstrated that although ppGpp is necessary for the production of the antibiotic actinorhodin, it was not solely sufficient (Strauch *et al*, 1991). It has been demonstrated that ppGpp is a factor in bialaphos production in *S.hygroscopicus* (Holt *et al,* 1992), but that it again was not sufficient to elicit the activation of the bialaphos operon via BrpA, the transcriptional activator responsible for switching on bialaphos biosynthetic genes (Holt *et al,* 1992). Recently, it has been postulated that the growth-rate of the organism could control antibiotic synthesis and is currently being investigated (Thompson, 1992).

A-factor was first identified as a regulator essential for streptomycin production and sporulation in *S.griseus* (Grafe *et al,* 1982; Hara and Beppu, 1982). It is believed to act by complexing with a regulatory DNA-binding protein (afsB) to derepress antibiotic and sporulation genes and is proposed to act at the transcriptional level (Miyake *et al,* 1990). In the search for the A-factor binding protein, other regulatory proteins have been identified. These include *abs*A from *S.coelicolor* , which has a global regulatory effect on the activation of antibiotic synthesis without affecting sporulation (Adamidis *et al,* 1990). A similar less severe phenotype, is observed for *afs*R from *S.coelicolor,* which when present in multi-copy results in antibiotic hyper-production (Horinouchi *et al,* 1983).

The *Streptomyces coelicolor bld* mutants were originally characterized by their inability to form an aerial mycelium (Chater, 1989). The best characterized of these is the *bld*A mutant which was found to encode the tRNA for the TTA codon (Lawlor *et al,* 1987). This codon is extremely rare in the d(G-C) rich streptomycetes and has been proposed to play a role in the morphological development of the organism. TTA codons are only found in a few genes and these appear to be activators of antibiotic pathways or resistance genes. This topic has been covered in great detail elsewhere (Chater, 1989: etc).

Analysis of promoter sequences in streptomycete species has recently taken on more signifigance due to the compilation of 139 promoter sequences (Strohl, 1992). Strohl concluded that although a σ^{70} like consensus sequence is found in 29 of the 139 promoters analysed, similar to that found in *E.coli* and *B.subtilis*, the majority of promoters displayed a great diversity of sequences. As a result of this study a new consensus sequence for the streptomycete σ^{70}- like -35 and -10 regions has been derived; TTGACPu (-35) and TAgPuPuT (-10). However, the remaining 80% of promoter sequences exhibit a wide range of sequences, some of which can be sub-divided to those which have sequence similarities in the -10 and -35 regions (Strohl, 1992). Strohl concludes that considerable effort must be employed to show functionality of the apparent promoter sequences shown in the study.

There are a number of examples of genes with tandem promoters in *Streptomyces* including the galactose operon from *S.coelicolor* (Westpheling and Brawner, 1989), *brp*A from *S.hygroscopicus* (Raibaud *et al,* 1991) *dag*A from *S.coelicolor* (Buttner *et al,* 1987), and the thiostrepton resistance gene of *S.azureus* (Janssen and Bibb, 1990). Other genes use tandem promoters to express more than one protein from the same ORF including the ØC31 repressor gene (Smith and Owen, appendix), the *tip*A gene (Caso *et al,* 1990) and the *bld*A gene from *S.griseus* (Babcock and Kendrick, 1990). Of all the streptomycete genes analysed to date, over 30% possess multiple promoters (Strohl, 1992), either divergent overlapping promoters e.g. *otrA*/*otrZ* (Doyle *et al,* 1991) and *act* II (*orf*I)/ *act* II (*orfs* 2,3) (Caballero *et al,* 1991) or tandem promoters e.g. *gal* utilization gene (Fornwald *et al,* 1987). These data suggest that gene regulation in streptomycetes is very complex.

A number of DNA-binding proteins from streptomycetes have recently been identified. These include both a putative activator and a repressor for the actinorhodin biosynthetic pathway (Fernandez-Moreno *et al,* 1991), a transcriptional repressor for tetracenomycin resistance genes in *S.glaucescens* (Guilfoile and Hutchinson, 1992), an activator of spiramycin biosynthesis in *S.ambofaciens* (Geistlich *et al,* 1992), and *brp*A, a transcriptional activator for the bialaphos resistance gene in *S.hygroscopicus* (Anzai *et al,* 1987). However, few regulatory proteins involved in primary metabolism have been identified. The few include the glycerol repressor from *S.coelicolor* which has been isolated (Smith and Chater, 1988; Hindle and Smith, 1991) and recently, a transcriptional activator for a metalloprotease gene from *S.coelicolor*. This was purified and shown to possess DNA-binding activity (Dammann and Wohlleben, 1992). To date all the potential DNA-binding proteins identified from *Streptomyces* possess a putative helix-turn-helix motif. Recently, however, the zinc dependent binding of a protein is believed to be responsible for the activation of the esterase gene in *S.scabies* and has been shown to bind zinc (Babcock *et al,* 1992), similar to other zinc-binding regulatory proteins in eukaryotes (Klug and Rhodes, 1987; Berg, 1989).

1.6.1 <u>RNA Polymerase Heterogeneity in *Streptomyces coelicolor*</u>

RNA polymerase holoenzyme in the eubacteria consists of 2 α sub-units, one ß, a ß' sub-unit and a single σ factor. The σ factor provides the specificity for binding of core RNA polymerase to the correct promoter in the appropriate orientation. The holoenzyme can then polymerise ribonucleoside triphosphates to form RNA (Burgess, 1969; Helmann and Chamberlin, 1987).

The complex life cycle exhibited by *Streptomyces coelicolor* A3(2) probably requires a precise physical and temporal co-ordination of gene expression. One mechanism of achieving such control is to use alternative σ factors to alter the specificity of the RNA polymerase holoenzyme and consequently allow the expression of discrete sub-sets of genes. Such a strategy has been known for some time in *Bacillus subtilis* where alternative σ factors mediate the temporal control of gene expression during sporulation and the regulation of lytic development of bacteriophage SPO1as well as other processes (Losick and Pero, 1981, Pero *et al*, 1982 Stragier and Losick, 1990; Errington and Illing, 1992). Σ factors also play an important role in T4 phage development in *E.coli* (Kassavetis and Geiduschek, 1984), nitrogen utilization in enteric bacteria (Hirschman *et al*, 1985; Hunt and Magasanik, 1985), the production of the heat shock proteins in *E.coli* (Grossman *et al*, 1984; Gross *et al*, 1987) and the expression of flagellar and chemotaxis genes in *Bacillus subtilis* , *Escherichia coli and Salmonella typhimurium* (Helman and Chamberlin, 1987; Helman, Marquez and Chamberlin, 1988; Mirel and Chamberlin, 1989; Chamberlin, personal communication). The σ factor responsible for chemotactic genes in *B.subtilis*, σ^D (formerly σ^{28}) bears remarkable homology at the amino acid level to an alternative sigma factor, *whi*G, from *Streptomyces coelicolor* which has been shown to be responsible for the development of spores from aerial hyphae during development (Chater *et al*, 1989).

RNA polymerase heterogeneity was first demonstrated by testing *Streptomyces coelicolor* partially-pure RNA polymerase extracts for *in vitro* 'run-off' transcribing activity on the *B.subtilis ctc* and *veg* promoters (Westpheling, Ranes and Losick, 1985). The *veg* promoter resembles the consensus canonical *E.coli* promoter sequence, widespread in the eubacteria (Moran *et al*, 1982). Fractionation of crude RNA polymerase preparations permitted the isolation of two distinct forms of RNA polymerase, associated either with σ^{35} or σ^{49}. The *veg* promoter is recognised by σ^{35} whilst σ^{49} seems specific for the *ctc* promoter. Th *ctc* promoter is recognised by a relatively minor σ factor in *B.subtilis* (σ^B) (Haldenwang and Losick, 1980; Johnson, Moran and Losick, 1983; Moran, Johnson and Losick, 1982). However, the promoter for the *endo*H gene from *Streptomyces plicatus* has been demonstrated to interact with σ^{49} (Westpheling, Ranes and Losick, 1985).

The *rpo*D gene encodes σ^{43} in *B.subtilis* and σ^{70} in *E.coli*; the major "house-keeping" σ factor in both species. Using an oligonucleotide probe to a conserved region of the two *rpo*D homologues, four *rpo*D homologues were found in *S coelicolor* A3(2), which were cloned and sequenced (Tanaka *et al*, 1988). The *hrd*B, *hrd*C and *hrd*D genes were later cloned independently using the *rpo*D of *Myxococcus xanthus* as a probe (Buttner, Chater and Bibb, 1990). All the *hrd* genes from *S.coelicolor* show

extensive homology with each other, and with the *hrd* genes of *E.coli* and *B.subtilis* (Tanaka *et al*, 1988).

Gene disruption of *hrd*A (Buttner, 1992), *hrd*C and *hrd*D by the mutational cloning method using ØC31 (Chater and Bruton, 1983) resulted in viable mutants which were apparently unaffected in differentiation, gross morphology and antibiotic production (Buttner, Chater and Bibb, 1990). No mutants of the *hrd*B gene could be isolated which implied that *hrd*B was an essential gene in *Streptomyces coelicolor*. Recently, a 66kDa polypeptide was isolated and N-terminally sequenced. This protein was found to be the product of the *hrd*B gene (Brown, Wood and Buttner, 1992). It is believed that the *hrd*B gene product is the major vegetative σ factor in *Streptomyces coelicolor*. It has been observed that ØC31 can infect *S.coelicolor hrd*A, *hrd*C and *hrd*D mutants, and does not require these σ factors (Buttner and Lewis, 1992) .

1.7 The Actinophage ØC31

The actinophage ØC31 is a temperate bacteriophage with a genome size of 41.2kbp of double-stranded DNA (Lomovskaya, 1972). The DNA has a moderate d(G-C) content (63%), based on a molar percentage of bases, and contrasts markedly with the high d(G-C) bias characteristic of the *Streptomyces*. Of 137 strains of *Streptomyces* tested, over half were identified as sensitive to infection by ØC31, illustrating the broad host range of this phage (Chater 1986). ØC31 was first isolated by testing culture filtrates of *S.coelicolor* A3(2) on the indicator strains *S.lividans* 66 and 130 and *S.anthocyanicus* 31, and observation for the formation of plaques (Lomovskaya *et al*, 1972). Typically *S.lividans* 66 is used as an indicator strain, as the phage shows a distinctive, large-plaque morphology. Lysogens of ØC31 are not inducible by UV irradiation or mitomycin C, only through a high multiplicity of infection with *c* gene defective or virulent phages and c^{ts} phages (Lomovskaya *et al*, 1980).

ØC31 is member of Bradley's group 2 phages and has a polyhedral head with a long tail (Bradley, 1967). The phage adsorbs most efficiently to the freshly-germinated spore tube, between 6-8hours post-germination, phage adsorbed after 13 hours do not enter the lytic cycle. The explanation of this observation has not been determined. One-step growth curve experiments for ØC31 on *S.coelicolor* A3(2) show a latent period of about 40min., a rise period of 50min. and a burst size of between 20-30 pfu. Using electron microscopy, it has been identified that the polyhedral head is 50nm wide and the long non-contractile tail is 100nm x 10nm (Suarez *et al*, 1984). At the base of the tail is a plate, 15nm in diameter with at least one pin and between the head and tail, a 'knot' is present. Analysis of the structural proteins of ØC31 by SDS-polyacrylamide gel electrophoresis has shown the existence of 17 polypeptides, of which four, the 51,

38.5, 29.5 and 28kDa proteins, constitute 84% of the total protein content of the phage particle (Chater, 1986).

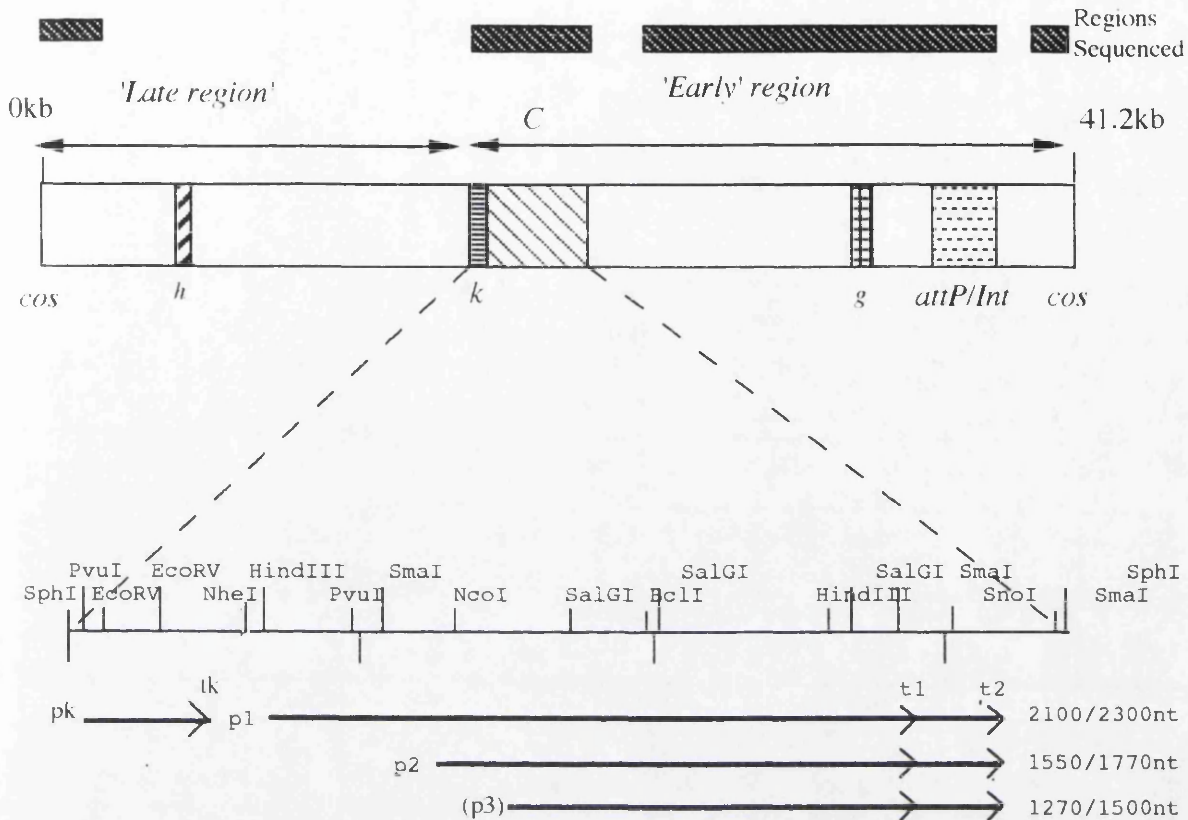
At either end of the ØC31 genome are cohesive ends (cos) suggesting that the DNA is packaged by specific cleavage of long concatamers of ØC31 DNA rather like that of λ phage (Hendrix *et al,* 1983). The *att*P integration site has been mapped to the right-hand end of the ØC31 genome and lies within the early region (Lomovskaya *et al,* 1980). This region has been sequenced and proven useful in the construction of a number of ØC31 derived vectors for site-specific integration within *Streptomyces* species (Chater, 1986; Kuhstoss *et al,* 1991).

The resistance of wild-type *Streptomyces coelicolor* A3(2) to plaque formation by ØC31 was originally considered to indicate the presence of a defective ØC31 prophage (Lomovskaya *et al,* 1970). This was later found not to be the case by southern blotting the *Streptomyces* genome with ØC31 DNA. The resistance mechanism was found to be chromosomally encoded, and was termed the factor for "phage growth limitation" (*pgl*) (Chinenova *et al,* 1982). The *pgl* locus was found to be unstable, *pgl*$^+$ *S.coelicolor* revert to the *pgl*$^-$ phenotype at a frequency of 10^{-2} to 10^{-3}, with the reverse phenotypic change occurring at frequencies between 10^{-3} to 10^{-4} (Chinenova *et al,* 1982). It has been postulated that such a high level of instability may indicate phase variation as found in *Neisseria gonorrhoae* pilin variation (Saunders, 1989) and the expression of outer membrane polysaccharides in *Haemophilus influenzae* (Weiser *et al,* 1989).

The adsorption and the ability of a *S.lividans* 66-derived ØC31 phage to form lysogens in either the *pgl*$^+$ or *pgl*$^-$ strains of *S.coelicolor* has been studied. Chinenova *et al,* 1982 found that during ØC31 one-step growth experiments, that a phage isolated from a *pgl*$^-$ host can undergo a single cycle of infection of a *pgl*$^+$ host. If the phage isolated from the *pgl*$^+$ host is then used to re-infect another *pgl*$^+$ host, then the phage has a much reduced capacity to infect although the same phage will give a normal productive infection in a *pgl*$^-$ strain. This results in phage that lose plaque-forming ability after going through two infective cycles of a *pgl*$^+$ host..

Recent work has shown that the Pgl system applies only to homo-immune phages and is specific to *S.coelicolor* A3(2) (M. Buttner, personal communication). A number of *pgl* mutants have been isolated, some of which have been found to be complemented by a 16kb fragment of *S.coelicolor* A3(2) DNA. This fragment did not confer phage

Figure 1.4. Genetic Map of ØC31 Genome



Transcription Map of the Repressor Gene

resistance to *S.lividans*. The two distinct types of mutants, those that can, and those that cannot be complemented by the exogenously- cloned 16kb of DNA containing the *pgl* locus have been classified as either PglA or PglB strains. Crosses between PglA$^-$ andPglB$^-$ strains revert to a Pgl$^+$ phenotype with a high frequency. It was also found that if the 16kb piece of DNA was passaged through a PglB$^-$ strain then it was no longer capable of complementing the PglA$^-$ strain. No ØC31 mutants resistant to the Pgl resistance system have been found (N.Lomovskaya, personal communication).

A model has been proposed to explain the Pgl phenomenon which takes account of the fact that phage are released after a single cycle of infection of a Pgl$^+$ strain, are modified so that they are unable to reinfect other Pgl$^+$ hyphae. They can, however, re-infect a Pgl$^-$ host. The model therefore states that the Pgl system must contain at least two components; one for phage alteration (*pal*) and the other for phage growth limitation (*pgl*) (Chinenova *et al*, 1982). It has been suggested (Buttner, pers. comm.) that the Pgl system may encode a restriction-modification system in which a restriction endonuclease would specifically degrade unmodified phage DNA. Several restriction-modification systems have previously been found in *Streptomyces* species (MacNeil, 1988).

1.7.1. Organization of the ØC31 Genome

Clear plaque mutants of ØC31 map to the centre of the phage genome (Figure 1.4), to the repressor locus (*c*) (Lomovskaya *et al*, 1980). The only other genetically defined loci include *'h'* which specifies hosts on which ØC31 can form plaques (Chinenova and Lomovskaya, 1975); *'g'* mediates events between the phage and host restriction-modification systems (Chater, 1986); *lyg* , believed to participate in prophage establishment, although some *lyg* mutants map to the *att* site (Sladkova *et al*, 1980) and *y* , a locus which has been demonstrated to alter the ratio of lytic to lysogenic events (Mkrtumian and Lomovskaya, 1972). Other than their genetic location the exact function of each of these genes is not known.

Heat induction of a c^{ts} allele induces expression of the lytic pathway (Rodriguez *et al*, 1986). To the left of the *c* gene lie a cluster of 'late' genes involved in phage packaging and other such processes, see Figure 1.2 (Suarez *et al*,1992; Lomovskaya *et al*, 1980; Ingham and Smith, 1992). As their name suggests, these genes are subject to temporal regulation and are expressed 20 minutes post-induction. To the right of the repressor lie the 'early' genes. These genes have been found to be expressed maximally at 10 minutes, with little or no mRNA detectable at 20 minutes post induction. Induction of the prophage leads to a 30-fold decrease in rRNA synthesis in *S.coelicolor* A3(2). The mechanism for this decrease in rRNA synthesis is unclear but which is believed to be

due to a putative activator enhancing the specific transcription of ØC31 genes at the expense of host-encoded genes (Clayton and Bibb, 1990). The evidence for an activator protein comes from findings that inhibition of RNA synthesis required *de novo* protein synthesis and although crude preparations of RNA polymerase from induced cultures had enhanced specificity for ØC31 promoters, this specificity was lost upon further purification (Clayton and Bibb, 1988). Data from this laboratory also suggests the possible presence of an activator protein in ØC31. It has been observed that phage promoter-*xyl* E reporter fusions are switched-on only in cells undergoing lytic development i.e. induced lysogens and is thought to be due to the expression of an activator protein (MCM Smith, personal communication). Transcription of the early region is believed to be mediated through the use of phage-specific promoter sequences. Sequence analysis of these promoter regions demonstrate a significant degree of conservation and a consensus promoter sequence can be identified (C. Howe and M.C.M Smith, personal communication). This sequence does not show any similarity to vegetative *Streptomyces* promoters (unlike the *c* gene promoters) and possibly requires activation before induction (C. Ingham and M.C.M Smith, personal communication). No potential σ factor or activator has yet been identified from the 10kb of the early region that has so far been sequenced (M.C.M Smith, personal communication).

Virulent mutations in ØC31 c^{+} arise very infrequently (10^{-10}). Most of the virulent mutants obtained are defective (*vd*) and cannot form plaques on non-lysogenic hosts (Lomovskaya *et al*, 1980). They are able to induce the resident prophage upon infection of lysogens and depend entirely upon the properties encoded by the prophage to function (Lomovskaya *et al*, 1980). Recent characterisation of a virulent mutant which has the ability to form plaques on a lysogen has demonstrated the *vir* phenotype to be mediated through a 7kb inversion in the early region upstream of the *att* site (M.C.M Smith, personal communication). It has been postulated that this inversion may have led to the derepression of a putative activator gene (M.C.M Smith, personal communication).

1.7.2. The ØC31 Repressor Locus

The repressor gene of ØC31 is responsible for the establishment and maintenance of lysogeny. Clear plaque mutants arise with a frequency between 1×10^{-3} to 5×10^{-3}. Attempted cross-infections of different *c* mutant isolates did not result in the formation of lysogens indicating that all *c* mutants map to a single complementation group (Mkrtumian and Lomovskaya, 1972). Temperature sensitive *c* mutants (c^{ts}) have also been isolated which form a clear-plaque at the restrictive temperature of 37°C but revert to a turbid plaque morphology at 28°C with a frequency of 1×10^{-4} (Novikova *et al*, 1973). Cross-infections of *c* mutants with c^{ts} mutants failed to form lysogens,

indicating that all mutations in the genes responsible for the establishment of lysogeny belong to the single complementation group (Novikova *et al,* 1973).

Deletion mutants of the *c* locus were determined using electron microscopy (Sladkova *et al,* 1980). Recombination between different *c* mutants, including the deletion mutants led to the conclusion that the *c* gene was almost certainly contained within a 3.4kb *Sph*I-G restriction fragment (Vasil'chenko *et al,* 1981). The entire *Sph*I-G fragment was cloned and sequenced, as were two of the deletion derivatives and the putative c^{ts} mutant (Sinclair and Bibb, 1988). From this work a potential ORF was identified which was predicted to produce a protein of 74kDa. A potential helix-turn-helix (HTH) DNA binding domain was identified near the C-terminus of the repressor protein which possessed a great deal of homology with the helix-turn-helix motif of λ cII (Figure 1.1). This homology was most striking in helix 2 of the HTH where 5 of 8 residues were identical and the other three residues shared similarity (Sinclair and Bibb, 1988). The degree of homology breaks down somewhat following this region and little homology is observed in helix 3. The conserved glycine at position 9 of the HTH (refer to section 1.2.3) has been replaced with a lysine although this substitution has been seen in other HTH motifs e.g. λ B58 protein (Sanger *et al,* 1982) and the *Shigella* isoIS1 *ins*A protein (Ohtsubo *et al,* 1981). It has recently been suggested that HTH motifs without a glycine at position 9 form a separate sub-section of HTH's (Brandon and Tooze, 1991). Helix 2 is involved in DNA recognition whilst helix 3 is believed to be responsible for specificity of binding (Ho *et al,* 1986 ; section 1.2.4)).

The deletion mutants have been compared to the wild-type gene. Deletion mutant Δ25 had lost 1102bp and Δ28 1401bp. The Δ25 mutant should result in a protein of 26 amino acids; 9 from the N-terminus and the others from out-of-frame translation to a stop codon at position 1958. The Δ28 mutant results in an in-frame fusion of 216 amino acids including the N-terminal 86 residues and the C-terminal 140 residues. This protein would be predicted to include the putative HTH motif but it failed to produce lysogens when tested (Sinclair and Bibb, 1988). Sequence analysis led to the discovery that the c^{ts} mutation at position 2265 was the result of a T to G transversion substituting a glycine for a tryptophan in the protein sequence (Sinclair and Bibb, 1988).

In vitro transcription-translation of the cloned *c* gene resulted in the appearance of a protein with a predicted molecular weight of 72kDa which matched the expected repressor protein. In addition, proteins of 51, 36 and 32kDa were also observed (Sinclair and Bibb, 1988). High resolution mapping demonstrated that a putative promoter (*cp*1) was present upstream of the ORF for the 74kDa protein which becomes derepressed upon induction of c^{ts} lysogen (Sinclair and Bibb, 1989). A

Figure 1.6. Amino Acid Composition of ØC31 Proteins Compared to Average Proteins

	Repressor:	74kDa	54kDa	42kDa	Average Protein
	Predicted Isoelectric Point:	4.93	5.03	4.93	
	Number of Amino Acids:	683	495	384	

Type of Residue	Residues	Mole %	Mole %	Mole %	Mole %
Aliphatic	A+G	20.059	19.394	20.052	16.9000
Hydroxyl	S+T	11.859	10.909	10.677	13.100
Acidic	D+E	18.302	18.586	19.792	11.600
Acidic+Acid Amide	D+E+N+Q	23.719	24.242	25.000	19.800
Basic	K+R+H	15.959	16.566	17.188	13.500
Hydrophobic	L+V+I+M	19.766	21.212	20.833	20.200
Aromatic	F+Y+W	3.807	3.030	3.125	8.300
Charged	D+E+K+R+H	34.261	35.152	36.979	25.100

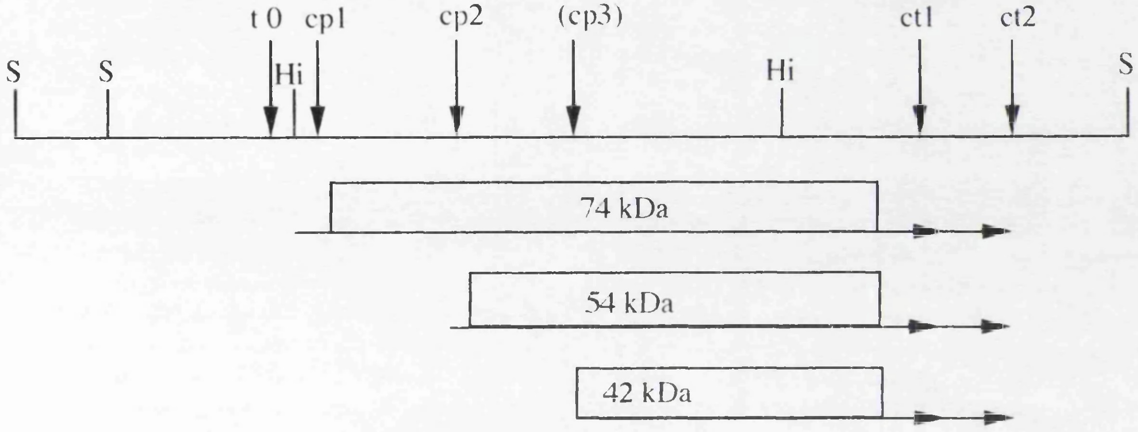


Figure 1.5: Organisation and expression of the c locus. The *Sph*I-I and *Sph*I-G fragments of ØC31 are shown (see figure 1.4). Horizontal arrows tepresent transcripts of the *c* gene. Translation of the c tran scripts produces the 74-, 54- and 42-kDa proteins as shown by open boxes. Downward arrows represent the positions of transcriptional signals; *c*p1 and *c*p2 are promoters. ct1, ct2 and t0 are terminators (see Figure 1.4). Hi, *HindIII*,:S, *SphI*

transcription terminator similar to the Rho-independent terminators in *E.coli* was mapped downstream (Sinclair and Bibb, 1989).

More recent work from this laboratory has developed the findings of Sinclair and Bibb and shown that expression of the *c* locus is more complex than initially anticipated. These data are presented in a published paper in the appendix of this thesis, but can be briefly summarised as follows. Low resolution transcript mapping of the *c* gene resulted in the identification of a nest of transcripts (Figure 1.5) which if translated would be predicted to produce N-terminally different in-frame proteins with identical C-termini from ORF's (Figure 1.6) and predicted molecular weights of 74, 54 and 42kDa respectively. The amino acid composition of the three repressor proteins is shown in Figure 1.6. This demonstrates that the repressor proteins probably have an overall negative charge due to preponderance of negatively charged amino acids and so may possess an acidic isoelectric point (pI). This is quite an unusual feature of DNA-binding proteins, an exception is the tryptophan repressor (TrpR) which has a pI of 5.9 (Joachimiak *et al*, 1983). The role of the overall negative charge has not as yet been determined, but may play a role in oligomerization (Carey, 1989).

Previous work has identified that at least one internal promoter is present in the *c* gene (Figure 1.4). Promoter *c*P2 initiates transcription downstream of *c*P1 and results in a transcript of 1550nt terminating at *c*t1. A second internal promoter (*c*P3) initiating a transcript of 1250nt has also been tentatively identified by low resolution S1 mapping. However, promoter-probing experiments with this promoter show only weak activity. Line-ups of the three *c* gene promoters are presented in Figure 1.7. This figure demonstrates that the -10 region of *c*P1 is identical to that of *tsr*p1 (Janssen and Bibb, 1990) which also does not possess a well-defined -35 region, the -10 region of *c*P2 is identical to that of *rrn*Dp4 (Bayliss and Bibb, 1988) and *c*P3 is similar to that of *gal*p3 (Fornwald *et al*, 1987). It should be observed that none of the three promoters contain a putative -35 sequence. This feature has been previously observed for promoters which require transcriptional activation (Busby, 1986; Ho *et al*, 1986).

Transcription terminates either at *c*t1 or *c*t2, which is located ~200bp downstream of *c*t1. In the wild-type lysogen, transcription termination of the *c* gene takes place preferentially at *c*t2, however, in a heat-induced c^{ts} lysogen transcription preferentially terminates at *c*t1. A number of terminators in ØC31 have now been identified and seem to be very well conserved (Smith *et al*,1992). The ribosome binding sites for each of the transcripts in the *c* gene are identical consisting of the sequence AAGGG which is similar in complementarity to a region close to the 3' end of the 16S rRNA of *S.lividans* 66 (Sinclair and Bibb, 1988).

Upstream of *c*P1 and *c*P2 has been identified a perfectly conserved inverted-repeat of 16bp (Sinclair and Bibb, 1988). Recent work has involved constructing a CIR oligonucleotide of the consensus 16bp motif plus the 4 nucleotides upstream and using this as a probe in DNA-DNA hybridization studies of the ØC31 genome (C.J. Ingham and MCM Smith, personal communication). Together with sequence analysis this has resulted in the identification of 11 different regions which bind the oligonucleotide (C.J. Ingham and M.C.M Smith, unpublished work). The 16bp perfect inverted-repeat is absolutely conserved in all the sequences identified except one. The sequence of the 12th isolate with the imperfect 16bp inverted repeat was identified as a result of analysis of sequence of the early region of the phage (G. Murphy, N. Hartley and K. Chater, personal communication). In addition, the region of homology extends further than the 16bp core-sequence so that larger inverted-repeats may be formed. This sequence has been titled the "conserved inverted-repeat" (CIR) and is predicted to be a possible repressor binding site. This prediction is based on two pieces of evidence. Firstly, a *c*P1-*lac*Z fusion in *E.coli* is repressed by the 42kDa repressor alone. Secondly, a derepressed *c*P1-kanamycin fusion in pIJ486 was isolated by selection for kanamycin resistance and found to have a point mutation within the CIR, suggesting that it may have a role as an operator.

Reporter gene fusions of *c*p1 with *lac*Z in *E.coli* have demonstrated that the putative 74kDa repressor protein may possess anti-repressor activity. If the fusion was co-transformed with a plasmid carrying an insert expressing the predicted 42kDa repressor, or both the 54 and 42kDa repressor proteins then the *c*P1 promoter became repressed and white colonies resulted. However, if the fusion was co-transformed with a plasmid containing an insert which expressed the 74, 54 and 42kDa repressor proteins, then blue colonies were once more observed indicating possibly that the 74kDa protein is antagonizing the true repressor proteins to derepress *c*P1 (N. Wood and MCM Smith, personal communication; Smith *et al,* 1992). This result concurs with data presented by Smith and Owen, shown in the appendix, where immunity to superinfection was tested. In experiments where *S.lividans* containing streptomycete plasmid derivatives of pMS201, pMS211, pMS221 and pMS231 were challenged with phages containing different *c* gene deletions, then immunity to superinfection was observed in the presence of the 54 and 42kDa repressor proteins only suggesting that the 74kDa repressor is not essential for immunity to superinfection. *Streptomyces lividans* expressing the 42kDa repressor only showed immunity to a phage with a complete deletion of the *c* gene, partial immunity to *c*Δ25 and poor immunity to Δ28. This observation suggests that the putative in-frame truncated protein may indeed be produced from *c*Δ28,and possibly interferes with the function of the 42kDa repressor.

```
                  20mer
      683     — — ►◄ — —                       -10          +1
cp1   gcaccCAGCCTACTCACGTGAGTAGCttggAgCGTgGgcTAGGGTGagcgAag
tsrp1           ggtcaggGcagccAtCcgccAtCGTcGcgTAGGGTGtcacA

cp1   caccacaacgacaacgAAGGGgcgggaaCATG
                         rbs

       1254           20mer
                 — — ►◄ — —
cp2    tggctCAGCCTACTCACGTGAGTAGtgTAGAGTCGCctACGCAcGACaacgacGAc
rrnDp4:     aAatCggaTCggaaagaTctgaTAGAGTCGGaaACGCAaGACcgaaggGA

cp2    aacAAGGGgacgggcacCATG
          rbs

       1531                                  +1                          +1
cp3    gctCGaaGgtCGTgCcGaccGgTGAGGTTtCcgtctccgaaagagtccgtTGAGGTtccgggcc

galp2  ttcCGccGtaCGTcCgGcaaGcTGAaGTTcCtcg

cp3    tgagtatgacggtcagtcagaagagctacacgcccgttgagggcgacgacAAGGGagagaaCATG
                                                           rbs
```

Figure 1.7. A comparison of the promoter sequences of the *c* locus with other *Streptomyces* promoters.

1.8. Aim of Experimental Work.

The repressor proteins from ØC31 are proposed to control the events that govern whether the phage enters the lytic cycle or becomes an integrated prophage during lysogenic growth. To be able to understand these mechanisms it was desirable to isolate the repressor proteins so that they may be characterised. Overexpression of the repressor proteins would aid the isolation of each of the repressor forms to a stage whereby the proteins may be characterised biochemically and their DNA-binding properties analysed. The biochemical characterisation of the repressor proteins would enable a picture of their interactions with themselves, and with other repressor proteins to be constructed. Finally DNA-binding studies would give valuable information about the sequences bound by the repressor proteins and the likely effect of DNA-binding on gene expression.

Chapter 2

Methods and Materials

2.1 Bacterial Strains. The bacterial strains used were all derivatives of *Escherichia coli* K12 and are listed in Table 2.1

2.2 Plasmids. The plasmids used and constructed are listed in Table 2.2

2.3 Synthetic Oligonucleotides were synthesized on an Applied Biosystems 391 PCR-Mate oligonucleotide synthesizer and are described in section 2.7.14.

Table 2.1 Bacterial Strains

Strain	Genotype	Source
E.coli Strains		
DS941	AB1157, *rec*F143, *lac*I^{q}, *lac*ΔM15 *thr*1, *leu*6, *his*G4, *thi*1, *ara*14, *pro*A2, *arg*E3, *gal*K2, *sup*37, *xyl*15, *mtl*12, *tsx*33, *str*31, *sup*E44	Stirling *et al* (1988)
TG1	*thi, lac, hsd*DΔ5, *sup*E, F' [*tra*D36, *pro*AB^{+},*lac*I^{q}, *lac*ZΔM15]	Gibson, (1984)
CB51	*dam, ara, thi,*Δ(*lac-pro*)	C. Boyd

Table 2.2 Plasmids

Plasmid	Relevant Genotype	Source
pUC18/19	Amp^R	Yanisch -Perron *et al*, (1985)
pBS1	Full length *c* gene in pUC18 (2.75kb *Hin*dIII-*Sph*I-G)	R. Sinclair
pIJ2410	Full-length *c* gene in pUC18 (3.4kb *Sph*I-G Fragment)	Sinclair and Bibb (1988)
pMS201	Full-length *c* gene in pUC19 (As for pBS1)	Smith and Owen (appendix)
pCO100	As for pMS201 with fill-in frame-shift at *Nco*I site	Chapter 3
pCO101	As for pCO100 but with deletion frame-shift at *Nco*I site	Chapter 3
pCO102	As for pMS201 but frame-shift at *Bcl*I site	Chapter 3
pMS231	Encodes *c* gene from *Pvu*I-*Sph*I in pUC19 with p*lac*	Smith and Owen (appendix)
pMS232	Encodes *c* gene from *Pvu*I-*Sph*I in pUC19 opposite p*lac*	Smith and Owen (appendix)
pMS221	Encodes *c* gene from *Nco*I-*Sph*I in pUC19 with p*lac*	Smith and Owen (appendix)
pMS222	Encodes *c* gene from *Nco*I-*Sph*I in pUC19 opposite p*lac*	Smith and Owen (appendix)
pMS211	Encodes *c* gene from *Bcl*I-*Sph*I in pUC19 with p*lac*	Smith and Owen (appendix)
pMS212	Encodes *c* gene from *Bcl*I-*Sph*I in pUC19 opposite p*lac*	Smith and Owen (appendix)
pCO103	Encodes *c* gene from *Nhe*I-*Sph*I in pUC19 with p*lac*	Chapter 3
pCO104	Encodes *c* gene from *Nhe*I-*Sph*I in pUC19 opposite p*lac*	Chapter 3
pCO105	Encodes *c* gene from *Eco*RV-*Sph*I in pUC19 with p*lac*	Chapter 3
pCO106	Encodes *c* gene from *Eco*RV-*Sph*I in pUC19 opposite p*lac*	Chapter 3
pCO107	Encompasses 475bp *Nhe*I-*Sma*I containing *c*p1 in pUC19	Chapter 6
p19H3-300	Encompasses 300bp *Hin*dIII-*Pvu*I containing *c*p1 in pUC19 on *Hin*dIII linkers	M. Smith
p19-B300	Encompasses 300bp *Pvu*I-*Nco*I containing *c*p2 in pUC19 on *Bam*HI linkers	M.Smith

2.4 Chemicals

Chemicals	Source
General Chemicals, biochemicals	BDH, May and Baker, Sigma and organic
solvents	Aldrich
Urea	Promega
Agarose	BRL
Media	Difco
X-Gal, IPTG	BRL
T7 Sequencing Kits	Promega
T4 Polynucleotide Kinase	Promega
Restriction Enzyme Buffers	Promega, BRL, New England
Nucleotides	Pharmacia, Promega
Prepared Polyacrylamide	Severn Biotechnology
Radiochemicals	NEN, ICN

2.5 Buffer Solutions

a) DNA Electrophoresis

10x TAE Buffer: 48.4g Tris, 16.4g NaAcetate, 3.6g $Na_2EDTA.2H_2O$, made up to 1 litre in distilled water, pH adjusted to 8.2 with glacial acetic acid.

10x TBE Buffer: 109g Tris, 55g Boric acid, 9.3g $Na_2EDTA.2H_2O$, made up to 1 litre in distilled H_2O; pH 8.3

5x Agarose Gel Loading Buffer: 0.025% (w/v) bromophenol blue, 0.025% (w/v) xylene cyanol, 25% (w/v) ficoll, 0.5% (w/v) SDS, 50mM EDTA.

b) Protein Electrophoresis

Laemmli Gels (Laemmli, 1970)

4x Separating Gel Buffer: 1.5M Tris.HCl pH8.8, 0.4% SDS
4x Stacking Gel Buffer: 0.5M Tris.HCl pH6.8, 0.4% SDS
10x Running Buffer: 30g Tris, 144g Glycine, 10g SDS, made up to 1 litre in distilled water

Tricine-SDS Gels (Schagger and von Jagow, 1987)

Anode Buffer:	0.2M Tris (24.2g/l) in volume of 300ml pH8.9
Cathode Buffer:	0.1M Tris (12.1g/l), 0.1M Tricine (17.9g/l), 0.1%(w/v) SDS (1g/l) in volume of 200ml; pH adjusted to 8.25 with addition of Tricine. Do not use HCl with this buffer
Gel Buffer:	3M Tris (36.3g/100ml), 0.3% (w/v) SDS (0.3g/100ml) pH8.45
Sample Buffer:	SDS (2% w/v), 2-mercaptoethanol (2% v/v), bromophenol blue (0.01% w/v), glycerol (10% v/v)

Native Gels:(Davis, 1964; Ornstein, 1964; Goldenberg, 1989)

4x Separating Gel Buffer:	18.2g Tris 0.23ml TEMED 1M HCl to pH8.9 Distilled water to 100ml.
8x Stacking Gel Buffer:	5.7g Tris 0.46ml TEMED 1M H_3PO_4 to pH6.9 Distilled water to 100ml.
10x Running Buffer:	6.0g Tris 28.8g Glycine Distilled water to 1000ml.
5x Marker Dye Solution:	5.0g Glycerol 4ml Water 1ml 0.1% Bromophenol Blue.

c) Western Blotting (Towbin *et al*, 1979)

Blot Electrode Buffer (BEB):	25mM Tris $3.0gl^{-1}$ 192mM Glycine $14.4gl^{-1}$ 20%(v/v) Methanol $200ml.l^{-1}$
Phosphate-Buffered Saline (PBS):	10x Dulbecco's PBS (Sigma Chemical Co., Poole, Dorset).
Blocking Buffer:	0.2% (w/v) Gelatin 0.1% (v/v) Triton X-100 -made up to 500ml in Dulbecco's PBS.
Secondary Antibody:	Goat anti-rabbit IgG-Horse Radish Peroxidase conjugate - (Northeast Biomedical Laboratories Ltd., Uxbridge, Middlesex.).
Developing Solution:	2.5mg Diaminobenzidine.4HCl in 50ml of 50mM Tris.Cl pH7.4.

d) *in vitro* DNA Manipulation

10x Restriction Buffers:	The buffers recommended and supplied by enzyme manufacturers were used. Stored at 4°C.
5x/10x Ligase Buffer:	Ligase buffer supplied by BRL or Promega, stored at -20°C.
T4 Polynucleotide Kinase Buffer	Buffer supplied by Promega stored at -20°C.
1x Klenow Buffer	50mM Tris.HCl, 10mM$MgSO_4$, 0.1mM DTT; pH7.2
1x Mung Bean Buffer	30mM CH_3CO_2Na pH5.0, 50mM NaCl, 1mM $ZnCl_2$, 50% Glycerol.
Mung Bean Dilution Buffer	10mM CH_3COONa pH5.0, 1mM Cysteine, 0.1mM Zinc acetate, 0.1% Triton X-100, 50% Glycerol
(1x) TE Buffer	10mM Tris.HCl, 1mM EDTA; pH8.0.

e) Other Buffers All phenol used was purchased in the pre-distilled form, contained 0.1% 8-hydroxyquinoline and was buffered against 0.1M Tris.HCl pH8.0.

Alkaline Lysis Plasmid Preparation Solutions:

Birnboim-Doly I:	50mM Glucose, 25mM Tris.HCl pH8.0, 10mM EDTA
Birnboim-Doly II:	0.2M NaOH, 1% SDS: freshly prepared
Birnboim-Doly III:	5M Potassium acetate: equal volumes of 3M CH_3COOK and 2M CH_3COOH, pH will be 4.8

STET Buffer Solution:	8% Sucrose, 5% Triton X-100, 50mM EDTA, 50mM Tris.Cl pH8.0,

2.6 Microbiological Methods

2.6.1 Culture Media Chemicals used for the preparation of growth media and other microbiologically useful purposes were obtained from BDH Chemicals Ltd, Poole, Dorset; Difco Laboratories, Detroit, Michigan, USA and Sigma Chemical Co. Ltd, Poole, Dorset.

L-Broth:	10g Tryptone, 5g Yeast extract, 5g NaCl, 1g Glucose, 20mg thiamine made up to 1 litre in distilled water and adjusted to pH7.5 with NaOH.
L-Agar:	as for L-broth but with addition of 15g/l Difco Bacto-agar and no added glucose.

2.6.2 Sterilization All growth media were sterilized at 120°C for 15 minutes; supplements and buffer solutions at 108°C for 10 minutes and $CaCl_2$ at 114°C for 10 minutes. Heat labile solutions were sterilized by filtration through Nalgene 0.22μm pore membranes.

2.6.3 Selection Supplements

a) X-Gal (5-bromo-4-chloro-3-indoyl-ß-galactoside)

X-Gal provides a chromogenic substrate which forms a blue colour on agar when in the presence of ß-galactosidase from the *E.coli lacZ* gene induced by IPTG. Many plasmids and bacteriophage vectors use the *lacZ* gene in conjunction with an appropriately located multiple cloning site as an indicator for the presence of a cloned insert. A cloned insert disrupts the expression of the *lacZ* gene and results in the formation of a white colony. X-Gal at a stock concentration of40mg/ml in Dimethylformamide (DMF) was stored at -20°C and added to growth medium at a final concentration of 20ug/ml.

b) Antibiotics. The antibiotic concentrations used for broth and plate selection were as follows:

Antibiotic	Source	Stock Solution	Selective Concentration
Ampicillin (Amp)	Plasmid	20mg/ml (water)	200ug/ml
Streptomycin (Sm)	Chromosome	10mg/ml (water)	100ug/ml

All antibiotic stock solutions were stored at -20°C. Antibiotics were only added to growth medium pre-cooled to <55°C.

2.6.4 Bacterial Growth Conditions. Liquid cultures of *E.coli* were routinely grown at 37°C shaking vigorously. Plate cultures were grown at 37°C. Antibiotics were added as required.

2.6.5 E.coli Strain Preservation. *E.coli* strains were stored frozen at -70°C in 50% L-broth, 20% glycerol, 1% peptone and 29% water.

2.6.6 Introduction of Plasmid DNA into E.coli Two methods of introducing plasmid DNA into *E.coli* were routinely used:

a) $CaCl_2$ Treatment An overnight culture of an *E.coli* strain was diluted 1 in 100 in 40ml of L-broth and incubated for about 2 hours shaking at 37°C until the optical density of the culture at OD_{600} had reached a reading of between 0.40-0.60 (approx. $10^8 ml^{-1}$). The transformation protocol used is described in Sambrook *et al* (1989).

b) Modified Hanahan Procedure (Hanahan, 1983) This procedure was used to obtain highly efficient competent cells with a transformation frequency greater than 10^8.

Reagents: TFB;	10mM 2[N-morpholino]ethanesulphonic acid (MES) buffer (pH6.3), 10mM RbCl, 45mM $MnCl_2.4H_2O$, 10mM $CaCl_2.2H_2O$.
DTT/KAc;	2.25mM DTT, 40mM K O_2CH_3 (pH6.0).

Preparation of Competent Cells: An overnight culture of *E.coli* grown up in 2TY broth was diluted 1 in 100 and grown up in 40ml of fresh 2TY broth. The culture was grown at 37°C, shaking until an OD_{600} of between 0.40 and 0.60 was reached. The cells were then centrifuged (4,000g, 4°C, 10min) in pre-washed 50ml polypropylene tubes. The cell pellet was gently resuspended in 2.5ml pre-cooled TFB and left on ice for 15min. The suspension was then treated with 100ul of Dimethylsulphoxide (DMSO) and left on ice. After 5min., 100µl of DTT/KAc was added and left for 10min. on ice. The final step before the cells were ready for use involved the addition of a further 100µl DMSO and incubation for a final 5min. on ice.

c) Introduction of Plasmid DNA into the *E.coli* Cell: The method of heat shock outlined by Sambrook *et al* (1989) was routinely used.

2.7 Manipulation of DNA

2.7.1 Large Scale Plasmid Isolation (modified Birnboim and Doly 1976)

A 200ml culture of stationary phase plasmid-containing cells were harvested by centrifugation (12,400xg, 10min, 4°C). The bacterial pellet was resuspended in 4ml of Birnboim and Doly Solution I. To this was added 8ml of freshly-prepared Solution II and the solution left on ice. After 5mins on ice 6ml of cold Solution III was added, and mixed until the contents of the tube became visibly clear. Cell debris and chromosomal DNA were removed by centrifugation (39,200xg, 30mins, 4°C). The plasmid DNA was precipitated from the supernatant with 12ml of isopropanol and left for 15mins at room

temperature. The DNA was isolated by centrifugation (27,200xg, 15mins, 20°C). The DNA pellet was washed with 70% ethanol and allowed to air dry. The DNA was resuspended in (1x)TE to a final volume of 2.1ml. To this was added 270ul 15mg/ml ethidium bromide solution. The DNA/EtBr solution was added to 4.37ml of a 5g CsCl in 3ml (1x)TE solution. This solution was then loaded into a Beckman Ti70 ultracentrifugation tube and overlayed with paraffin oil. The tubes were balanced to within 0.01g and gradients were centrifuged in a Beckman Ti70 fixed-angle rotor at 200.000xg for 16 hours at 20°C. The DNA was removed by first piercing the top of the tube with a needle, the DNA band was then drawn away using a needle and syringe piercing the side of the tube. If two bands were visible, only the lower band was removed. The ethidium bromide was removed by repeated isopropanol or butanol extractions. The DNA was dialysed to a large volume of (1x) TE before being precipitated in ethanol. The plasmid was resuspended in (1x)TE.

2.7.2 Small-scale Plasmid Isolation

a) The method of Holmes and Quigley (1981) was used occasionally.

b) A modifed alkaline lysis method initially developed by Birnboim and Doly (1979) was routinely used for isolating plasmids from *E.coli*.

A 1.5ml stationary phase culture of plasmid containing *E.coli* was harvested by centrifugation (12,000xg, 10mins,4°C). The cell pellet was resuspended in 200uμl of Solution I, to this was added 100μl of solution II.The suspension was mixed and 150μl of solution III was added and mixed. The cell debris and chromosomal DNA was cleared by centrifugation (12,000xg, 5mins, 4°C). The supernatant was drawn off and added to a new tube containing 800υl ethanol and 40ul 1M $MgCl_2$:10% sodium acetate. The solutions were mixed and allowed to stand for 2mins at room temperature. The plasmid DNA was collected by centrifugation (12,000xg, 10mins, 4°C). The plasmid pellet was washed with 70% ethanol and allowed to air dry. The DNA was resuspended in 50υl (1x)TE. RNaseA was added to restriction digests of this DNA (to 50υg/ml) before loading onto an agarose gel.

2.7.3 Precipitation of DNA. A solution containing DNA was precipitated of the DNA by using a solution comprising 1/10th volume 1M $MgCl_2$:10% sodium acetate and 2 volumes of ethanol or an equal volume of isopropanol.

2.7.4 Digestion of DNA with Restriction Endonucleases All digests using restriction endonucleases were carried out using procedures recommended by the manufacturer using the stated buffer and digesting at the recommended temperature for each enzyme. The reactions were stopped either by heat-killing suitable restriction endonucleases at 70°C for 15min. or by loading directly onto an agarose gel.

2.7.5 Construction of Frameshift Mutations

a) Klenow Fill-in Klenow DNA polymerase I will fill-in the 3' site of a 5' DNA overhang left after restriction enzyme digestion. The filling-in abolished the original restriction site and resultedin a modified sequence when re-ligated.

About 1 μg of DNA was cut to completion with a specific restriction endonuclease. The cut DNA was electrophoresed on an agarose gel and later stained with ethidium bromide. The desired band was isolated from the gel after visulaization with long wave U.V light. The DNA was extracted from the gel by electoelution and cleaned on a Schleicher and Schuell Elutip-D column. To the clean DNA was added 1x Klenow buffer, 0.2mM dNTP's, 3units Klenow in a volume of 25μl. This was left at room temperature for 15 minutes. The reaction was stopped by heating at 70°C for 10 minutes. The reaction was extracted once with phenol and then precipitated with ethanol. The DNA was re-ligated using T4 DNA ligase overnight at 16°C and the DNA introduced into *E.coli* the following day.

b) Mung-Bean Nuclease Removal of Ends The DNA was prepared in the same way as for the Klenow fill-in described above. To the clean DNA was added Mung Bean nuclease buffer to provide a 1x solution and 15units of diluted Mung Bean Nuclease. The DNA and the nuclease were incubated at 30°C for 15min. The reaction was stopped by placing on ice and adding NaCl to a final concentration of 300mM. The DNA was then cleaned by phenol extraction and ethanol precipitation, before being re-ligated and introduced into *E.coli* as described above.

2.7.6 Ligation of DNA Fragments The ligation of DNA molecules was usually carried out in a volume of 20μl comprising 1x ligase buffer (Promega, BRL) containing about 1unit T4 DNA ligase. Ligations were incubated either at room temperature for >3hours or overnight at 16°C. The molar ratio of insert to vector was adjusted given the requirements of the ligation. Directional cloning into two different sites usually consisted of a 3:1 insert to vector ratio. Blunt-ended or cloning where religation could be a problem, higher insert to vector molar ratios were used up to 10:1.

2.7.6.1 Dephosphorylation of Free Ends of DNA 5' terminal phosphate groups were removed by treatment with calf intestinal alkaline phosphatase (CIP). Reactions took place in 50μl of (1x) BRL React™ 2 Buffer at 56°C for 15 minutes in the presence of 1U of Boehringer Mannheim CIP. A further unit of CIP was then added and incubated at 56°C for a further 15 minutes. Reactions were stopped by two extractions with phenol and a final ethanol precipitation.

2.7.7 Phenol and Chloroform Extraction of DNA Solutions Phenol and chloroform extractions were performed as described by Sambrook *et al* (1989). Chloroform was mixed in a 24:1 ratio with isoamyl alcohol.

2.7.8 Agarose Gel Electrophoresis DNA was visualized on horizontal agarose gels. Routinely, 0.75-1.00% (w/v) agarose gels were used, although higher percentage gels up to 2% were occassionally used to separate fragments less than 1kb. Most commonly, TAE buffer was used, but on occassions TBE buffer was used.

2.7.8.1 Mini-Agarose Gel Electrophoresis Mini gels were generally run using the BRL model H6 gel kits. This system provides a rapid, reliable method for analyzing and quantitating DNA by visualization stained with ethidium bromide. Gels were pre-cast in formers in volumes of 20ml and stored at 4°C until needed. Agarose was dissolved in (1x) TBE or TAE by boiling at 100°C. Once dissolved, the solution was left to cool to 55°C before pouring into a caster. DNA was separated by applying a voltage of 2-10V.cm^{-1} for about 40-60min. Gels were electrophoresed in 500ml of TBE or TAE buffer. Once run, the gel was stained in a solution containing 200ng.ml^{-1} EtBr for about 20min., followed by a brief destain in distilled water for 10min. The separated DNA molecules were visualized on a 302nm UV transilluminator.

2.7.8.2 Large Agarose Gel Electrophoresis Large gels were used to provide good separation of fragments from restriction digests for accurate sizing and isolation. A volume of 100ml of molten agarose was used to form the gel in a gel former. The gels were usually run overnight by applying a voltage of less than 1V.cm^{-1}, or for about 2.5hours with an applied current of 4V.cm^{-1}. The gel was electrophoresed in a 1 litre volume of buffer. After electrophoresis, gels were stained for up to 30min. in 200ml of a solution containing 100ng.ml^{-1} ethidium bromide and detsained in distilled water until all excess ethidium bromide was removed.

2.7.8.3 Native Polyacrylamide Electrophoresis of DNA Polyacrylamide separation of DNA fragments is very useful when analyzing and separating fragments less than about 1kb. Native polyacrylamide DNA gels were generally run using the Mini-Protean™ kit manufactured by Bio-Rad. The glass plates, spacers and comb were cleaned with distilled water andthen with ethanol. The apparatus was assembled according to the manufacturers instructions. Invariably, (1x)TBE was used. A solution comprising 4% polyacrylamide in (1x) TBE was poured into the gel former and allowed to set with the aid of freshly-prepared ammonium persulphate and TEMED. Gels were usually run at a constant voltage of 100V for about 1 hour. After electrophoresis, gels were stained with 100ng.ml ethidium bromide and visualized under UV light at a wavelength of 302nm.

2.7.8.4 Denaturing Polyacrylamide Sequencing Gel Electrophoresis A BRL sequencing apparatus (Model S2) was used for high voltage polyacrylamide gel electrophoresis.

Gel Preparation: 6% (w/v) denaturing polyacrylamide gels were used for sequencing and footprint analysis of DNA:protein complexes. The acrylamide used in the gels was prepared in the liquid form by Severn Biotechnology Ltd. at a concentration of 40% (w/v) with a ratio of 19 parts acrylamide to one part bisacrylamide. The gels were prepared in the following manner:

40%(w/v) acrylamide	15ml
Urea (Promega)	50g
10x TBE	10ml
Distilled Water	35ml

- to give a final volume of approximately 100ml.

The urea was dissolved by heating in water in a microwave at the low setting for 30 seconds. The gel solution was allowed to cool to near room temperature before the gel was cast. Before pouring the gel, the solutions were degassed for 20min., then 200ul freshly-prepared 10% (w/v) ammonium persulphate and 60µl TEMED were added. For sequence analysis of clones a 0.4mm deep sharkstooth comb was used to form wells. When footprinting gels were cast, 0.4mm thick BRL "Delrin" combs were used to form 6ul capacity wells in the gel.

Assembly and Pouring of the Gel: The glass sequencing plates (40cm x 33cm) were cleaned thoroughly in distilled water and then swabbed with ethanol. The smaller glass plate was coated with a thin layer of vegetable oil ("PAM" cooking oil) which was spread evenly across the plate. Spacers (0.4mm thich were positioned along the two long sides of the gel and a strip of Whatmann 3MM filter paper across the bottom. The two plates were then clamped together using bulldog clips. The gels was poured at a 30° angle taking care to avoid bubbles. Once poured, the combs were placed in position at the top of the gel and the gel left between 30-60min. to polymerize.

Electrophoresis of Large Denaturing Polyacrylamide Gels: The gel was pre-electrophoresed for at least 30 minutes. prior to loading the samples at a constant power of 60W to allow the gel to pre-heat. Before loading, all samples were denatured at 95°C for about 5min. DNA was separated by electrophoresis at a constant power of 60W for about 2 hours. Post electrophoresis, gels containing ^{35}S-dATP were soaked in a solution consisting 10% methanol, 10% acetic acid for 30min. prior to drying the gel and autoradiography.

2.7.9 Photography of Ethidium Stained Gels Gels stained with ethidium bromide were viewed on a 302nm UV transilluminator and photographed using Polaroid Type 67 land

film, or using a Pentax 35mm SLR camera loaded with Ilford HP5 film. In all cases photographs were taken using Kodak Wratten filters (No. 23A). Ilford HP5 film was developed according to the manufacturers instructions.

2.7.10 Recovery of DNA from Agarose Gels Two methods were generally used for the isolation of DNA fragments from agarose gels after separation from contaminating DNA. For low resolution electrophoresis, when only 3-4 bands are expected and are well separated, then low melting point (LMP) agarose was utilized. When efficient separation of DNA fragments was required, or when fragments greater than 1kb needed purification, then regular gel electrophoresis was used followd by isolation of the band and electro-elution in a dialysis membrane.

2.7.10.1 Isolation of Small DNA Fragments from LMP Agarose Low melting point agarose gels usually consisted a 1% agarose solution in (1x)TBE. The gels were run using the BRL mini-gel system and were electrophoresed at room temperature at voltages less than a constant 50V to avoid heating effects and dissolution of the gel. The gel was stained with 100ng.ml^{-1} ethidium bromide and then viewed using a hand-held long-wave UV transilluminator. The band of interest was isolated and then weighed. To the band was added 3 volumes of (1x)TE and the gel solution incubated at 65°C for about 20min. or until the gel slice had completely melted. The gel solution was then extracted twice with an equal volume of phenol and twice with an equal volume of 24:1 chlorofom:isoamyl alcohol. The DNA was then precipitated by the addition of 2 volumes of ethanol and 1/10th volume of 1M $MgCl_2$:10% sodium acetate. The precipitate was collected by centrifugation (12,000g, 10min. 4°C), washed with 70% ethanol and allowed to air dry. The DNA pellet was resuspended in about 20µl (1x)TE. Routinely, 2µl would be run on a gel to estimate the amount of DNA obtained and consequently the efficiency of the recovery.

2.7.10.2 Isolation of DNA from Agarose Gels Using Electroelution Gel slices were isolated from agarose gels after staining with ethidium bromide and viewing under long wave UV. The gel slice containing the DNA fragment of interest was excised with a scalpel and trimmed of excess agarose. As a matter of routine, gels in which fragments were to be isolated for use later were always run in (1x) TAE buffer. The gel slice was placed in a dialysis membrane with about 400µl of (1x)TAE. All air was excluded from the membrane and the open-ends clipped shut. The gel slice was layered along one side of the sack. The dialysis membrane was placed in a large electrophoresis tank containing (1x)TAE buffer. It was positioned such that the gel slice was at the farthest side for the anode. A voltage of 100V was applied for about 60min. The progress of the elution could be visualized using a hand-held UV source. Once all the visible DNA has left the gel slice and has stacked up along the opposite side of the dialysis membrane, the polarity

of the applied voltage was reversed for 1min. This had the effect of releasing all the eluted DNA into solution. The solution was then drawn up from the dialysis bag, the bag was washed with 200µl fresh (1x)TAE. The DNA was now ready for further purification. This usually consisted of a phenol:chloroform:isoamyl alcohol extraction and an ethanol precipitation. The DNA recovered was usually resuspended in about 20µl (1x)TE. Quantification of the recovery is usually performed by running 2µl on an agarose gel.

2.7.11 Elution of DNA from Polyacrylamide Gels DNA was eluted from excised fragments in polyacrylamide gels in 400µl (1X)TE at 65°C for 30 minutes. The DNA was collected by ethanol precipitation.

2.7.12 Quantification of Nucleic Acid The concentration and purity of nucleic acid was determined spectrophotometrically by measuring the absorbance of a 1ml DNA solution through a pathlength of 1cm at 260nm. BRL quartz cuvettes were used to measure absorbances. An absorbance of 1.0 at a wavelength of 260nm is equivalent to a solution concentration of $50ug.ml^{-1}$ double-stranded DNA or $\sim 20ug.ml^{-1}$ of single-stranded DNA and oligonucleotides. By determining the ratio of A_{260}/A_{280} the purity of the nucleic acid sample can be determined. A ratio af 2.0 is a good indication of a pure DNA preparation, any contamination will significantly lower this value.

2.7.13 Plasmid DNA Sequencing Plasmid DNA sequencing was used to confirm the identity of small DNA inserts and also as a comparative sequence on footprinting gels. A modified method of that published by Hsiao (1991) was used. The T7 sequencing kit (Promega) was the source of all buffers, chemicals and enzymes used in the method described. To determine DNA sequence, between 4-5µg of supercoiled plasmid DNA is required in a volume not exceeding 5µl (1x)TE. To the DNA was added $10ng.\mu l^{-1}$ of oligonucleotide sequencing primer and 1µl of 1M NaOH and incubated at 37°C for 10min. to denature the double-stranded DNA. This was neutralized by the addition of 1ul 1M HCl and 2ul of (5x) T7 buffer and left for a further 10min. at 37°C. At this time four 0.5ml tubes were labelled A,G,C and T. To each small tube was added 2.5ul of the appropriate dideoxy termination mix.

For n reactions the following was mixed on ice:

0.50n x µl	5x dG labelling mix
1.00n x µl	100mM DTT
1.00n x µl	DMSO
1.55n x µl	dH_2O
0.50n x µl	^{35}S-dATP ($1,300Ci.mmol^{-1}$) ~10µCi

The T7 polymerase was diluted in (1x) T7 buffer to a concentration of 1.5U.μl^{-1}. The termination tubes were placed at 42°C. To the DNA tube was added 3.5µl of the labelling mix and left on ice for 2min. After 2min., 2µl of the diluted T7 polymerase (3U) was added to the DNA tube and left to incubate at room temperature (23°C) for 3min. Once the labelling reaction had initiated, 3.5µl of the labelling mix was added to each of the four termination tubes, and left for 5min. at 42°C. The reaction was stopped by the addition of 4µl of the STOP mix. The reactions may be frozen at -20°C and stored for up to one week at this stage. The samples were subjected to denaturing polyacrylamide gel electrophoresis after initially denaturing the DNA at 90°C for 3min. Up to 3ul was loaded onto a typical sequencing gel.

2.7.14 Oligonucleotides-Synthesis and Use

2.7.14.1 Oligonucleotide Synthesis Oligonucleotides were synthesized on an Applied Biosystems PCR Mate DNA Synthesizer according to instructions supplied by the manufacturer. DNA synthesis reagents were purchased from Applied Biosystems Ltd. and Cruachem Ltd.

2.7.14.2 Deprotection of Oligonucleotides Oligonucleotides are synthesized 3'-5' via chemical addition of individual nucleotides to an initial nucleotide attached to a matrix of inert beads within a plastic column. To deprotect the oligonucleotide and release it from the matrix, the column was first broken open and the beads carefully poured into a 2ml Nunc cryo-tube. To this was added 1ml of fresh 30% ammonia and then left at room temperature for between 1 and 2hours. The solution was then vortexed and the beads allowed to settle. The ammonia was drawn off and placed in a new Nunc tube. A further 1ml of fresh ammonia was added and the tube left to incubate at 50°C overnight. The tube was allowed to cool before opening. The ammonia solution of the deprotected oligonucleotide was stored in ammonia at -20°C. Usually, one quarter of the deprotected oligonucleotide solution was precipitated by the addition of 2 volumes of ethanol and making the solution up to 0.5M ammonium acetate. Precipitated oligonucleotides were resuspended in 200µl (1x)TE and the concentration measured by spectrophotometry (see Section 2.7.12).

2.8. Preparation of Radio-labelled DNA Probes

2.8.1. 5'-End-labelling of Oligonucleotides Oligonucleotides were labelled at their 5' ends using T4 polynucleotide kinase and gamma-ATP. Reactions were carried out in a volume of 10µl:

	Volume (µl)
10x Polynucleotide kinase buffer	1
60pmoles Oligonucleotide	3
50uCi [gamma] ATP (6,000Ci.$mmol^{-1}$)	5

T4 Polynucleotide kinase (8U)	1

Reactions were incubated at 37°C for between 30 and 60min. Reaction was stopped by heating at 70°C for 10min. The unincorporated ATP was removed by Sephadex G50 chromatography.

2.8.2. 5' End-labelling of DNA Fragments DNA fragments were end-labelled at a single end only for DNAseI footprinting reactions. DNA (5µg) was cut with an appropriate restriction enzyme. 5' terminal phosphates were removed as described in section 2.7.6.1. The dephosphorylated DNA was resuspended following precipitation in 10ul γ-^{32}P andincubated with 3U of T4 polynucleotide kinase at 37°C for 40 minutes in (1X) Promega T4 kinase buffer. Reactions were stopped by heating at 70°C for 10 minutes. The DNA was then cut with a second restriction enzyme to release the DNA fragment of interest. The fragments were purified after electrophoresis on 4% polyacrylamide gels in (1X) TBE and eluted according to the method described in section 2.7.11.

2.8.2 Sephadex G50 Column Chromatography A 1ml syringe was pre-packed with Sephadex G50 prepared as described in Sambrook *et al* (1989). The column was pre-equilibrated with 2vol. (1x)TE. The labelling reaction was mixed with 5µl blue dextran and 5µl phenol red. This mixture was loaded onto the column and washed through with (1x)TE. Single drop fractions were collected and counted in a scintillation counter. The blue dextran co-eluted with the DNA fragments.

2.8.3 PCR to Produce End-Labelled Probes from the Repressor Gene A single oligonucleotide was 5' end-labelled as described above. The template DNA was linearized and if possible a restriction fragment containing the DNA to be polymerased was isolated and purified from an agarose gel. Most fragments in pUC vectors were isolated on a *Pvu*II restriction endonuclease fragment, fractionated on an agarose gel, purified from the gel and used in the PCR reaction.

Component	Volume (µl)
Template DNA (2-5ng)	2
Labelled Oligonuleotide (30pmoles)	5
Oligonucleotide 2 (30pmoles)	1
10x Taq DNA Polymerase Buffer	5
1.25mM dNTPs	8
Taq DNA Polymerase (1U)	1

Final reaction volume of 50µl. The contents were overlaid with a drop of mineral oil and the PCR assay initiated. The "hot-start" method was not used because of the dangers of radioactive contamination.

2.8.3.1 PCR Protocol for Repressor Promoter Fragments:

Step 1: 94°C Denaturation - 4min.
Step 2: 94°C Denaturation - 30sec.
45°C Annealing - 2min.
72°C Elongation - 20sec.
-repeat for 30 cycles.
Step 3: 94°C Denaturation - 30sec.
45°C Annealing - 2min.
72°C Elongation - 5min.
-repeat for 1 cycle.
Step 4: 4°C soak.

2.8.3.2 Oligonucleotides Used for PCR End-labelling

Oligo 1: M13/pUC -40 forward primer - 17mer

5'-GTTTTCCCAGTCACGAC-3'

Oligo 2: M13/pUC Reverse Primer - 18mer

5'-CAGGAAACAGCTATGACC-3'

Oligo 3: Downstream of *c*P1 - 17mer Lab. Name: Oligo 473

5'-TGATGAGCTGGTGCGGG-3'

Oligo. binds downstream of *c*P1 between nucleotides 1333-1349.

Oligo 4: Downstream of *c*P2 - 17mer Lab. Name: Oligo 451

5'-CTACAAGCCCCGCTGCA-3'

Oligo. binds downstream of *c*P2 between nucleotides1887-1903.

2.8.4. Preparation of Early Gene DNA Fragments by PCR PCR was used to prepare DNA from the early gene region for use in band-shift assays.

Component	Volume (ul)
Template Mini-prep. DNA (2-4ng)	2
Oligonucleotide 1 (30pmoles)	1
Oligonucleotide 2 (30pmoles)	1
10x Taq Polymerase Buffer	5
1.25mM dCTP, dTTP, dGTP, dATP	6
Taq Polymerase (1U)	1

2.8.4.1 PCR Protocol for Preparation of Early Gene Promoter Fragments

Step 1: 94°C Denaturation - 30sec.
Step 2: 94°C Denaturation - 30sec.
65°C Annealing - 30sec.
72°C Elongation - 30sec.
- repeat for 30 cycles
Step 3: 94°C Denaturation - 30sec.
65°C Annealing - 2min.
72°C Elongation - 2min.

- repeat for 1 cycle.

Step 4: 4°C soak.

2.8.4.2 Oligonucleotides used as Primers for PCR of the Early Gene Region

CIFC 24mer 5'-TCAGTGGTCACGCGCGAATAACGA-3'

-sequence between 7590-7613 of early gene region.

CIGC 23mer 5'-ACATCATCTGACACCCCACCTAAG-3'

-sequence between nucleotides 7643-7666 of early gene region.

CHFC 24mer 5'-GTACGGCGATGGCATATGACAGAA-3'

-complementary strand oligonucleotide between nucleotides 7722-7699 of early region.

CIIC 24mer 5'-TTCCGGTCCGTCGTGCTCAACGAA-3'

-complementary strand oligonucleotide between nucleotides7867-7844 of early gene region.

2.8.5 Clean Up and Analysis of PCR Products PCR products were purified from mineral oil by extraction with 100µl 24:1 chloroform: isoamyl alcohol and 50µl (1x)TE. This was centrifuged at 12,000g for 2min. in a microfuge. The upper aqueous layer was removed and passed down a Sephadex G50 column. The fractions containing the labelled PCR product were pooled and to them was added 2 volumes of ethanol and $1/10^{th}$ volume 1M $MgCl_2$:10% sodium acetate and left on ice for 20min. The precipitate was collected by centrifugation (12,000g, 4°C, 10min.) and left to dry before being resuspended in 100µl (1x)TE. The product was analyzed on a 4% polyacrylamide gel run on the Bio-Rad Mini-Protean gel kit. The gel was stained with $200ngml^{-1}$ ethidium bromide and visualized on a 302nm UV transilluminator. An autoradiograph of the gel was also taken.

2.8.6. Incorporation of Radio-labelled Nucleotides The percentage of radiolabelled nucleotides incorporated into DNA probes was calculated by spotting one microlitre of the labelling reaction onto a Whatman GF/C 2cm. diameter filter in duplicate. The filters were allowed to air-dry. One of the filters was then washed in 100ml of ice-cold 5% (w/v) trichloroacetic acid (TCA) and 20mM sodium pyrophosphate for 2min. The washing step was repeated three times in fresh TCA. The filterwas then washed in 70% ethanol for 30secs. and allowed to air-dry. Both filters were then placed in scintillation vials and the amount of radioactivity measured by the Cerenkov method. The proportion of radio-labelled precursors and specific activity of the probe were calculated using the following equations:

$$\frac{\text{cpm in washed filter}}{\text{cpm in unwashed filter}} = \text{proportion incorporated}$$

proportion incorporated X total weight = total amount of product

cpm of incorporated = specific activity

total amount of product

2.8.7 Autoradiography Autoradiography was performed in metal cassettes using Kodak X-OMAT x-ray film. A du Pont Cronex Lightening Plus intensifying screen was used to enhance autoradiographic images by placing at -70°C. X-ray films were developed in a X-Ograph.

2.9 Methods Used in the Analysis and Purification of the Repressor Proteins

2.9.1 Chromatography Media Various types of chromatographic media were used during the course of the purification of the repressor proteins. DEAE-Sephacel and double-stranded DNA-cellulose (Burgess and Jendrisak, 1975) was obtained from Pharmacia plc (Milton Keynes, Bucks) and heparin-agarose from Bio-Rad plc. Chromatographic media was packed in 10ml syringe columns plugged with glass wool and equilibrated according to the manufacturers instructions.

2.9.2 Protein Separation using Liquid Chromatography Liquid chromatography was performed using the MonoQ (HR5/5) and Superose 6 (10/30) pre-packed columns (Pharmacia plc) on a Waters Advanced Protein Purification System. A description of the nature and use of the MonoQ and Superose 6 columns is given in Chapter4 and 5 respectively.

2.9.3 Quantification of Protein Protein concentrations were measured according to the method of Bradford (1976) with bovine serum albumin as standard. Absorbance was measured at 595nm on an LKB spectrophotometer.

2.9.4 Electrophoresis of Protein Gels

2.9.4.1 SDS-Polyacrylamide Gel Electrophoresis Denaturing polyacrylamide gel electrophoresis of proteins was performed according to the method described by Laemmli (1970). The ratio of acrylamide:bis-acrylamide was 29:1. The polymerization was catalyzed by the addition of 0.05% ammonium persulphate and 0.03% N,N,N',N' tetramethylethylenediamine. Routinely, a 10% separating gel with a 4% stacking gel was used. Most gels were run overnight at a constant current of 17mA for 8 hours using an LKB Ultraphor electrophoresis power pack.

2.9.4.2 Native Continuous Polyacrylamide Gel Electrophoresis Native continuous gel electrophoresis was performed with a 4% polyacrylamide gel (29:1 acrylamide:bis-acrylamide) in 50mM sodium phosphate pH6.6. Gels were run at 4°C at 100V.

2.9.5 Visualization of Protein Polyacrylamide gels were stained with either Coomassie R-250 when large amounts of protein were present. The silver staining procedure of

Morrissey (1981) was used when greater sensitivity than that gained by Coomassie staining was required.

2.9.5.1 Coomassie Staining. Polyacrylamide gels were immersed in a 0.1% Coomassie R-250 solution in distilled water. The gel was soaked for about 20min. on a Denley mechanical rocker before being destained. A mixture of 10% methanol:10% acetic acid comprised the destain solution. The gel was left in destain until all excess Coomassie stain had been washed from the gel. Fresh destain solution was routinely added at least twice during the procedure. To remove the last traces of stain, a few beads of Amberlite mixed-bed resin were added to the destaining gel.

2.9.5.2 Silver Staining. The procedure described by Morrissey (1981) was used.

2.9.6 Drying Polyacrylamide Gels Polyacrylamide protein gels, once having been destained and photographed were dried down. The gel was soaked in a solution comprising 45% ethanol: 10 acetic acid:1% glycerol for a period of 30min. Gels were routinely dried onto sheets of cellophane, between double layers of Whatman 3MM paper at 80°C under a constant vacuum for 1 hour.

2.9.7 Molecular Weight Markers

Laemmli Gels:	Rabbit Muscle Myosin	205,000 Da
	ß-galactosidase	116,000 Da
	Phosphorylase B	97,400 Da
	Bovine Albumin	66,000 Da
	Egg Albumin	45,000 Da
	Carbonic anhydrase	29,000 Da
Native Gels:	α-lactalbumin	14,200 Da
	Carbonic anhydrase	29,000 Da
	Chicken Egg Albumin	45,000 Da
	Bovine serum albumin (Monomer)	66,000 Da
	(Dimer)	132,000 Da
	Jack Bean Urease (Trimer)	272,000 Da
	(Hexamer)	545,000 Da

2.9.8 Cross-linking of Protein Subunits Using Dimethylsuberimidate Cross-linking involves the formation of inter-molecular covalent bonds between adjacent protein monomers in an oligomeric protein complex to form a stable structure which is visible when subjected to SDS-polyacrylamide gel electrophoresis. Inter-molecular cross-linking agents include simple bis-imido esters which can be further sub-divided into two types. Those that may be cleaved by chemical modification and those which are stable once formed. Cross-linking using dimethylsuberimidate was performed. A modified method to that used by Lumsden and Coggins (1977) was used. Pure protein was dialysed to

0.1M Triethanolamine.Cl pH8.0 with at least 2 changes. It should be noted that Tris.Cl buffer cannot be used in cross-linking experiments. To 50µl of repressor solution (37.5µg) was added:

10µl	1M Triethanolamine.Cl pH8 (100mM)
10µl	100mM Dimethylsuberimidate.4HCl (10mM)
1µl	5M NaCl (50mM)
24µl	distilled water

This resulted in a solution in which the concentration of repressor was 375µg.ml^{-1}. The solution was gently mixed and left at room temperature for 1 hour for cross-links to form. The cross-linking was stopped by the addition of 5µl 1M NH_4HCO_3 (50mM). Half the sample of cross-linked protein (50µl) was then either loaded onto a Laemmli gel or on occa sions electrophoresed on native continuous 50mM NaPi pH7.0 polyacrylamide gels.

2.10.0 Production of Anti-Repressor Anti-serum Antibody against the ØC31 repressor proteins was produced in New Zealand white rabbits at the University of Stirling. All manipulations with animals were carried out by a suitably qualified technician. A solution of pure 42kDa repressor (1mg) was mixed with 2 volumes of Freunds incomplete adjuvant (Pierce Chemical Company) in a total volume of 150µl and injected intra-peritoneally into the rabbit. A test bleed was taken, centrifuged (5,000rpm, 20min.) and stored at -20°C. Following 6 weeks, a booster injection was given comprising a further 1mg. of purified repressor in 2 volumes of Freunds incomplete adjuvant. A test bleed was taken and treated as above. After 3 weeks the rabbit was bled and the serum purified, divided into 20µl aliquots and stored at -20°C.

2.11.0 Western Blotting Towbin *et al*, (1979)

2.11.1 Transfer of Proteins to a Membrane Support Protein gels were blotted onto nitrocellulose with a pore diameter of 0.45µm (Schleicher and Schuell BA85). The blotting procedure carried out was essentially that described by Towbin, Staehelin and Gordon, 1979. At the end of electrophoresis, the separating gel was removed from the stacking gel and immersed in blot electrode buffer (BEB; see section 2.5) and left agitating for 5min. A section of nitrocellulose, slightly larger than the gel to be blotted was cut and soaked in distilled water for 5min. The blotting sandwich is then assembled in the following manner immersed in a BEB solution. Onto a perspex support is placed a "Scotch-Brite" sponge pad, on top of which is then layered 3 sheets of Whatman 3MM filter paper. The nitrocellulose is placed carefully on the 3MM filter paper followed by the gel. Care must be taken to avoid air bubbles. On the gel is placed 3 more sheets of Whatman 3MM and another "Scotch Brite" sponge pad. The complete assembly is

enclosed in a plastic support which is then placed with an electro-blotting apparatus. Care must be taken to orient the assembly such that the nitrocellulose membrane is toward the anode.A Bio-Rad 'Mini-Protean' blotter or a Bio-Rad large capacity blotter was used to transfer the protein onto the membrane. The 'Mini-Protean' system was electrophoresed at 150V for about 90min. and the large blotting apparatus at 0.1A overnight at 4°C. To check that transfer of the proteins was successful, the membrane was usually stained with Ponceau S (Sigma) and the gel stained with Coomassie-R250.

2.11.2 Incubation of Western Blot with Antibody (Harlow and Lane, 1988)
The nitrocellulose filter onto which proteins had been transferred was placed in a clean tray, larger than the nitrocellulose sheet, so to allow free movement on a Denley Rocker. The blot was incubated in blocking buffer for not less than 3 hours with constant agitation. If longer incubation times were used, NaN_3 was added to a final concentration of 0.05% to inhibit microbial contamination. The blot was rinsed briefly with blocking buffer. The blot was incubated with primary antibody (section 2.10.0) in a volume of about 40ml. such that a 1/1000th dilution of anti-serum was used. This was left agitating for 60min. The blot was washed in blocking buffer for three periods of 5min. A 40ml solution containing a 1/2000th of the secondary antibody (goat anti-rabbit IgG-HRP conjugate) was added to the blot and left to agitate for a further 60min. The blot was again stripped with three 5min. washes with blocking buffer, plus a final wash with PBS. The blot was developed by incubating with a 0.05% solution of Diaminobenzidine.4HCl in 50mM Tris.Cl pH7.4. For every 1ml. of DAB.HCl was added 1ul of a 30% H_2O_2 solution (Sigma). Once the blot had developed to a sufficient intensity, the reaction was stopped by the addition of 20% SDS followed by several washes in distilled water. The blot was air-dryed on Whatman 3MM and stored between two sheets of Whatman 3MM.

2.12.0 N-terminal Protein Sequence Analysis N-terminal protein sequencing according to the Edman degradation method was carried out at the SERC Protein Sequencing facility at Aberdeen University by Mr Bryan Dunbar and Professor J. Fothergill on an Applied Biosystems 4077A pulsed-liquid phase automatic sequencer.

2.12.1 Preparation of the Protein Sample for Sequencing (Matsudaira, 1987; Owen and Chapman, 1991).
It is important that all apparatus and solutions used in this procedure are free even of trace quantities of glycine as this can seriously debilitate any attempt to sequence the protein. Enough protein was used such that the protein band of interest contained at least 100pmoles.

2.12.1.1 Electrophoresis of the Protein Sample The Bio-Rad Mini-Protean gel system was used in all blotting procedures.A tricine-SDS polyacrylamide gel prepared as described in section 2.5 was pre-run for 1 hour at a constant current of 6mA with an anode buffer containing 600µl of a glutathione solution (10mM, stored at -20°C). The glutathione is a reducing agent and will eradicate any free-radicals in the polyacrylamide gel which may react with the protein sample and prevent sequence analysis. When glutathione was not available, 600µl of a 100mM DTT solution was used. The samples were prepared by mixing with 1/10th volume loading buffer and boiling for 5min. The pre-run anode buffer was removed and replaced with fresh buffer containing 120µl of sodium thioglycollate (100mM). The gel was electrophoresed at a constant 100V until the dye reached the bottom of the gel. A 10% (29:1 acrylamide:bis-acrylamide) gel was used for all N-terminal sequencing gels.

2.12.1.2. Electroblotting of Protein Samples onto Pro-Blott^{TM} Membrane Pro-Blott^{TM} membrane provides a solid support onto which proteins can be electrophoresed. It is designed such that proteins stack-up on its surface and never enter the membrane. Thus, the protein to be sequenced is readily accessible when subjected to Edman degradation. The actual electroblotting procedure is as follows:

10x Stock Buffer:	100mM CAPS (3-[cyclohexylaminol]-1-10x propanesulphonic acid). 22.13g CAPS was dissolved in900ml distilled water. The pH was adjusted to pH11.0 with 2N NaOH, then adjusted to 1 litre with distilled water.
Blotting Buffer:	10mM CAPS pH11.0 10% Methanol

Following electrophoresis, soak the polyacrylamide gel in 100ml blotting buffer for 5min. Equilibrate the Pro-Blott^{TM} membrane; soak in 100% methanol for a few minutes before transferring to blotting buffer and leaving for 5min. Cut 6 strips of Whatman 3MM and pre-soak in blotting buffer. Assemble the blotting sandwich in the following order immersed in blotting buffer. Within the plastic holder, place a sponge. On top of the sponge place 3 layers of pre-soaked Whatman 3MM, then the pre-treated Pro-Blott^{TM}. Overlay the membrane with the pre-soaked gel, excluding all air bubbles. Cover the gel with 3 more sheets of Whatman 3MM, a sponge and encase the complete sandwich in the plastic case. Orientate the cassette such that the membrane is toward the anode. Electroblotting took place at 250mA for 60min. Following transfer, the blot was stained with Ponceau S and the band of interest excised. The polyacrylamide gel was stained with Coomassie R-250 to ensure complete transfer. Up to six lanes from any one blot were excised and sent for sequence analysis.

2.13 Purification of the Repressor Proteins

2.13.1 Construction and Maintenance of Chromatographic Columns Both the DEAE-Sephacel and Heparin-agarose steps in the purification of the repressor proteins made use of 'home-made' columns. The barrel of a 10ml syringe in which the barrel had been removed was plugged with glass-wool. To the syringe was then added 10ml of the chromatographic medium. The column was then washed with at least 3 volumes of Buffer A (section 2.13.2). Following a chromatographic separation, the columns were cleaned with the following:

DEAE-Sephacel:	20mM Tris.Cl pH8.0 2000mM NaCl 1mM EDTA
Heparin-agarose:	20mM Tris.Cl pH8.0 1.5M NaCl 8.0M Urea

For long term storage columns were left at 4°C in 20mM Tris.Cl pH8.0 buffer with 0.05% NaN_3.

2.13.2 Buffers Used During Purification of the Repressor Proteins

Buffer C: (Cracking Buffer)	Buffer A: (Normal Buffer)
20mM Tris.Cl pH8.0	20mM Tris.Cl pH8.0
200mM NaCl	5mM $MgCl_2$
5mM $MgCl_2$	1mM EDTA
1mM EDTA	0.1mM DTT
0.1mM DTT	1mM Benzamidine
1mM PMSF	

.c.2.13.3 Purification of the Putative 74, 54 and 42kDa Repressors The *E.coli* strain DS941 containing the construct p19H3S2.2 was grown up overnight after being inoculated with 1ml of an overnight culture in a volume of 4 litres of Luria broth containing 200μgml^{-1} Ampicillin as selection for the plasmid shaking at 37°C.

The bacterial pellet was resuspended in 20ml of Buffer C and passaged through a French Press three times at 950psi. To the crude lysate was added a few grains of DNaseI and left for 30mins on ice to allow the nuclease to completely digest the nucleic acid. The lysate was centrifuged at low speed (42,000xg, 4C, 30mins) in a JA20 rotor in order to remove high molecular weight cellular debris from the protein preparation. The supernatant was removed and placed in SW50 ultracentrifugation tubes. The material was then centrifuged at high speed (100,000xg, 4C,150mins) in an SW50 ultracentrifuge rotor to remove low molecular cellular debris. Once the supernatant had been cleared of the majority of particulate material, the volume was measured.

To the known volume of crude protein solution (22ml) was added solid ammonium sulphate (6.9g), stirring continuously at 4C, until a state where 50% saturation had been achieved. Once all the ammonium sulphate had been added the solution was mixed for a

further 50 minutes, before the protein precipitate was removed by centrifugation (6300xg,4°C,30 mins) in a 30ml glass corex tube. The supernatant was collected and placed in dialysis tubing. The salt was removed through dialysis to 2 litres of Buffer A with two changes.

The sample was split into two halves. One half was stored in the fridge at 4C and the other half was used for the remainder of the purification. From the ammonium sulphate step, the 40ml of protein solution was loaded onto a 10ml heparin-agarose column. The column was constructed as described in section 2.13.1. Once loaded, the protein was eluted using a step-gradient of increasing NaCl in Buffer A (section 2.13.2). Fractions from the heparin-agarose column were analyzed by SDS-PAGE. Fractions containing the repressor proteins were pooled and dialysed to 2 litres of Buffer A (at least 2 changes) to remove all salt. The dialysate was then applied to a Pharmacia MonoQ HR5/5 column on a Waters Advanced Purification System. A linear gradient of increasing NaCl from Buffer A to Buffer A + 1M NaCl was run at a flow-rate of $1ml.min^{-1}$. From this, 1ml fractions were collected and the absorbance at 280nm wavelength measured usinga Waters on-line detector attached to a chart recorder. Protein containing fractions were analysed using SDS-PAGE and visualization by Coomassie staining. Fractions containing the repressor proteins were pooled and treated as described in section 2.15.

2.13.4 Purification of the 42kDa Repressor Protein A culture of *E.coli* expressing the pMS221 construct was grown up in 1.2l of L-broth containing 200µg/ml of Ampicillin overnight. The culture was harvested by centrifugation in 250ml polypropylene beakers at 12,482xg at 4°C for 15 minutes using a JA14 rotor.

The bacterial cell pellet was resuspended in 20ml of Buffer C. The cell suspension was lysed in a French Press using two passes at 950psi (6555kPa). Insoluble cell debris was removed by a low speed centrifugation (4°C, 42,000 x *g*, 60 min) in a JA20 rotor. The volume of the cleared lysate was measured and solid ammonium sulphate was added to achieve 50% saturation of the protein solution. The solution was stirred for 30 minutes and the precipitate then collected by centrifugation (13,000xg 4°C, 30mins) in a JA20 rotor.

The supernatant from the ammonium sulphate precipitation was retained and dialyzed against 2 litres of Buffer A for at least 2 hours. This was repeated twice more with fresh Buffer A. Once all the ammonium sulphate had been removed, the protein containing solution (32ml) was loaded onto a DEAE-Sephacel anion-exchange column. The DEAE-Sephacel column contained 10ml of the anion-exchange matrix and was attached to a manual gradient-former and peristaltic pump in a cold-room. The column was pre-equilibrated by washing with 2 volumes of Buffer A, follwed by 1 volume of Buffer

A+1000mM NaCl and finally with 2 volumes of Buffer A alone. Fractions from the DEAE-Sephacel column were initially measured for protein using the Bradford assay. Individual fractions were then analysed by SDS-PAGE and staining to visualize the proteins to known standards. The activity was followed using band-shift assay (section 2.15.1. Fractions containing activity were pooled and loaded onto a pre-made 10ml heparin-agarose column pre-equilibrated with 3 column volumes of Buffer A. The DEAE-Sephacel fractions usually eluted in about 200mM NaCl and could be applied directly to the heparin-agarose column. Fraction were eluted using a step-gradient of Buffer A+NaCl in 100mM increments to 600mM and finally washed with a one volume of Buffer A+2000mM NaCl. Fractions were again visualized on Laemmli gels and activity measured using the band shift assay. At this stage the pooled 42kDa protein was greater then 95% pure.

2.14 Storage of the Purified Protein The purified proteins were stored at -20°C after diluting with 100% glycerol such that the final glycerol concentration was 50%.

2.15 Assays for DNA Binding Activity Primarily, only two methods were used to detect DNA binding activity. The bandshift assay first described by Fried and Crothers (1981) and Garner and Revzin (1981) was used as a fairly gross, but sensitive test for DNA binding activity. The DNaseI footprinting method (Galas and Schmitz 1978) was used to provide more detailed information about the precise location of DNA binding.

2.15.1 Band-shift Assay The band-shift assay provides a relatively easy and reliable way in which to assay the binding of a protein to a specific fragment of DNA. The method relies upon a shift in the molecular weight of a protein:DNA complex to free DNA and thus an observed difference in mobility when electrophoresed on a native polyacrylamide gel. A modified buffer for binding reactions was used composed of:

10x DNA Binding Buffer:	200mM NaPi pH6.6
	500mM NaCl
	10mM EDTA
	1.0mM DTT
	5% Glycerol
	100ugml^{-1} BSA
10x Running Buffer:	300mM NaPi pH6.6
Native Gel Composition:	4% Polyacrylamide (19:1 acrylamide:bisacrylamide)
	30mM NaPi pH6.6

A typical binding reaction contained about 1ng of end-labelled DNA (less than 300bp in length), produced by the PCR (section 2.08.3). Non-specific competitor DNA in the form of pUC18 plasmid or poly(dI.dC)(dI.dC) was used at a concentration at least 200-fold greater then that of the labelled probe. Bovine serum albumin was added to a final

concentration of 100ug.ml^{-1} to aid the formation and stability of the DNA:protein complex (Kozmik, Urbanek and Paces, 1990). Reactions were incubated for 15min. at room temperature, quenched on ice for 3min. and then loaded onto a pre-equilibrated native gel. Native polyacrylamide gels were pre-run at a constant 50V for at least 30min. before the samples were loaded. Once loaded the samples were electrophoresed at 75V for at least 3 hours. Gels were dried down onto Whatman 3MM and subjected to autoradiography.

2.15.2 DNaseI Footprint Analysis of DNA:Protein Complexes DNaseI footprinting was carried out essentially as originally described by Galas and Schmitz (1979). Binding reactions were carried out in a volume of 20μl, contained 17.5μg.ml^{-1} of pUC19 plasmid DNA and approximately 1ng of labelled fragment. Binding buffer described in section 2.15.1 was used to incubate protein with DNA except that $CaCl_2$ to a concentration of 5mM was also added. Binding reactions were performed at room temperature for 15 minutes. DNaseI (Pharmacia fplc pure DNaseI) was added as a diluted 4μl solution containing 15U (Pharmacia units) of activity and allowed to incubate with the protein:DNA mixture for 30 seconds. The nuclease reactions was terminated with the addition of 100μl phenol and rapid mixing. To each reaction was added 180μl (1x)TE, mixed, and extracted from the organic phase. The DNA was precipitated by the addition of 400μl ethanol and 20μl 3M NaAcetate and recovered by centrifugation. DNA pellets were resuspended in 2.5μl Sequenase™ STOP buffer and heated to 90°C for 3 minutes before loading onto a denaturing urea-polyacrylamide sequencing gel. Eppendorfs were siliconized before use according to the method described in Sambrook *et al*, 1987.

2.15.3 Maxam and Gilbert Chemical Sequence Analysis Maxam and Gilbert (1977) The reactions were performed with labelled DNA linked to Amersham M+G™ paper according to the protocol supplied by the manufacturers.

Chapter 3

Overexpression of the Repressor Proteins from ØC31 in *E.coli*

3.1 Introduction

Overexpression of the repressor proteins encoded by the ØC31 *c* gene locus was required to facilitate their purification. The overexpression of the repressor proteins was attempted by cloning the repressor gene into the polylinker of pUC19 in front of the *lac* promoter in *E.coli* and examining the protein content of crude lysates for overexpressed proteins unique to the *c* gene.

3.2 Overexpression of the Repressor Proteins from ØC31 in *E.coli*

A 2.7kb *Hin*dIII- *Sph*I fragment was cloned into pUC19, using *Bam*HI linkers, in both orientations with respect to the *lac* promoter to give pMS201 and pMS202(M.C.M Smith, personal communication). The orientation of each of the *Hin*dIII-*Sph*I clones; pMS201 and pMS202 were determined by digestion with *Hin*dIII (Figure 3.1.a). This fragment contains the complete repressor ORF and a short upstream sequence of 95bp from the ATG translation start (Smith and Owen, appendix). Crude lysates of plasmid-containing *E.coli* cultures following induction with IPTG and were analysed by SDS-PAGE. Overexpression of at least three proteins was observed with molecular weights of 110,68 and 50kDa, the appearance of which seems to increase with time and which were not present in *E.coli* pUC19 extracts (Figure 3.1.b.). The observed molecular weights of the overexpressed proteins did not agree well with any of the predicted sizes (74,54 and 42kDa).

Expression of the repressor proteins from pMS201 and pMS202 was found to be independent of the *lac* promoter on pUC19 and was not affected by IPTG induction (Figure 3.1b). The expression of the repressor proteins increases to a very high level, estimated to comprise between 10-20% of the total cellular protein.

3.3 Introduction of Frameshift Mutations into the Repressor Gene to Identify Repressor Proteins

To demonstrate the ORF responsible for the overexpressed proteins observed on SDS-PAGE from crude lysates, frameshift mutations were introduced at unique restriction sites within the *c* gene (pMS201). A mutation was introduced first at the *Nco*I site which overlaps the putative ATG initiation codon for the predicted 54kDa protein. The mutation was carried out by *Nco*I digestion followed by filling-in the 5' overhang with the Klenow

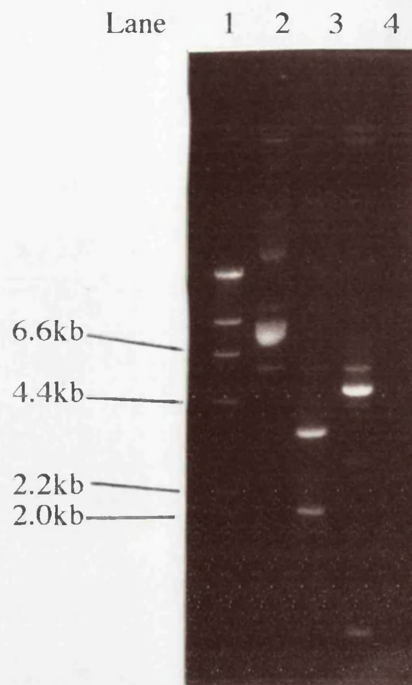


Figure 3.1.a. Orientation of pMS201 and pMS202.
Both pMS201 and pMS202 were digested with 10U *Hin*dIII and electrophoresed on a 0.8% agarose gel in (1x) TAE. Gels were stained in ethidium bromide and visualized on a short wave transilluminator.

Lane	Plasmid		
1	λHindIII Markers	3	pMS201- *Hin*dIII digested
2	pMS201 Uncut	4	pMS202- " "

MWt *Hin*dIII Digested Fragments

Plasmid	Fragment	Predicted Size	Observed Size
pMS201	1	3.518	3.47
	2	1.926	1.91 (by extrapolation)
pMS202	1	4.626	5.01
	2	0.818	0.74 (by extrapolation)

Figure 3.2. Nucleotide Sequence of Frameshift Mutations

	---*Nco*I---	Plasmid
Wild-type *Nco*I Sequence	5'.....CACCATGGCCAC...........3'	
*Nco*I Fill-in Frameshift	5'.....CACCATGCATGGCCAC...3'	pCO100
*Nco*I Deletion Frameshift	5'.......CACCGCCAC.........3'	pCO101
	---*Bcl*I-----	
Wild-type *Bcl*I Sequence	5'......CCTGATCAAG.....3'	
*Bcl*I Fill-in Frameshift	5' CCTGATCGATCAAG....3'	pCO102

Lane 1 2 3 4 5 6 7 8 9 10 11 12 13

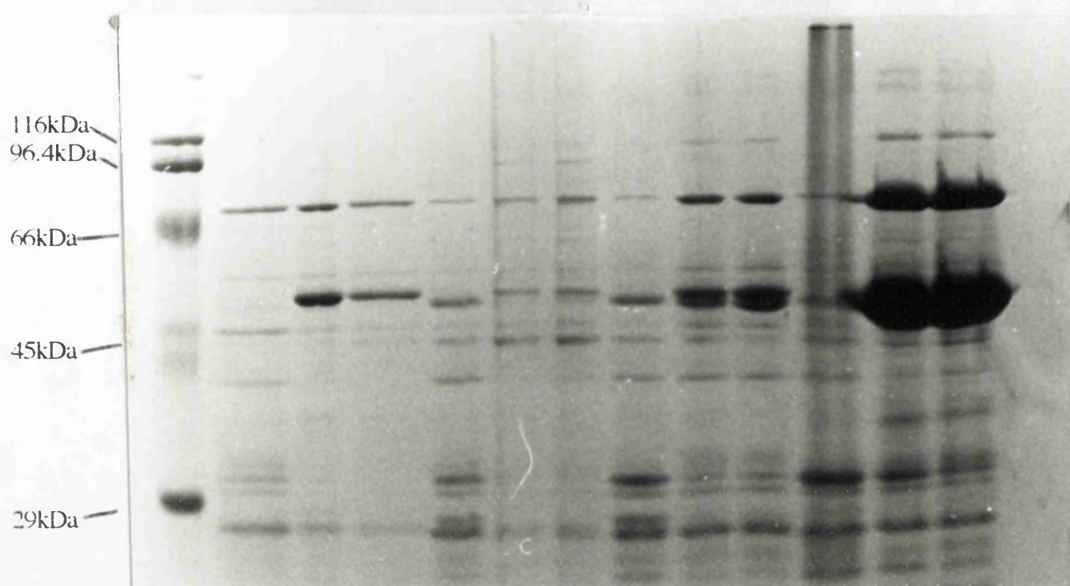


Figure 3.1.b. Overexpression of the Repressor Proteins in *E.coli*
Samples of *E.coli* DS941:pMS201 and DS941:pMS202 were grown up after sub-culturing from an overnight culture to an $O.D_{600}$ of 0.2 . IPTG was added to a 1mM final concentration. Samples of 1ml were taken at various time-points to analyze protein content on a 10% SDS-polyacrylamide gel.

Lane	Sample	Time of Harvest(hr)	Lane	Sample	Time of Harvest (hr
1	High Molecular Weight Markers		8	DS941:pUC19	4
2	DS941:pUC19	0	9	" pMS201	4
3	" pMS201	0	10	" pMS202	4
4	" pMS202	0	11	" pUC19	16
5	" pUC19	2	12	" pMS201	16
6	" pMS201	2	13	" pMS202	16
7	" pMS202	2			

fragment of DNA polymerase and re-ligation of the blunt-ends. The plasmid containing the *Nco*I frameshift was designated pCO100. The mutation disrupted the open reading frame for the predicted 74kDa repressor and provided a second out-of-frame ATG just downstream of the correct putative initiation for the 54kDa protein (Figure 3.2). Digests to confirm the frameshift mutations in each of the three constructs are demonstrated in Figure 3.3. Crude lysates from DS941:pCO100 analysed by SDS-PAGE demonstrate a marked reduction in expression of a protein with an apparent molecular weight of 70kDa (Figure 3.4). A protein with an apparent molecular weight corresponding to 110kDa was no longer present in DS941:pCO100 (Figure 3.4). This observation suggests that the predicted 74kDa protein which should be completely absent in DS941:pCO100, has aberrant mobility (110kDa) during SDS-PAGE. A deletion mutant of the *Nco*I frameshift mutation was made in which mung bean nuclease was used to remove the 5' overhang after the restriction digest, followed by re-ligation of the blunt ends (Figure 3.2). The expression of the *E.coli* pCO101 confirmed an identical pattern of expression as the fill-in mutation (Figure 3.4).

A second target for the generation of frameshift mutation was a unique *Bcl*I restriction site which lies within the ORFs for all three predicted proteins (Figure 3.2). This mutation was constructed by filling-in the 5' sticky-ends with Klenow fragment after digestion with *Bcl*I and re-ligating the blunt ends. The plasmid containing the *Bcl*I frameshift was designated pCO102. Expression of such a frameshift mutant should result in no expression of the repressor proteins and this was confirmed after analysis of extracts from *E.coli* pCO102 by SDS-PAGE (data not shown).

3.4 Expression of Repressor Proteins From Plasmids Containing Alternative Subclones of the *c* Gene

DS941 carrying plasmid, pMS221, which encodes the *c* gene from the *Nco*I-*Sph*I restriction sites (Figure 1.2; MCM Smith, personal communication) produced only the 42kDa protein (Figure 3.4), whilst DS941:pMS231 (encodiing the *c* gene from the *Pvu*I-*Sph*I sites) produced both the 54 and 42kDa repressors. No expression of the repressor proteins was observed from DS941:pMS222 and pMS232 containing respectively the *Nco*I-*Sph*I and *Pvu*I-*Sph*I fragments in the opposite orientation to be read from the *lac* promoter.

The *Sph*I-G fragment cloned in pUC19 with the *c* gene in the same or opposite orientation to the *lac* promoter in *E.coli* expressed none or very little of the three repressor proteins (Figure 3.5). However, the truncated *Hin*dIII-*Sph*I-G fragment cloned in *E. coli* pMS201 resulted in abundant expression of all three repressor proteins independent of the *lac* promoter (Figure 3.1b). One explanation of this difference is that a regulatory region

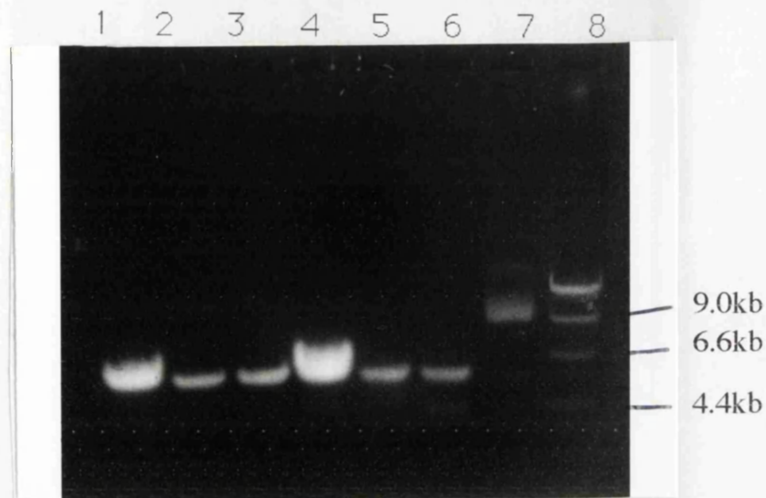


Figure 3.3. Construction of Frameshift Mutations.
Each plasmid was digested with for 1 hour with the indicated restriction enzyme at the optimum temperature. A 0.8% agarose gel in 1xTAE was run using the BRL 'Mini" system.

Lane	Plasmid	Enzyme
1	pMS201	*Bcl*I
2	pCO102	*Eco*RI+*Bcl*I
3	pCO100	*Eco*RI
4	pMS201	*Nco*I
5	pCO100	*Eco*RI+*Nco*I
6	pCO101	*Eco*RI+*Nco*I
7	pMS201	Uncut
8	High Molecular Weight Std	

Lane	Explanation
1	Control- *Bcl*I digests DNA (4.8kb)
2	A single band only (4.8kb) if a mutation exists at *Bcl*I site
3	Control- *Eco*RI digests plasmid DNA (4.8kb)
4	Control- *Nco*I digests DNA
5	Fill-in *Nco*I mutation- single band of 4.8kb due to *Eco*RI
6	Deletion *Nco*I mutation- single band of 4.8kb due to *Eco*RI
7	Uncut control

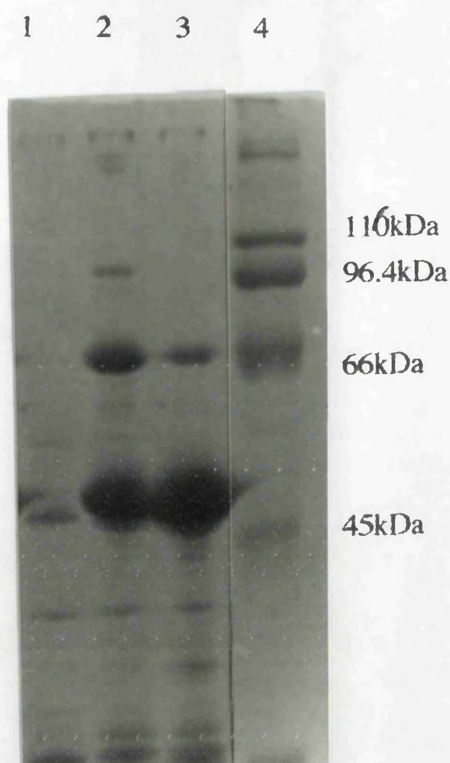
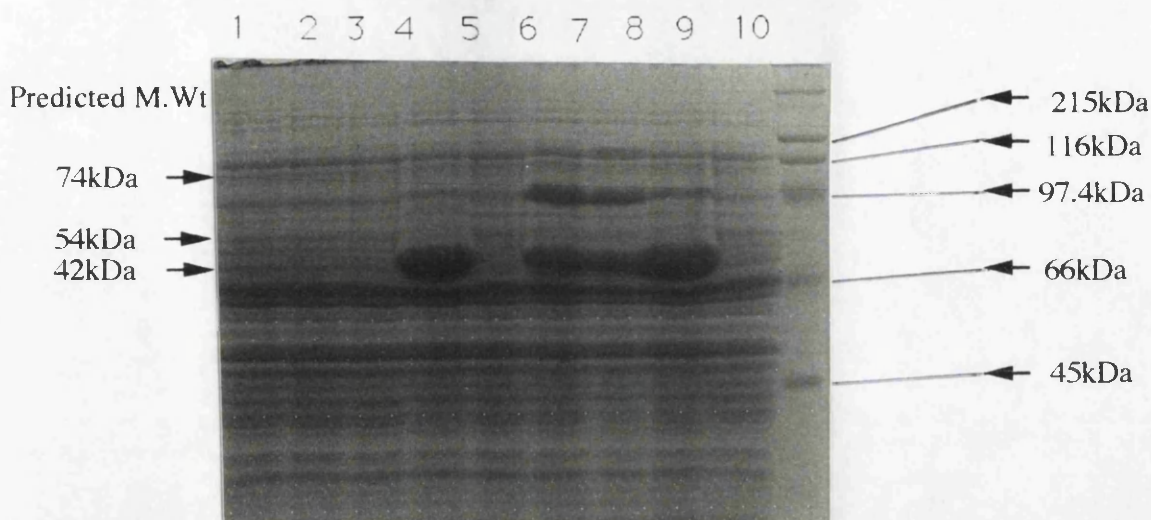


Figure 3.4. Expression of the Repressor Gene Containing Frameshift Mutations at the Unique *Nco*I Restriction Site. A frameshift mutation was introduced at the *Nco* I site within the *Hind* III-*Sph* I fragment using Klenow Pol..

Lane 1- *E.coli* DS941:pUC19
Lane 2- *E.coli* DS941:pMS201
Lane 3- *E.coli* DS941:pCO100
Lane 4- High Molecular Weight Markers.

Figure 3.5. Crude Lysates Analysed by SDS-PAGE to Demonstrate Expression from N-terminally-Deleted Sub-clones of the Repressor Gene *c* in *E. coli*



E. coli cultures containing their respective plasmids were incubated overnight in L broth at 37 C, with shaking and 200μg/ml ampicillin for plasmid maintenance. Samples were electrophoresed on a 10% SDS-Polyacrylamide gel and then stained in Coomassie Blue G-250. Samples were loaded in the following order:-

	Expression from p*lac*
Lane 1 -*E.coli* DS941:pMS212	-
Lane 2 -*E.coli* DS941:pMS211	+
Lane 3 -*E.coli* DS941:pMS222	-
Lane 4 -*E.coli* DS941:pMS221	+
Lane 5 -*E.coli* DS941:pMS232	-
Lane 6 -*E.coli* DS941:pMS231	+
Lane 7 -*E.coli* DS941:pMS201	
Lane 8 -*E.coli* DS941:pCO100	
Lane 9 -*E.coli* DS941:pUC19	
Lane 10 - High Molecular Weight Protein Standard	

upstream of the *Hin*dIII site has been deleted. To help confirm this, plasmids pCO104 and pCO106 were constructed with variable lengths of DNA upstream of the *Hin*dIII site (Figure 3.6).

A 4.3kb *Eco*RV-*Sph*I fragment and a 2.8kb *Nhe*I-*Sph*I fragment were inserted into pUC19 to form pCO105 &106 and pCO103 &104 in an orientation in which their expression was independent of the *lac* promoter, and analysed for the expression of protein (Figures 3.6 and 3.7). The *Hin*dIII-*Sph*I clone (pMS201) expressed all three repressor proteins in this orientation (Figure 3.1). Expression of the *c* gene is greatly reduced in DS941:pCO104 and DS941:pCO106 (Figure 3.5). Thus, the region responsible for regulating expression of the *c* locus must be located downstream of the *Nhe*I restriction site and upstream of the *Hin*dIII site.

3.6 Discussion and Conclusions

The data presented in this chapter demonstrated overexpression of 3 repressor proteins from the *c* locus of ØC31 in *E.coli*. These observations apparently reflect expression of the *c* gene in *Streptomyces* lysogens in which three transcripts are evident from the *c* locus each of which could be translated to produce proteins of 74,54 and 42kDa (Smith and Owen,1991). The expression of the 74, 54 and 42kDa repressors was confirmed in *Streptomyces* by raising antibody to the 42kDa repressor and using this in western blot analyses of streptomycete ØC31 lysogens (Smith and Owen, appendix; Chapter 5).

The expression of the repressor proteins from DS941:pMS201 was independent of orientation with respect to the *lac* promoter suggesting that at least one of the promoters within the repressor gene is active in *E.coli*. It is estimated that only 10% of streptomycete promoters are active in *E.coli* (Bibb and Cohen, 1982; Schottel *et al*, 1981). The criteria that pre-determine whether a streptomycete promoter is active or not in the heterologous host is not known, but is possibly dependent on evolutionary conservation between shared σ factor homologues. A number of *Streptomyces* promoters have been identified which bear a good resemblance to the well-characterized *E.coli* σ^{70} promoter at the -35 and -10 positions but sufficiently different to derive a streptomycete consensus promoter (Strohl, 1992; Baltz and Seno, 1989). The *c*P1 promoter of the ØC31 repressor contains a -10 sequence (TAGGGT) very similar to the *Streptomyces* consensus -10 sequence (TAGGAT), as do *c*P2 (TAGAGT) and *c*P3 (TGAGGT), but none possesses any recognizable -35 sequence (Sinclair and Bibb, 1988; Smith and Owen, appendix).

Data presented in this chapter implies that the two predicted internal promoters, *c*P2 and *c*P3, within the *c* gene do not function in *E.coli*. The consensus *E.coli* -10 sequence (TATAAT) bears some homology with all three repressor gene promoters. The reason why only *c*P1 appears to function in *E.coli* is unclear, although at the -10 sequence level, differences are obvious between *c*P1, *c*P2 and *c*P3. Other differences may lie in sequences upstream of the -10 region which may act in repressing / activating transcription from the various promoters. A direct comparison of the two promoters *c*p1 and *c*p2 at the primary sequence level reveals little difference in the -10 region (Figure 3.4). The promoter designated *c*p3, although possessing a -10 region very similar to *c*p1 and *c*p2 is not very active when tested *in vivo* in *S.lividans* using *c*p3-*xyl* E reporter fusions and possibly requires activation by a phage encoded regulator (Smith and Owen, appendix).

Some regulation of expression of the repressor locus may take place at the translational level. The presumed ribosome binding site for all three transcripts is highly conserved consisting of AAGGG situated between 7-10bp upstream of the ATG initiation codon (Smith and Owen, appendix). In *E.coli*, all three repressor transcripts are presumably produced from either *c*P1 or p*lac*. It is curious therefore why apparently little of the 74kDa repressor is produced, in preference for the 54 and 42kDa proteins. Transcription mapping of the 5' end-points of *c* mRNA's in streptomycete lysogens has provided possibly a clue. Two transcription start-points are found for *c*P1(P1 and P1'), one of which (P1') is downstream of the translation initiation signal (Smith and Owen, appendix). It is believed possible that transcripts initiating at P1 may be processed to the shorter P1' transcript so inhibiting translation of the 74kDa protein. Whether the same processing event takes place in *E.coli* is debatable and requires further investigation.

Frameshift mutations were introduced at restriction sites within the repressor gene to disrupt and subsequently identify ORFs encoding the repressor proteins. Mutations at the *Nco*I restriction site, either by filling-in or deleting the site resulted in the complete loss of expression of a protein migrating with an apparent molecular weight of 110kDa by SDS-PAGE (Figure 3.3). In addition, there was a marked reduction in expression of a protein with an apparent molecular weight of 70kDa. A frameshift mutation at the unique *Bcl*I locus which is situated in all three predicted ORFs results in the loss of expression of all 3 putative repressor proteins. The simplest explanation for this is that the observed 110, 70 and 50kDa proteins correspond to the predicted 74, 54 and 42kDa proteins respectively. Clearly the mobilities of the repressor proteins when analyzed on SDS-polyacrylamide gels do not correlate with the expected monomeric molecular weight of each protein as calculated based on the DNA sequence (Sinclair and Bibb, 1988). There are many examples of proteins which migrate with anomalous molecular weight on SDS-

polyacrylamide gels and include the σ factor RpoD from *E.coli* (Burton *et al*, 1981; Helmann and Chamberlin, 1988) and the FtsY protein from *E.coli* (Gill and Salmond, 1990). The reason for this apparent anomaly is probably due to the fact that proteins which possess a high proportion of acidic residues probably bind fewer SDS anions than other less charged proteins resulting in an apparent increase in molecular weight (Weber and Osborn, 1973). Glutamate and aspartate comprise over 20% of the total amino acids of *E.coli* RpoD (Burton *et al*, 1981). Similarly, the ØC31 repressor proteins possess between 18 and 20% of their total amino acids as either aspartate or glutamate (Table 3.1). Generally, the proportion of amino acids does not significantly differ between the repressor proteins but some e.g. acidic and basic residues are different to percentages quoted for the "average" protein (Table 1.5). The predicted pI for each of the repressor proteins was determined using the University of Wisconsin Genetics Computer Group package (Devereux *et al*, 1984). From the predicted pI and the composition of amino acids, it is assumed that the overall charge on the repressor proteins is predominantly negative (Table 3.1). The acidic residues are evenly distributed throughout the proteins and do not form acidic domains thought to be important for the function of transcriptional activators in many eukaryotic proteins (Struhl, 1989) and in σ^{54} from *E.coli* (Sasse-Dwight and Gralla, 1990). Similar to the ØC31 repressor proteins, the tryptophan repressor from *E.coli* also possesses an overall negative charge with a predicted isoelectric point of 5.9 (Joachimiak *et al*, 1983).

Expression of the repressor proteins is reduced in the presence of a 70bp region upstream of *c*P1. Analysis of extracts of pIJ2410 by SDS-PAGE (Figure 3.5) demonstrate that very little of the repressor proteins are producedwhilst pMS201 produces 5-10% of total cell protein as repressor. These data imply that the repressor proteins possibly autoregulate their own expression, which is a common feature amongst phage repressors (Wulff *et al*, 1983). Within the *Hin*dIII and *Nhe*I region are a number of possible DNA-binding sites including a putative direct repeat and a terminator, t_k, from the upstream gene 'k' (Smith *et al*, 1992). Terminators in ØC31 are remarkably well conserved, not only in the repressor gene, but within the 'early' region, and in the *int/att* region (Smith *et al*, 1992). Indeed it is possible that the terminator, t_k, may play a critical role in the regulation of gene expression in the longer clones of upstream DNA from *c*P1. It is possible that expression from *c*P1 may have been derived from a plasmid-borne transcript, in which event, terminator t_k may have reduced this expression. Even in the wrong orientation for the *lac* promoter, some c expression may be read off plasmid transcripts, which the presence of t_k would abolish.

In summary, three repressor proteins from bacteriophage ØC31 have been overexpressed in *E.coli*. The open-reading frame for each of the repressor proteins have been identified

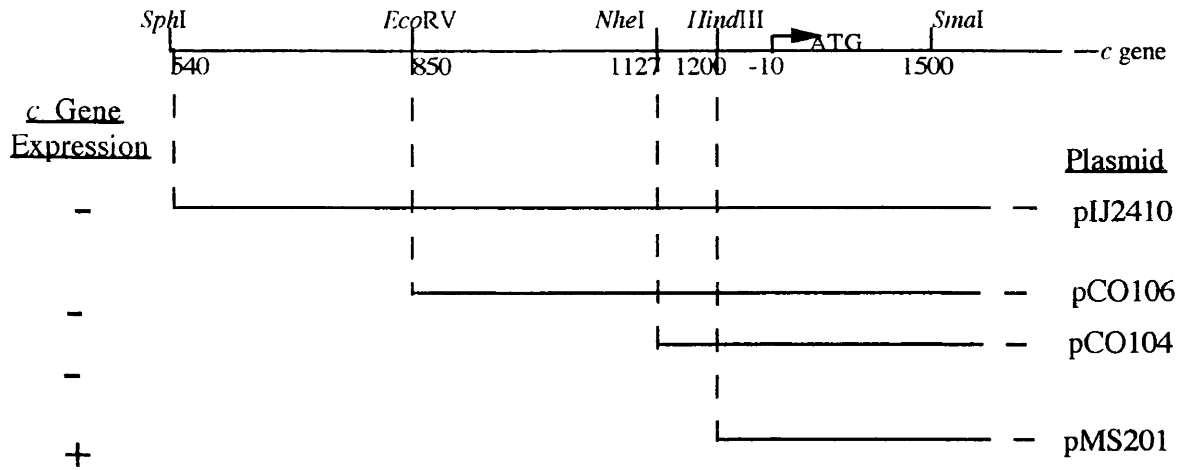


Figure 3.6. Restriction Map of DNA Upstream of *c*P1 Promoter

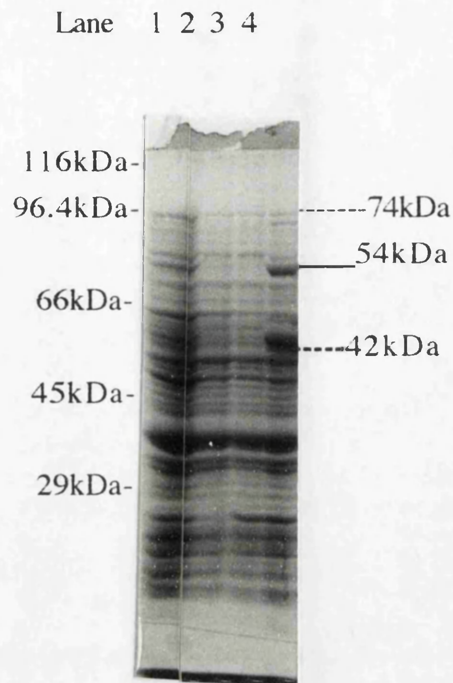


Figure 3.7. Protein Extracts from *E.coli* DS941 pIJ2410, pCO106, pCO104 and pMS201.
Crude lysates of 1ml. from overnight cultures were resuspended in 60μl SDS-loading buffer and loaded onto an 8% SDS-polyacrylamide gel.

Lane	Sample	Protein Loaded (μg)
1	*E.coli* DS941:pIJ2410	40
2	" :pCO106	30
3	" :pCO104	30
4	" :pMS201	30

Table 3.1: Predicted Isoelectric Points of the Repressor Proteins

Predicted Size	Predicted pI	Total Amino Acid
74kDa	4.93	683
54kDa	5.03	495
42kDa	4.93	384

The sequence is from Sinclair and Bibb (1988). Computer analysis was performed using the methods of Devereux *et al*, 1984.

by creating frameshift mutations within the *c* gene and by expressing N-terminally deleted sub-clones of the *c* gene. The repressor proteins migrate with aberrant mobility on SDS-PAGE, which is predicted to be a result of the relatively high proportion of acidic amino acid residues found within the repressor proteins. Finally, a regulatory sequence may be located downstream of a *Nhe*I restriction site, but upstream of the *Hin*dIII restriction site at the 5' end of the *c* gene.

Chapter 4

Development of a Purification Strategy for the ØC31 Repressor Proteins

4.1 Introduction

The previous chapter demonstrated the overexpression of all three repressor proteins in *E.coli.* The overexpression has enabled the purification of the repressor proteins to be attempted. This chapter illustrates the development of a purification protocol for the pool of repressor proteins away from *E.coli* proteins. Initial attempts to purify the three repressor proteins were made from extracts of DS941pMS201. An aim of this work was to study aspects such as solubility and oligomerization properties of the repressors and determine the N-terminal amino acid sequence for the 54 and 42kDa proteins. For these reasons, the purification of all three proteins was attempted, even though it was considered that the final separation of the individual proteins might be difficult ; very high resolution may be required to separate very similar molecules sharing partly the same amino-acid sequence .e.g. a dimer of the 42kDa protein has a molecular weight of 84kDa whilst a hetero-dimer of the 54 and 42kDa repressor proteins has a predicted size of 96kDa. A size difference of 12kDa is too small to resolve by gel filtration chromatography (Pharmacia plc Booklet on Gel Filtration Chromatography). However, the purification of the 42kDa protein in isolation was feasible using DS941:pMS221.

4.2 Theory of Separation Methods Used in the Purification of the ØC31 Repressor Proteins

4.2.1 Salt Fractionation

Fractionation of proteins by the step-wise addition of salt relies on differences in protein solubility (Scopes, 1985). The solubility of a protein is least, or the hydrophobicity greatest at its isoelectric point (pI), the point where the charge upon the protein is its least. Addition of salt reduces the interaction of hydrophobic amino groups with surrounding water molecules and simultaneously increases the interaction of charged amino acid residues with water. Therefore proteins with few charged amino acids will tend to aggregate and precipitate on salt addition due to the increasing number of protein-protein hydrophobic interactions; proteins with a large number of charged residues will have relatively few hydrophobic residues and remain in solution at relatively high salt concentrations, and vice-versa (Dreutscher, 1989). Ammonium sulphate is the preferred salt used in differential precipitation by 'salting out' because it is relatively inexpensive, possessing a high degree of purity, and is soluble at high concentration, even at low temperature.

4.2.2 Ion-Exchange Chromatography

Separation of proteins by ion-exchange chromatography relies on differences in charge at the protein surface. Initially, the sample is applied to the ion-exchange matrix and allowed to bind. The affinity of solute for matrix is determined by the pKa of the protein which can be altered by changing the pH of the surrounding mobile phase. Proteins with few charge interactions at the pH of the mobile phase remain unbound and are eluted. Hence ion-exchange is a very effective concentration step if the protein of interest binds to the matrix. Bound protein is eluted from the matrix by increasing ionic strength or pH which can contribute significantly to the high resolution of the procedure.

It was predicted, based on sequence data, that the large number of acidic residues in each repressor protein would possibly indicate a relatively large negative charge, thus suggesting that anion-exchange chromatography would be a suitable separation technique. Two different anion-exhange matrices were used in the purification of the repressor proteins, MonoQ and Diethylaminoethyl-sephacel (DEAE-Sephacel).

The MonoQ matrix, marketed by Pharmacia for use with the Fast Protein Liquid Chromatography system is used as a strong anion-exchanger and was used initially during the development of a purification scheme for the repressor proteins. The MonoQ matrix is composed of quarternary amine groups attached to Mono beads, developed to withstand the high flow rates used in FPLC without compromising resolution and is stable in a wide pH range 2-12.

DEAE-Sephacel is a weak anion-exchange matrix. The DEAE molecule is a qua ternary amine immobilised on beaded microcrystalline cellulose with a weak positive charge. It can be used for the separation of molecules up to 1,000kDa and has a relatively large binding capacity over a pH range of 2-9 (Pharmacia Plc. publication on 'Ion Exchange Chromatography-principles and methods').

4.2.3 Affinity Chromatography

Affinity chromatography provides a means of separation based on a predicted property, such as the ability to bind a co-factor e.g. NAD^+, ATP, or indeed to DNA. As the repressor proteins were expected to bind DNA, heparin-agarose was used as an affinity matrix.

Heparin is an analogue of DNA. It has a linear, polyanionic structure and possesses some ion-exchange properties (Bickle *et al,* 1977). DNA binding proteins generally bind to immobilised heparin with high affinity whilst non DNA-binding proteins do not bind to

the matrix. It was originally developed for the purification of blood coagulation factors and its specific mode of action as an affinity matrix for DNA binding proteins is unclear. During affinity binding of protein to DNA, the interactions are sensitive to salt concentration. Hence binding of the chosen molecules can be optimised by attention to the ionic environment (Record *et al,* 1981).

4.3 The Purification of the 74, 54 and 42kDa ØC31 Repressor Proteins

The purification protocol presented below details a procedure for the simultaneous purification of the 74, 54 and 42kDa repressor proteins away from *E.coli* proteins using extracts from DS941:pMS201. Salt fractionation and anion-exchange chromatography were used as initial purification steps based on the prediction that the repressor proteins had a highly negative charge (Section 1.7.2). During the purification of the three repressor proteins, the presence of each of the proteins was based solely on their appearance after SDS-PAGE. As the proteins were highly expressed, they were easily identified as the major protein at the expected molecular weight.

4.3.1 Procedure

The following procedure was developed for the purification of the ØC31 repressor proteins expressed by *E.coli*:DS941 pMS201 (Figure 4.1; Section 2.13.3).

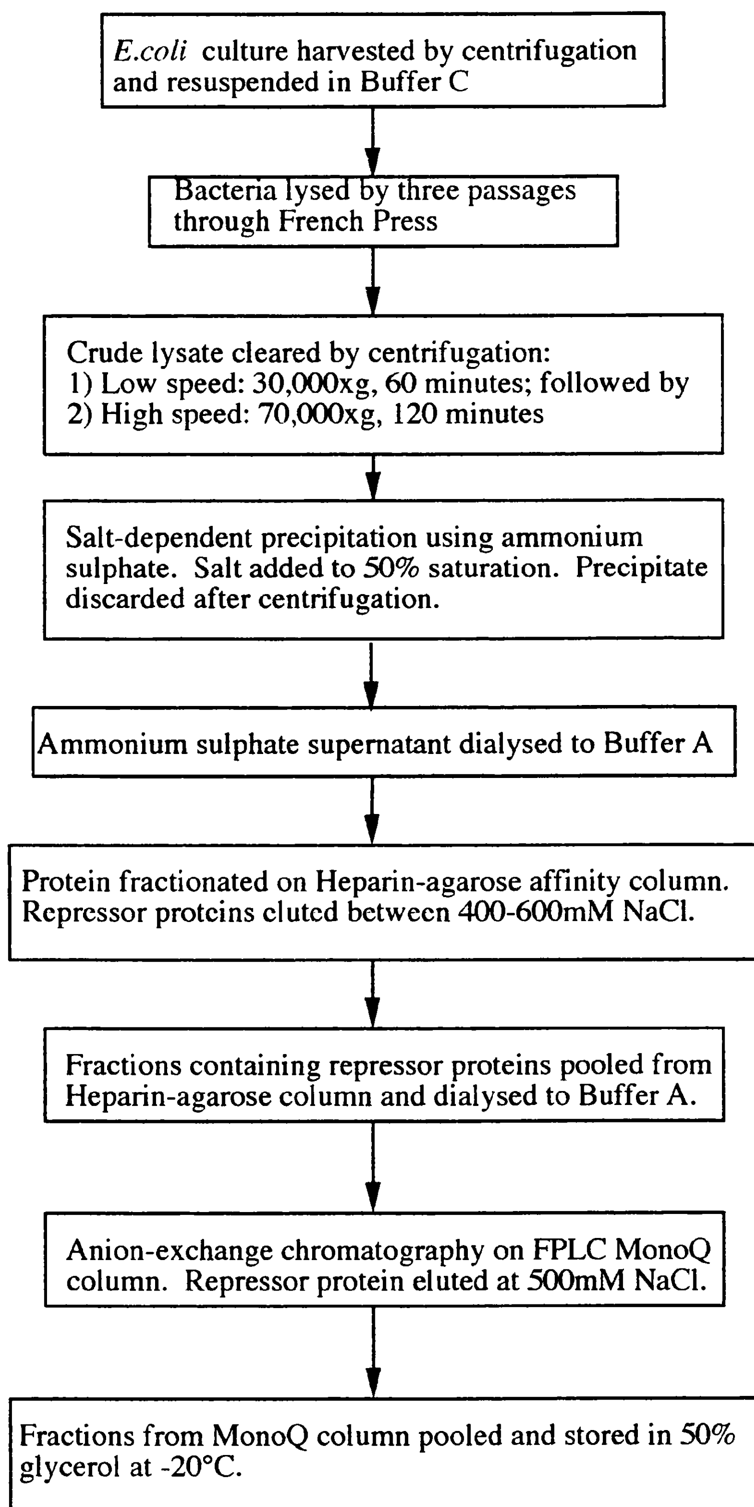
4.3.2 Salt Fractionation of the ØC31 Repressor Proteins

Ammonium sulphate was added to a cleared lysate of DS941:pMS201 to saturation levels of 0-30%, 30-50%, 50-60%, 60-75%, 75-90%. The majority of the putative repressor proteins are precipitated at an ammonium sulphate saturation level of between 50 and 75% (Figure 4.1). At ammonium sulphate concentrations greater than 50% saturation, it is common for proteins to float to the surface as a result of the high density of the salt solution (M.J. Chamberlin personal communication). To avoid this problem, the supernatant from a salt dependent precipitation at 50% $(NH_4)_2SO_4$ saturation was used for subsequent protein purification steps. Supernatant from the 50% ammonium sulphate stage was dialyzed overnight before loading onto a heparin-agarose column.

4.3.4 Separation of the ØC31 Repressor Proteins From Total Protein using Heparin-agarose Chromatography

The dialysate was loaded onto a heparin agarose column in low salt buffer A (Table 4.1). It was evident that the three repressor proteins bound to the column under the loading conditions of Buffer A based on the analysis of fractions from the column after SDS-PAGE (Figures 4.3 and 4.4). Two elution peaks are clearly visible in the plot of protein

Figure 4.1. Protocol for the Purification of the ØC31 74,54 and 42kDa Repressor Proteins



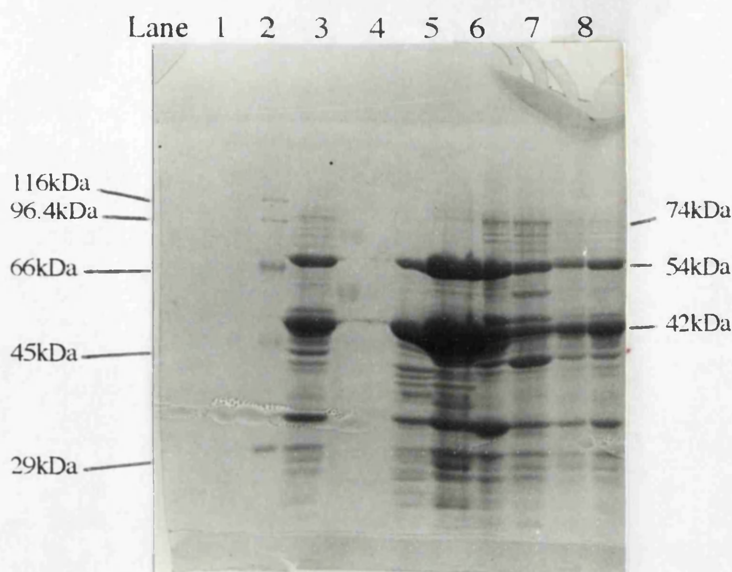


Figure 4.2. Ammonium Sulphate Fractionation of Proteins from DS941:pMS201. Solid ammonium sulphate was added to increasing levels of saturation to a solution containing the repressor proteins whilst being continuously stirred at 4°C. The precipitate at each saturation level was collected by centrifugation, the supernatant being retained to be brought to the next saturation level. Samples of fractions were loaded onto a 10% SDS-polyacrylamide gel.

Lane:	$(NH_4)_2SO_4$ Saturation	Volume Loaded (μl)
1	High Molecular Weight Standard	10
2	Crude Lysate	"
3	0-30% Saturation	"
4	30-50% Saturation	"
5	50-60% Saturation	"
6	60-75% Saturation	"
7	75-90% Saturation	"
8	90% Saturation Supernatant	"

Figure 4.3. Elution Profile of 74,54 and 42kDa Repressor Proteins Following Fractionation on a Heparin-agarose Column

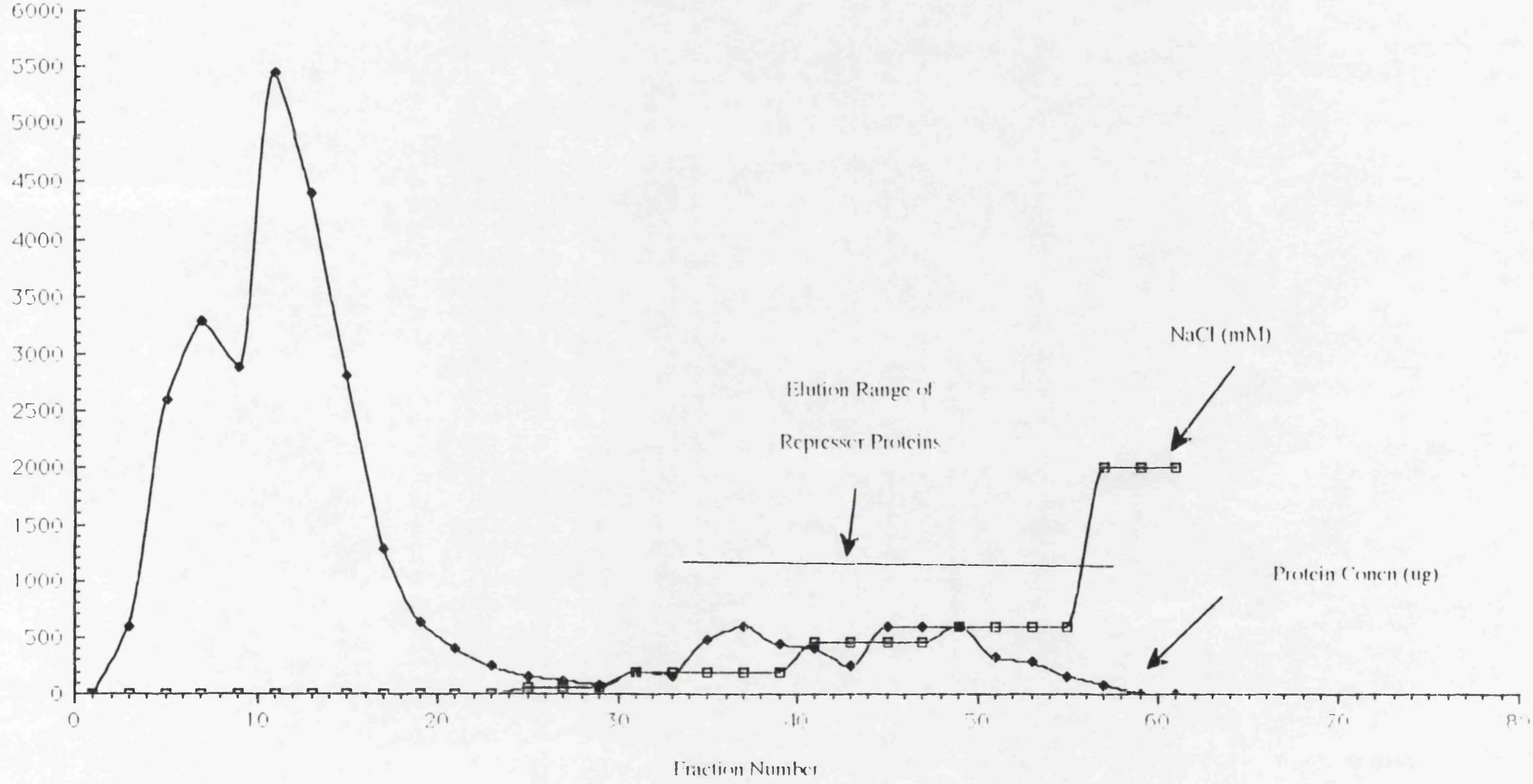


Table 4.1: Heparin-Agarose Elution Profile of the Three Repressor Proteins

Fraction No	Elution Volume (ml)	Eluant Composition
1-11	- -	Flow Through
12-24	20	Buffer A
25-30	10	" " + 50mM NaCl
31-40	15	" " + 200mM NaCl
41-47	10	" " + 450mM NaCl
48-56	15	" " + 600mM NaCl
57-61	10	" " + 2000mM NaCl

Table 4.2. Conditions used during Fast Protein Liquid MonoQ Chromatography

System:	Waters Advanced Protein Purification System
Column:	Pharmacia Mono Q HR5/5
Buffer:	Low Salt: Buffer A
	High Salt: Buffer A + 500mM NaCl
Flow Rate:	0.5ml min^{-1}
Chart Speed:	0.5cm min^{-1}
Back Pressure:	354psi

Gradient profile used during MonoQ ion-exchange chromatography.

Time	Volume	Buffer	Description
0-12mins	6ml	Buffer A	Flow Through
12-52mins	20ml	Linear Gradient of Buffer A to BufferA500	Protein Elution
52-74mins	11ml	BufferA500	Isocratic Elution

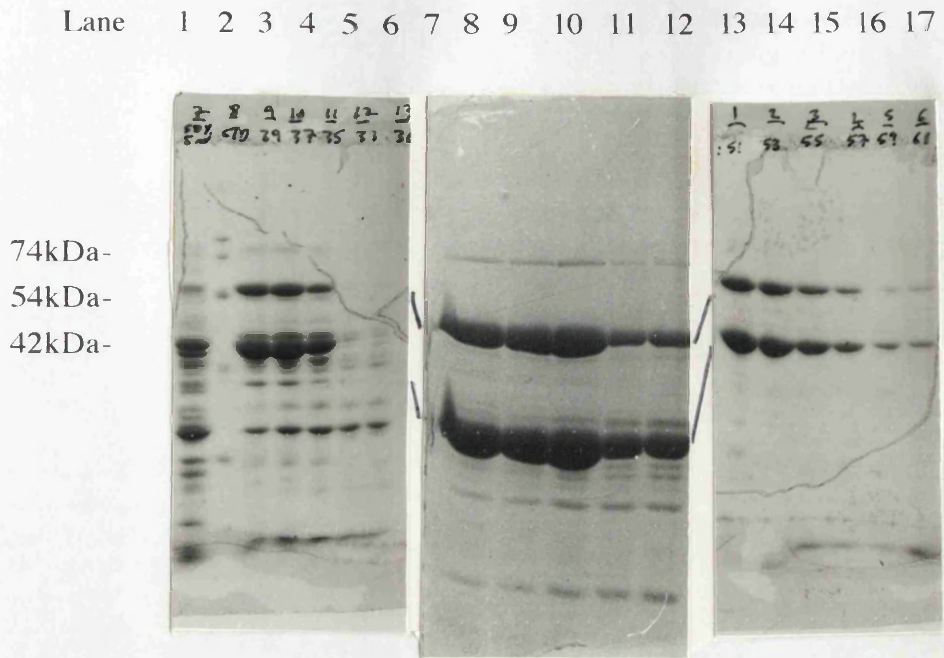


Figure 4.4. SDS-polyacrylamide Gel of Fractions from the Heparin-Agarose Column
A 10% SDS-polyacrylamide gel of fractions from heparin-agarose was run to analyse the presence of the the repressor proteins. Each lane consisted of 20μl of 1.5ml fractions collected from the heparin-agarose column.

Lane	Fraction	Volume (μl)	Lane	Fraction	Volume (μl)
1	50% $(NH_4)_2SO_4$	10	10	45	20
2	STD	10	11	47	20
3	39	20	12	49	20
4	37	20	13	51	20
5	35	20	14	53	20
6	33	20	15	55	20
7	31	20	16	57	20
8	41	20	17	59	20
9	43	20	18	61	20

concentration against fraction number (Figure 4.3). The majority of the repressor proteins elute between Fractions 41-58 at a salt concentration of between 450-600mM NaCl (Figure 4.4). Fractions 46-57 were pooled (5ml) and dialyzed to Buffer A before being used in a Mono Q anion-exchange separation stage.

4.3.5 Anion Exchange Chromatography of the Repressor Proteins

During the development of Mono Q anion-exchange chromatography it was found that the repressor proteins eluted reproducibly at about 500mM NaCl. To increase the effectiveness of the separation, the gel filtration properties inherent within the mono matrix were used to enhance separation of the repressor proteins by incorporation of an isocratic gradient at 500mM NaCl. The repressor proteins were loaded onto the column in buffer A and allowed to bind to the matrix by washing with 8ml of buffer A, during which proteins not bound were eluted. A 20ml linear gradient of buffer A to buffer A containing 500mM NaCl was used to elute the bound protein over 40 minutes followed by a 12ml isocratic gradient of buffer A containing 500mM NaCl for a further 22 minutes(Table 4.2).

Very little protein was eluted during the linear gradient to 500mM NaCl (Figure 4.5). The majority of the repressor protein was found to elute at 500mM NaCl in fractions 32 to 39 (Figure 4.6). Glycerol was added to these fractions to a final concentration of 50% and stored at -20°C. The Mono Q step resulted in the removal of several minor protein contaminants and increased the purity of the repressor proteins to greater than 95%, as deduced by silver stain (Figure 4.6)

4.4 NH_2 Terminal Amino Acid Sequence Analysis

The purification of all 3 represssors away from *E.coli* proteins enabled the N terminal amino acid sequence of the 42 and 54 kDa repressors to be deduced. This was considered essential in confirming that the repressor proteins overexpressed in *E. coli* (presented in Chapter 3) were those predicted from the DNA sequence of the *c* gene.

Samples were sent to the SERC Protein Sequencing facility in Aberdeen run by Dr J.E. Fothergill and Mr Bryan Dunbar, where they were sequenced using an Applied Biosystems 477A pulsed-liquid phase protein sequencer. The first 10 amino acids of the 54 and 42kDa proteins were identified as requested and compared to the predicted amino acid sequences. Both proteins matched perfectly to the predicted sequence (Figure 4.7). Unfortunately, too little of the 74kDa was present in the samples for reliable sequence to

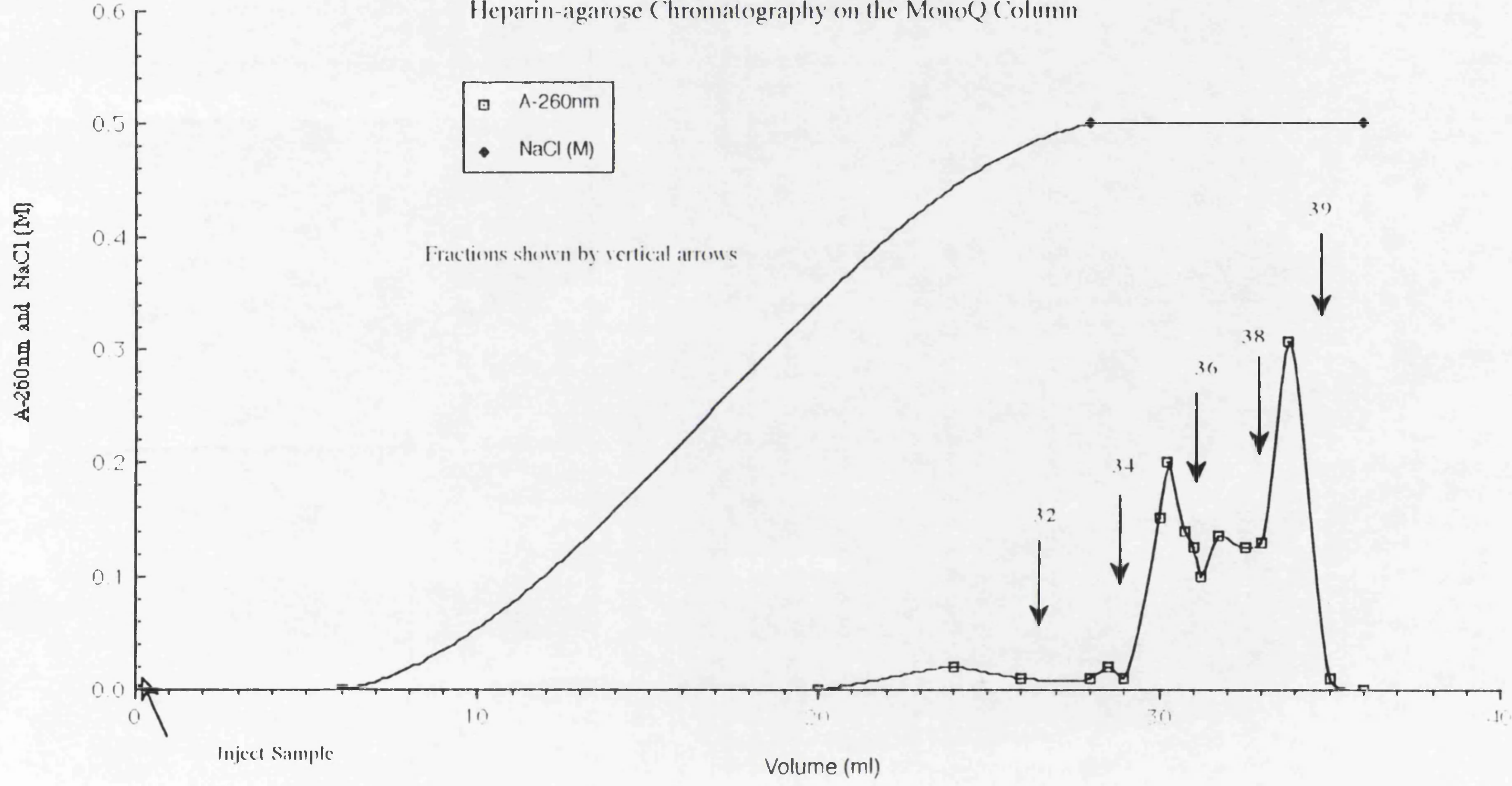


Figure 4.5. Elution Profiles of Fractionation of the Proteins After Heparin-agarose Chromatography on the MonoQ Column

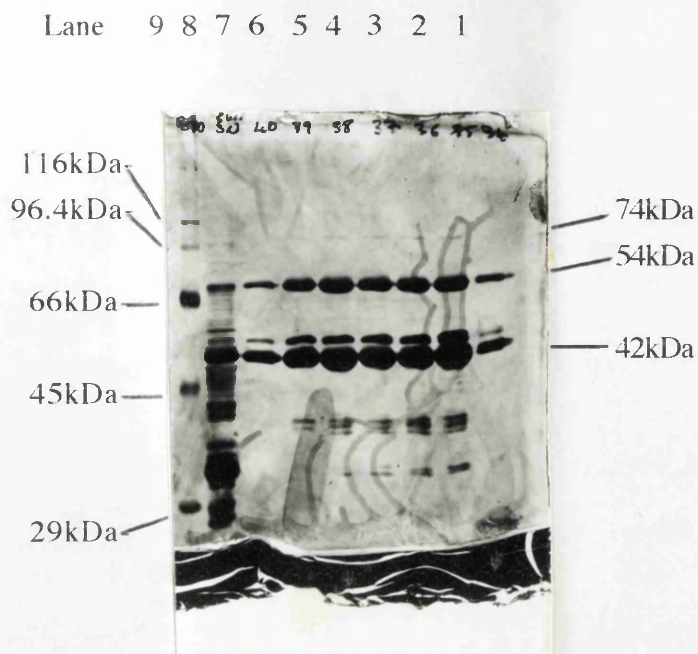


Figure 4.6. SDS-polyacrylamide Gel of Fractions from the MonoQ Column Containing the Purified 74,54 and 42kDa Repressor Proteins

Lane	Fraction	Volume Loaded (µl)	Lane	Fraction	Volume Loaded (µl)
1	34	20	6	39	20
2	35	20	7	40	20
3	36	20	8	50% $(NH_4)_2SO_4$	3
4	37	20	9	High M.Wt. Standards	10
5	38	20			

Figure 4.7.: N-terminal Sequence Analysis of the Repressor Proteins

Repressor	Amount Loaded onto Sequencer
74kDa	50pmoles
54kDa	500pmoles
42kDa	1000pmoles

Sequence	54kDa Repressor											42kDa Repressor										
Experimental	-	S/G/A	T	K	K	L	K	L	K	D	V	-	A	A	K	N	D	T	Q	D	V	D
Predicted	M	A	T	K	K	L	K	L	K	D	V	M	A	A	K	N	D	T	Q	D	V	D

be determined. It was noted that the N-terminal methionine appears to have been removed from both the 54 and 42kDa proteins.

4.5 Conclusions and Discussion

Ammonium sulphate precipitation of the repressor proteins resulted in the repressors remaining soluble to a 60% ammonium sulphate saturation which is consistent with the prediction that these proteins possess a high negative charge (Dreutscher, 1989). The ammonium sulphate saturation level was never taken beyond 50% saturation due to occassional problems encountered in preliminary experiments during which the density of the solution exceeded that of the proteins, causing proteins to float at the surface. In practice the supernatant following protein precipitation at 50% ammonium sulphate saturation was used. Fractionation on heparin-agarose was the single most effective purification step resulting in the removal of the majority of contaminating *E.coli* protein. That the repressor proteins bind to heparin-agarose is indicative of them possessing DNA-binding activity.

The final step of the purification was anion-exchange chromatography on an FPLC MonoQ column. The repressor proteins eluted during an isocratic elution step at 500mM NaCl and were approximately 90% pure, although minor contaminants, possibly due to degradation are visible after SDS-PAGE. The high salt concentration needed to elute the repessor proteins provides more evidence that they possess a negative charge consistent with predictions made from the DNA sequence, and observations made during ammonium sulphate precipitation.

The purification of the repressor proteins enabled N-terminal sequencing of the predicted 54 and 42kDa proteins. These data confirmed the identity of the ORF responsible for each repressor in *E.coli* consistent with predictions made in Chapter 3. During the purification of the 42kDa repressor presented in the following chapter, changes were made in the purification protocol. FPLC chromatography using a MonoQ column was eliminated in favour of a DEAE-Sephacel matrix. In the above experiments the repressor proteins eluted from the MonoQ column consistently at 500mM
NaCl and required dialysis against Buffer A before loading onto the heparin-agarose column. Dialysis is often a source of protein loss and reduced activity. By using a different anion-exchange matrix this problem was overcome, as the 42kDa repressor elutes from the DEAE-sephacel matrix in 200mM NaCl; a concentration which will allow binding to the heparin-agarose matrixwithout dialysis. This is explored in more detail in the following chapter.

Chapter 5

Purification and Biochemical Characterization of the ØC31 42kDa Repressor

5.1 Introduction

The aim of this chapter was to purify the 42kDa protein to analyse its DNA-binding activity. In addition, some biochemical characterisation of the oligomeric nature of the 42kDa protein was performed. If time had allowed these activities would have been compared to a preparation containing both the 54 and 42kDa proteins. Purification of the 42kDa repressor enabled the production of polyclonal antibody and allowed *S.coelicolor* lysogens to be tested for the presence of the repressor proteins.

5.2. Purification Method for Isolating the 42kDa Repressor Protein

Purification of the 42kDa repressor was carried out as presented in Chapter 4 but with some modifications (Section 2.13.4). Notably DEAE-Sephacel chromatography was used before affinity chromatography with heparin-agarose. The presence of active 42kDa repressor was assayed during the purification using a band-shift assay to demonstrate DNA binding to the *c*P1 promoter fragment. A flow-diagram demonstrating the purification protocol used for the isolation of the 42kDa repressor is presented (Figure 5.1).

5.2.1 Purification of the 42kDa Repressor By Anion-Exchange Chromatography

The dialyzed protein solution following ammonium sulphate fractionation (section 2.13) was loaded onto a DEAE-Sephacel anion-exchange column in a volume of 30ml. Unbound protein was washed from the column with 1.5 column volumes of Buffer A (Section 2.13). Protein bound to the column was eluted with a linear gradient of Buffer A with increasing NaCl concentration (Table 5.1).

A graph of total protein eluted against each fraction demonstrates that a significant proportion of the 42kDa protein, as measured using the band-shift assay, that was loaded onto the column was washed through without binding and is present in fractions 7-25 (Figure 5.2). The major proportion of protein bound to the column and eluted with the linear salt gradient in fractions 39-61. Alternate fractions from 41 to 61 were analyzed by SDS-PAGE and visualized by staining with Coomassie Blue R250 . The majority of the repressor was present in fractions 45-57 (Figure 5.3).

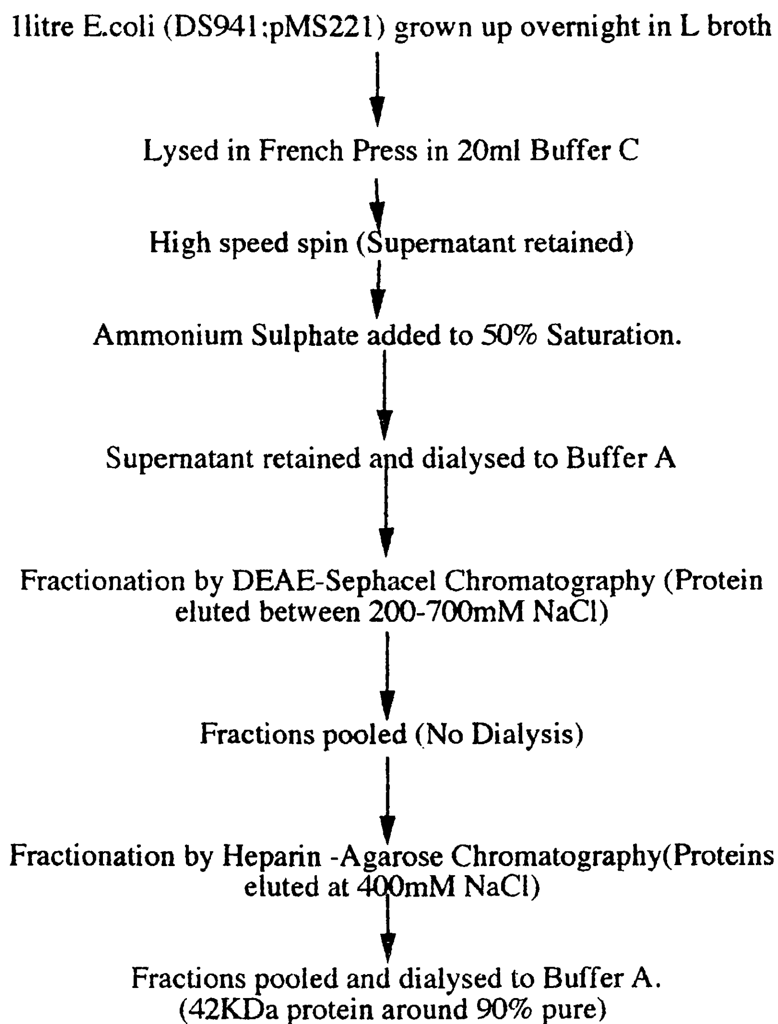


Figure 5.1: Diagram of the Method Used in Purification of 42kDa Protein

Figure 5.2. DEAE-Sephacel Elution Profile of the 42kDa Repressor

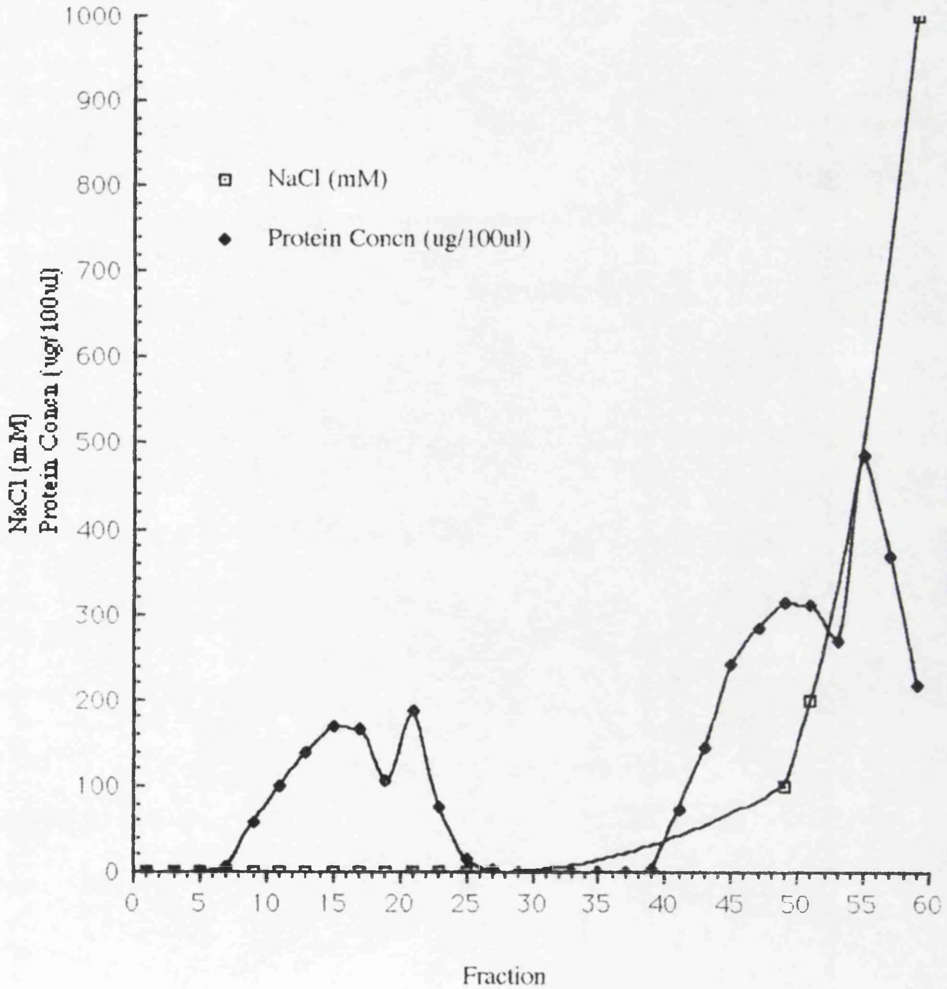
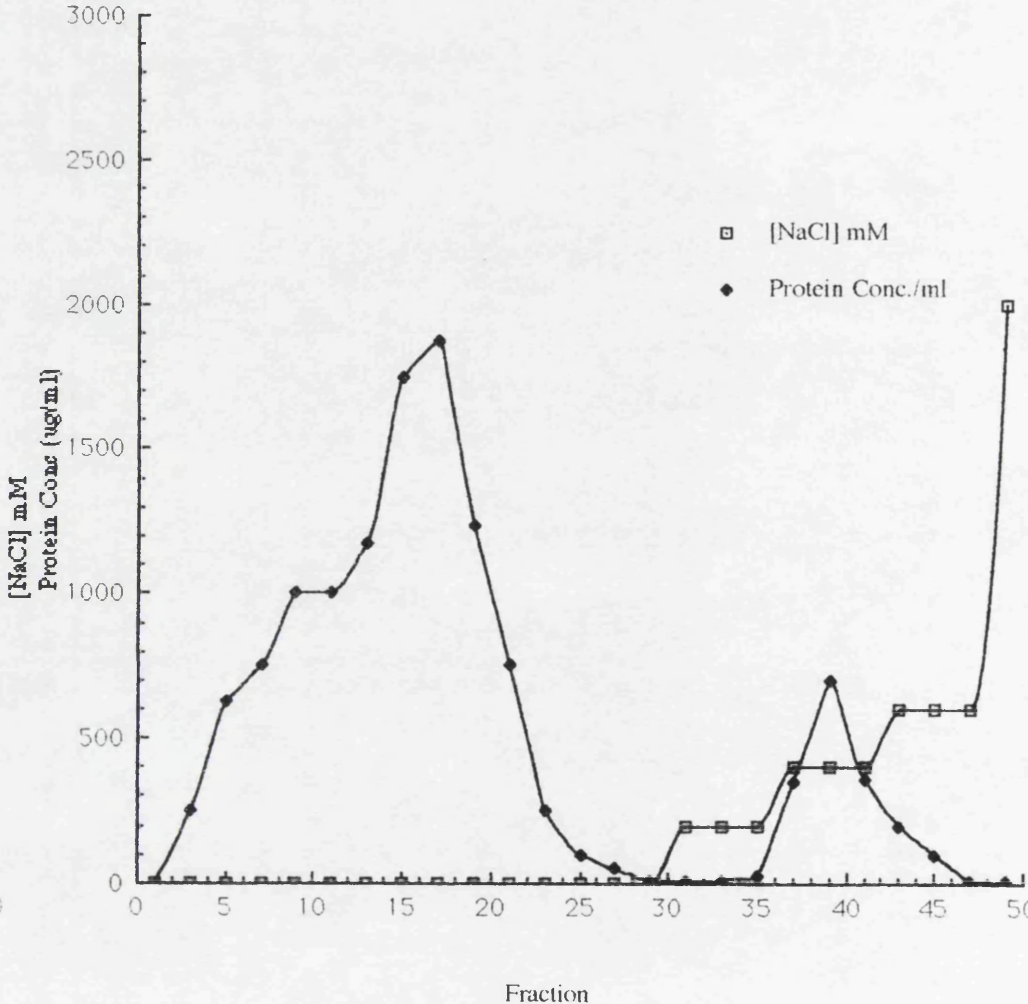


Figure 5.5: Elution Profile of Heparin-Agarose Fractionation of 42kDa Repressor



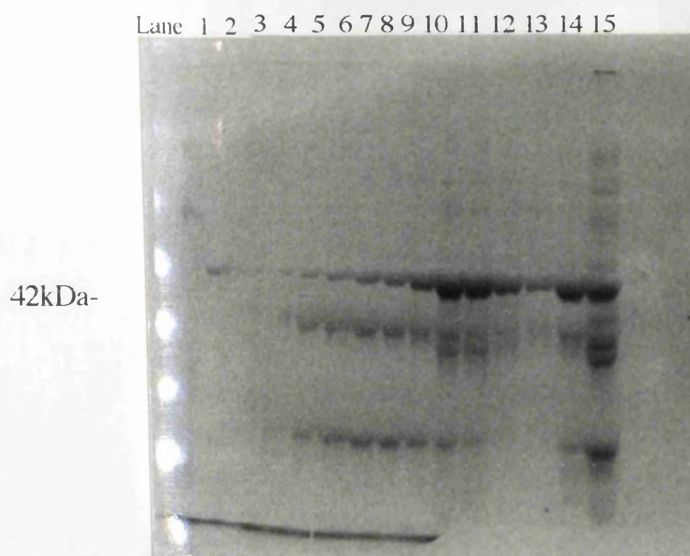


Figure 5.3. Fractions from the DEAE-Sephacel Column
Fractions 41 to 61 from the DEAE-Sephacel column were electrophoresed on a 10% SDS-polyacrylamide and stained with Coomassie R-250.

Lane	Fraction	Lane	Fraction
1	Molecular Weight Markers	9	53
2	Pure 42kDa Repressor	10	55
3	41	11	57
4	43	12	59
5	45	13	61
6	47	14	50% $(NH_4)_2SO_4$ Stage
7	49	15	Crude lysate
8	51		

Table 5.1. Elution Scheme Used During DEAE-Sephacel Fractionation of the 42kDa Repressor Preparation

Fraction	Elution Volume (ml)	Eluant (Buffer A +)
1-17	Crude Sample (30ml)	
18-28	15ml	Buffer A Wash
29-49	30ml	Linear Gradient: 0- 100mM NaCl
50-61	30ml	Linear Gradient: 200-1000mM NaCl

A graphical elution profile from the DEAE-Sephacel column is shown in Figure 5.2.

DNA-binding activity was assayed in fractions 11-23 and in 41-59. Roughly equivalent amounts of protein from each fraction were tested for activity by binding to a radioactively labelled DNA fragment containing the *c*P1 promoter. The DNA fragment used was made by PCR from linearised p19-H3-300 template using a single radioactively labelled oligonucleotide and a non-radioactive complementary primer (Section 2.7.14). From the autoradiograph, it is apparent that DNA binding activity of the 42kDa repressor protein is of sufficiently high specific activity to cause a band-shift in fractions 11-23 and in fractions 49-59 (Figure 5.4). Fractions 48-59 were pooled to a final volume of 18ml. The salt concentration of this solution was estimated to be around 350mM NaCl. This was diluted with 10ml. of ice-cold Buffer A to adjust the salt concentration below 250mM and loaded directly onto the heparin-agarose column.

5.2.2 Fractionation of the 42kDa Repressor Using Heparin-Agarose Affinity Chromatography

The diluted 42kDa repressor protein extract (28ml) from the DEAE-Sephacel anion-exchange column was loaded onto a 7ml heparin-agarose column in Buffer A. Unbound protein was washed from the column with 3 column volumes of Buffer A. whilst that which had bound to the matrix was eluted with a step-gradient of increasing salt in Buffer A (Table 5.3).

The protein content of each fraction was plotted against fraction number (Figure 5.5). The 42kDa repressor began eluting from the heparin-agarose column at salt concentration of 400mM NaCl and was observed in fractions 38-45 after SDS-PAGE (Figure 5.6). Bands of the 42kDa repressor coincided with a small peak observed on the elution profile (Figure 5.5). Fractionation on heparin-agarose resulted in the 42kDa repressor protein to be greater than 90% pure as determined by visualization on a Coomassie stained SDS-polyacrylamide gel (Figure 5.5).

Alternate fractions were tested using the band shift technique for DNA-binding activity to a fragment containing the *c*P1 promoter (Figure 5.7). In the band-shift assay, fractions 5-21 and 35-45 were assayed for binding to the *c*P1 promoter fragment. It is clearly apparent that DNA binding activity is present in fractions 37-43. Highest levels of the 42kDa protein were observed in fraction 39 which coincides with the most shifted band (Figure 5.7).

Lane 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19

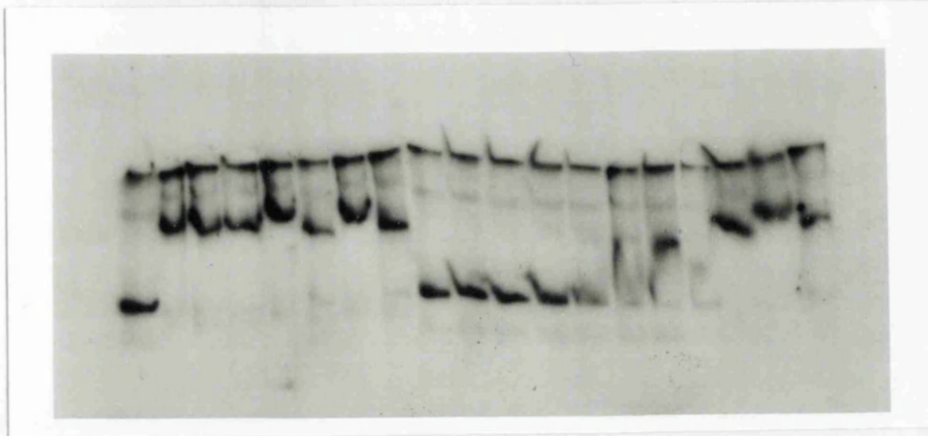


Figure 5.4. DNA Binding Activity of Protein Fractions from DEAE-Sephacel Column for the cP1 Promoter Fragment.
Fractions 11 to 23 and 39 to 59 were tested for DNA binding activity using the conditions described in Table 5.2.

Table 5.2. Amount of Protein Loaded From Fractions Eluting Off the DEAE-Sephacel Column to Assay DNA-Binding Activity.

Lane	Fraction	Protein (ng)	Lane	Fraction	Protein (ng)
1	No Protein	-	11	43	800
2	11	1200	12	45	1280
3	13	1600	13	47	1480
4	15	1800	14	49	1640
5	17	1760	15	51	1600
6	19	1200	16	53	1360
7	21	2100	17	55	2330
8	23	900	18	57	1920
9	39	80	19	59	1280
10	41	400			

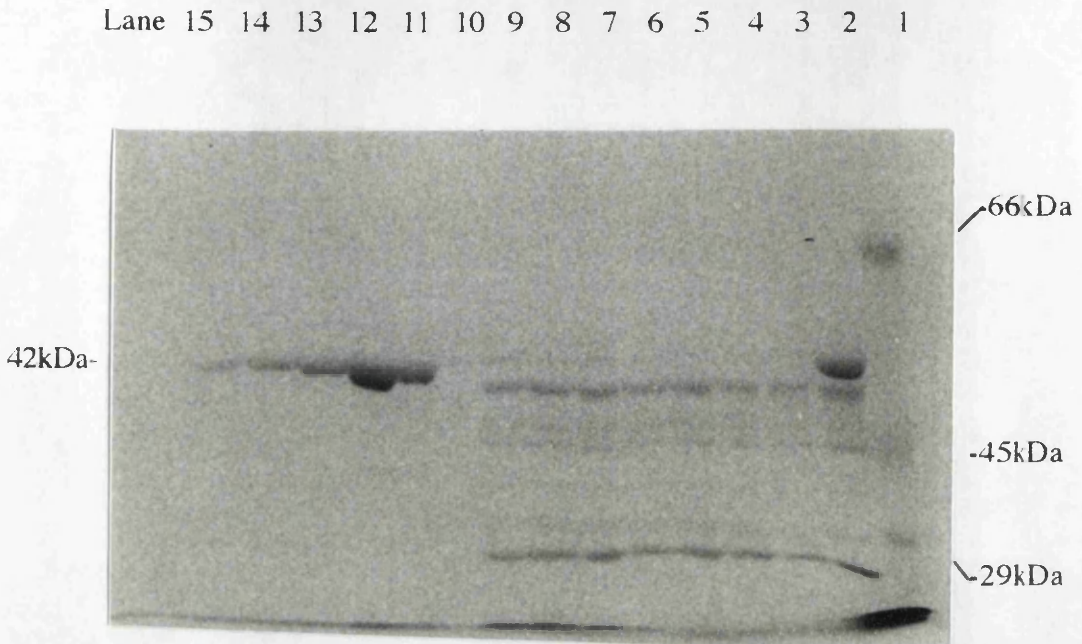


Figure 5.6. SDS-Polyacrylamide Gel of Fractions Collected from Heparin-Agarose Column During Fractionation of the 42kDa Repressor Protein.
Fractions 7 to 19 and 35 to 45 were analyzed on a 10% SDS-polyacrylamide gel and stained in Coomassie R-250.

Lane	Fraction	Lane	Fraction
1	Molecular Weight Markers	8	17
2	Pooled DEAE-Seph. Fractions	9	19
3	7	10	35
4	9	11	37
5	11	12	39
6	13	13	41
7	15	14	43
		15	45

Table 5.3. Elution Scheme Followed During Fractionation of the 42kDa Repressor Protein Using Heparin-agarose Chromatography.

Fraction	Volume (ml)	Buffer
1-17	25	Pooled DEAE-Sephacel Fractions
18-30	20	Buffer A
31-35	10	Buffer A + 200mM NaCl
36-41	10	Buffer A + 400mM NaCl
42-48	10	Buffer A + 600mM NaCl
49-54	10	Buffer A + 2000mM NaCl

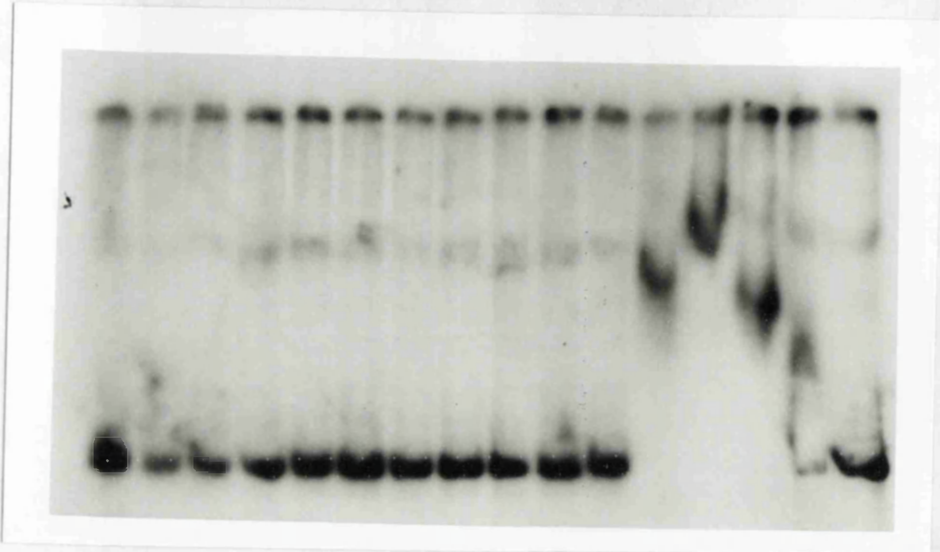


Figure 5.7. DNA Binding Activity in Fractions After Heparin-Agarose Fractionation of the 42kDa Repressor Protein.
Fractions 5 to 21 and 35 to 45 were tested for DNA binding activity using the gel retardation assay.

Lane	Fraction	Lane	Fraction
1	No Protein	9	19
2	5	10	21
3	7	11	35
4	9	12	37
5	11	13	39
6	13	14	41
7	15	15	43
8	17	16	45

Table 5.4. Table Demonstrating the Amount of Protein From Each Fraction From the Heparin-agarose Column Used During Testing for DNA-Binding Activity.

Fraction	Protein (ng)	Fraction	Protein (ng)
5	320	21	320
7	360	35	60
9	480	37	320
11	480	39	620
13	560	41	360
15	720	43	200
17	790	45	80
19	640		

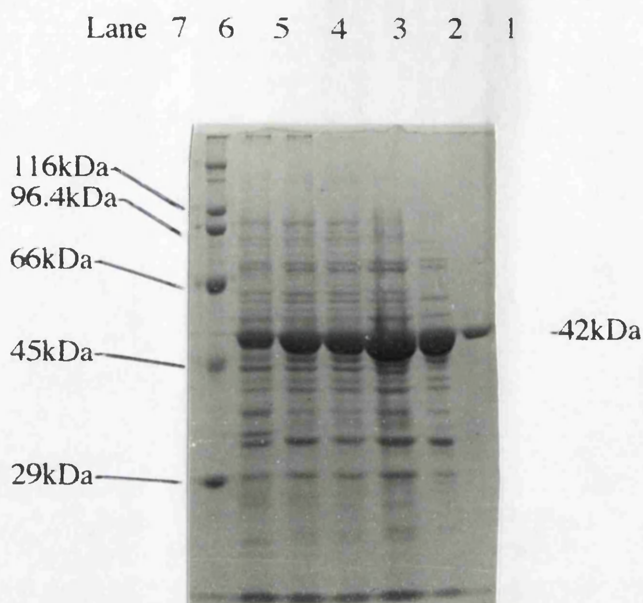


Figure 5.8. Purification Profile of the 42kDa Repressor
Protein from each purification stage were analysed by SDS-PAGE (10% Laemmli gel stained in Coomassie G-250). Lane 1- Pooled Heparin-Agarose fractions (5ug); Lane 2- Pooled DEAE-Sephacel Fractions (20ug); Lane 3- 50% Ammonium Sulphate supernatant (30ug); Lane 4- High Speed Spin Supernatant (20ug); Lane 5- Low Speed Spin Supernatant (20ug); Lane 6- Crude Lysate of E.coli DS941:pMS201 (20ug); Lane 7- High Molecular Weight Standard.

All fractions possessing DNA-binding activity were pooled into a volume of approximately 16ml and glycerol added to 50%(v/v). The purified protein was stored in 2ml aliquots at -20°C and appeared to be stable for at least 8 months.

5.3. Purification Summary for the 42kDa Repressor Protein

A purification table for the 42kDa repressor has been constructed to provide a measure of the relative success of the purification (Table 5.6). A profile of the protein content at each stage of the purification is demonstrated by SDS-PAGE (Figure 5.8).

Serial dilutions of protein after each purification stage of the protocol were used to quantitate the binding affinity of the 42kDa repressor protein for the *c*P1 promoter fragment. Bands from a band-shift assay were excised and quantitated by Cerenkov scintillation counting (Figure 5.9). These data were plotted (Figure 5.10) as a proportion of bound DNA against protein concentration. The crude lysate data shown in the experiment did not band-shift DNA probably because of the protein concentrations used. A similar experiment using double the protein concentration demonstrated that protein in the crude lysate could bind DNA . A purification table of the relative activities after each step of the purification procedure downstream of the ammonium sulphate precipitation was constructed (Table 5.6).

The purification table demonstrates a 3.5-fold purification of the repressor protein with a yield of 35% of total activity. The activity of the 42kDa repressor protein was calculated in $U.mg^{-1}$ (Majumdar et al, 1987). A unit (U) is defined as the quantity of repressor (ng) required to bind 50% of 5fmoles of $cP1^{H3}$ promoter DNA under the binding conditions described in section 2.15.1. The purified 42kDa repressor protein possesses a specific activity of $4028U.mg^{-1}$ repressor as measured using this assay. This represents a purification in which a substantial amount of the total activity has been retained. Most of the activity lost during the purification occurred during the anion-exchange step on the DEAE-Sephacel column where a large proportion of the 42kDa repressor washed straight through the column without binding.

5.4 Biochemical Characterisation of Pure 42kDa Repressor

In this section a study of the oligomeric structure of the ØC31 42kDa repressor protein using dimethylsuberimidate cross-linking and gel filtration chromatography will be presented.

Lane 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

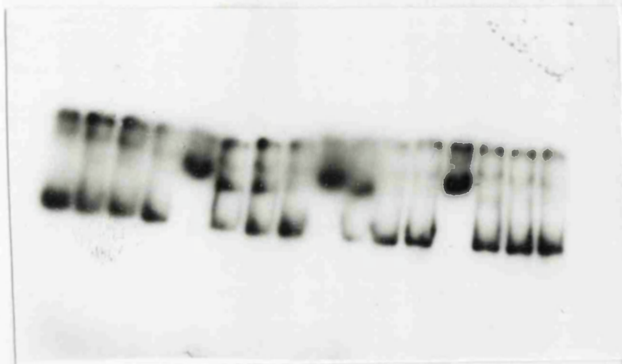


Figure 5.9. DNA-Binding Activity of the 42kDa Protein After Each Purification Stage
Protein from each stage of the purification was incubated for 15 minutes with 5fmoles end-labelled *c*P1^{H3-300} fragment before separation on a 4% native polyacrylamide gel in 30mM NaPi pH6.6 buffer (See lane numbers in Table 5.5). The amount of protein used in the assay was determined by the Bradford assay and is presented in Table 5.5.

Table 5.5. Total protein estimations used in quantitating DNA-binding activity of the 42kDa Repressor for the *c*p1 promoter fragment at each stage of the purificaton.

Purification Stage	Crude Lysate				50% $(NH_4)_2SO_4$				DEAE Sephacel				Heparin-agarose			
Lane Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Protein Concentration (μg)	6.0	3.0	1.5	0.75	4.0	2.0	1.0	0.5	3.5	1.75	0.87	0.44	0.57	0.29	0.14	0.07

Figure 5.10. Binding Activity in Fractions for cP1 Fragment Using Different Stages from 42kDa Purification

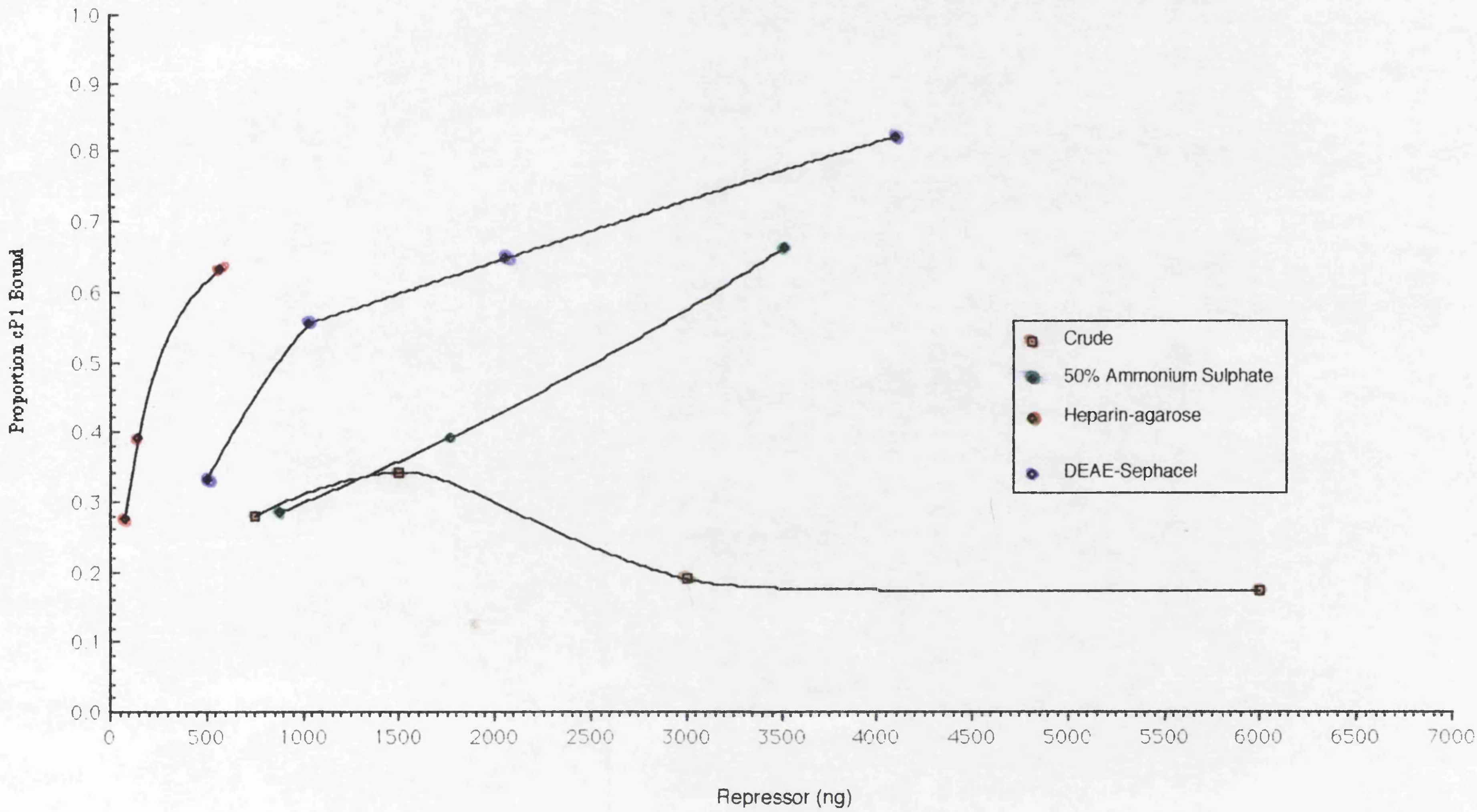


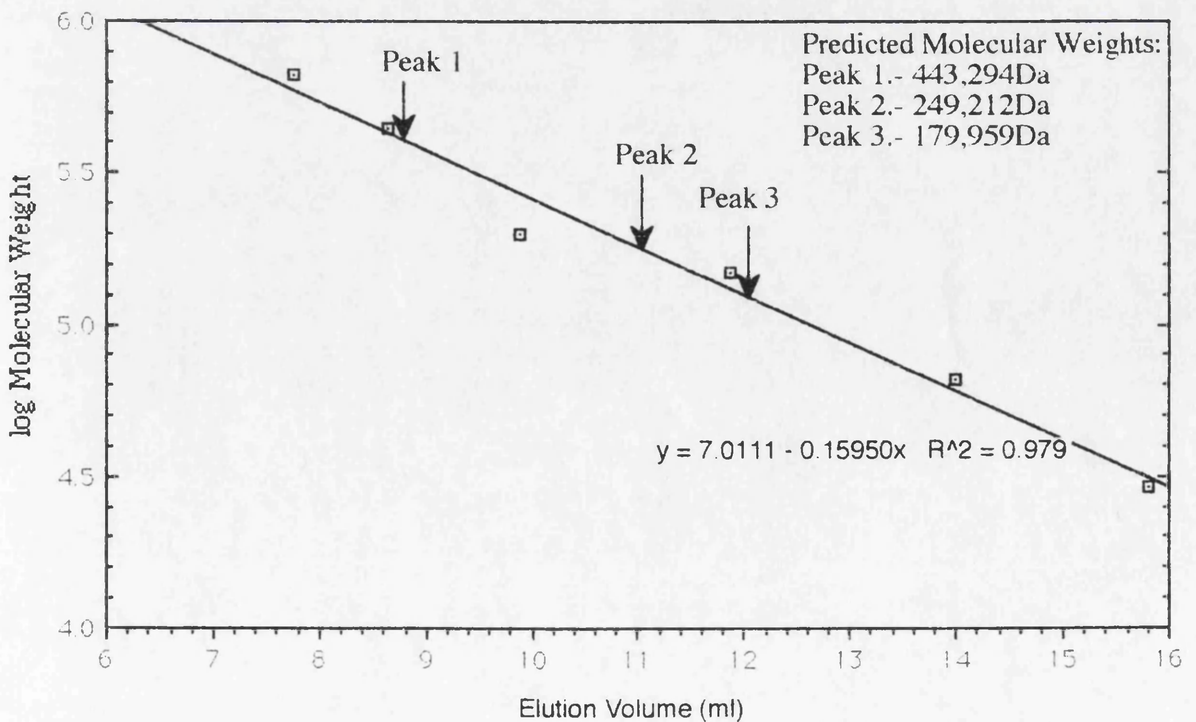
Table 5.6. Purification Table for the 42kDa Repressor.

Stage	Volume (ml)	Total Protein (mg)	Total Activity (U)	Specific Activity (U/mg)	Purification Factor	Yield (%)
1. Crude Lysate	27	245.7	approx. 8.2×10^4	332	-	-
2. 50% Supernatant	32	72.0	8.2×10^4	1138	-	100
3. DEAE-Sephacel Stage	18	56.7	2.2×10^4	395	0.35	27
4. Heparin Agarose Stage	16	7.2	2.9×10^4	4028	3.54	35

a: The DNA binding activity unit for the purpose of this purification has been defined as the amount of protein required to produce 50% binding of 5fmoles $cP1^{H3}$ fragment as measured in a band shift assay. This was quantitated by isolating bands from a gel dried onto Whatman 3MM paper and Cerenkov counted for 2 minutes in a scintillation counter. The results were then plotted and the 50% binding unit determined from the graph. b: Protein concentration was determined by the method of Bradford (1981). The crude data has been estimated from a separate experiment and hence has not been used in the calculations above.

Figure 5.11. Typical Trace of Fractionation of the 42kDa Repressor on the Superose 6 Gel Filtration Column

Figure 5.12 Superose 6 Gel Filtration Standard Curve



5.4.1 Superose 6 Determination of the Oligomeric Molecular Weight of the Monomeric 42kDa Repressor Protein

The purified 42kDa repressor protein was analysed on the Superose 6 gel filtration column to ascertain its native molecular weight and thus determine its oligomeric nature. Before and after each run, standards of known molecular weight were passed down the column. Samples of 100μl (25μg) of purified 42kDa repressor protein were applied to the column and fractionated through the column at a flow-rate of between 0.3-0.5ml.min.$^{-1}$. Figure 5.11 demonstrates a typical trace of the major peaks of protein detected at 280nm. The sizes correspond to proteins with apparent molecular weights of 443,000Da, 249,000Da, and 180,000Da. The molecular weight of the monomeric repressor sub-unit is assumed to be 42,000Da based on the known sequence of the protein. Consequently, the repressor could be considered to form either a tetramer, a hexamer or a decamer.

Gel filtration chromatography was performed several times changing several parameters including; the type of buffer, salt concentration, protein concentration and pH. The manipulations included using Tris.Cl buffer at pH8, potassium phosphate buffer at pH6.6, buffers containing no salt, 200mM NaCl and 500mM NaCl. The concentration of protein loaded onto the column was also adjusted between 250ng.μl^{-1} to 1.5μg.μl^{-1}. No observable change in the protein trace was obtained after each alteration of the running conditions. The sensitivity of the detection system was also changed from the relatively insensitive 280nm to 220nm when sodium phosphate buffer was usedwith no obvious effect. The only factor which influenced the oligomeric structure of the 42kDa repressor protein in solution appeared to be the amount of proteolytic degradation of the protein sample applied to the column. The protein trace shown is typical of that obtained after gel filtration of the purified 42kDa repressor and was chosen because of the particularly noticeable peak corresponding to a protein of 443,000Da.

The two major peaks eluting from the gel filtration column appeared to relate to putative tetrameric and hexameric protein species (Figure 5.12). The peak which related to the predicted decameric complex was usually of relatively weak intensity, and was not observed when low quantities of protein were used. The appearance of a tetrameric and a hexameric form of the same protein is unusual. Tetramers are most often the result of two dimers co-operating to form a tetramer (White and Richardson, 1987). Samples from fractions corresponding to the putative hexamer, tetramer and dimer were tested for DNA binding activity (Figure 5.11). DNA binding activity was observed for the

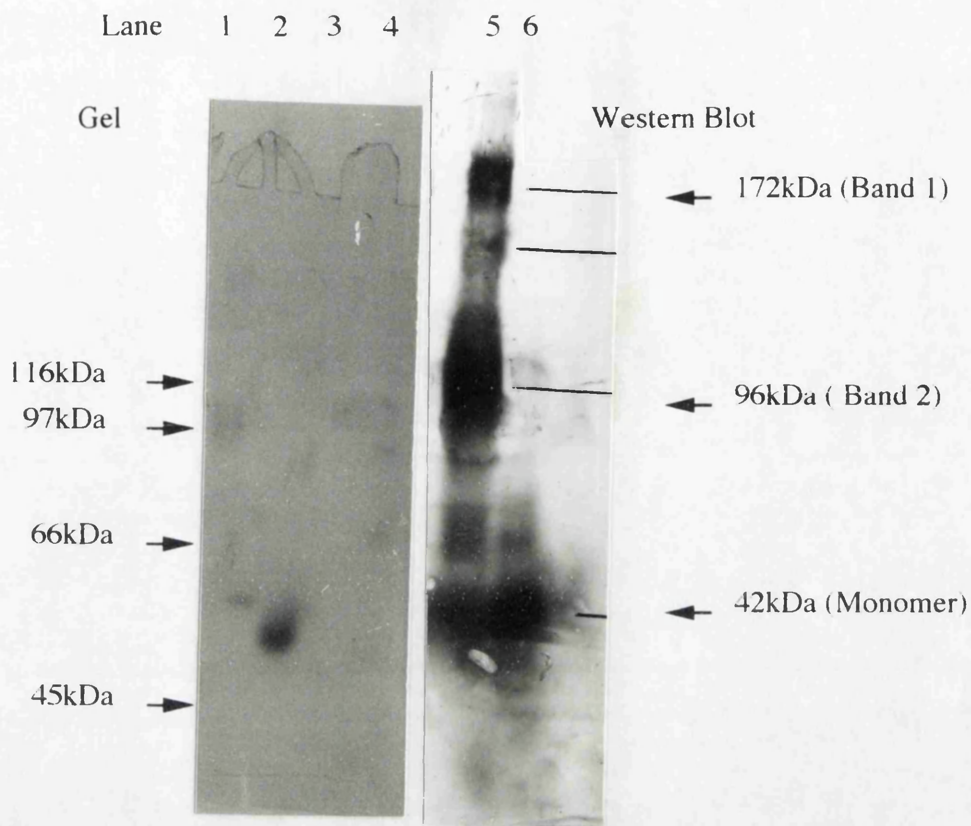


Figure 5.13. Dimethylsuberimidate Cross-linking of the Purified 42kDa Repressor Protein.
The 42kDa repressor protein was crosslinked with 10mM dimethylsuberimidate. Proteins run on 10% Laemmli gel and stained with Coomassie Blue G250. Lane 1- 42kDa (Crosslinked 10μg). Lane 2- 42kDa Repressor (Not Crosslinked-5μg); Lane 3- Crosslinked Aldolase (10μg); Lane 4- High Molecular Weight Markers;
Lanes 5 and 6 represent a western blot of the same cross-linked gel above. Lane 5- Cross-linked 42kDa repressor (10ug); Lane 6- 42kDa repressor (5ug) (not cross-linked)

Figure 5.14. Dimethylsuberimidate Crosslinking Standard Curve

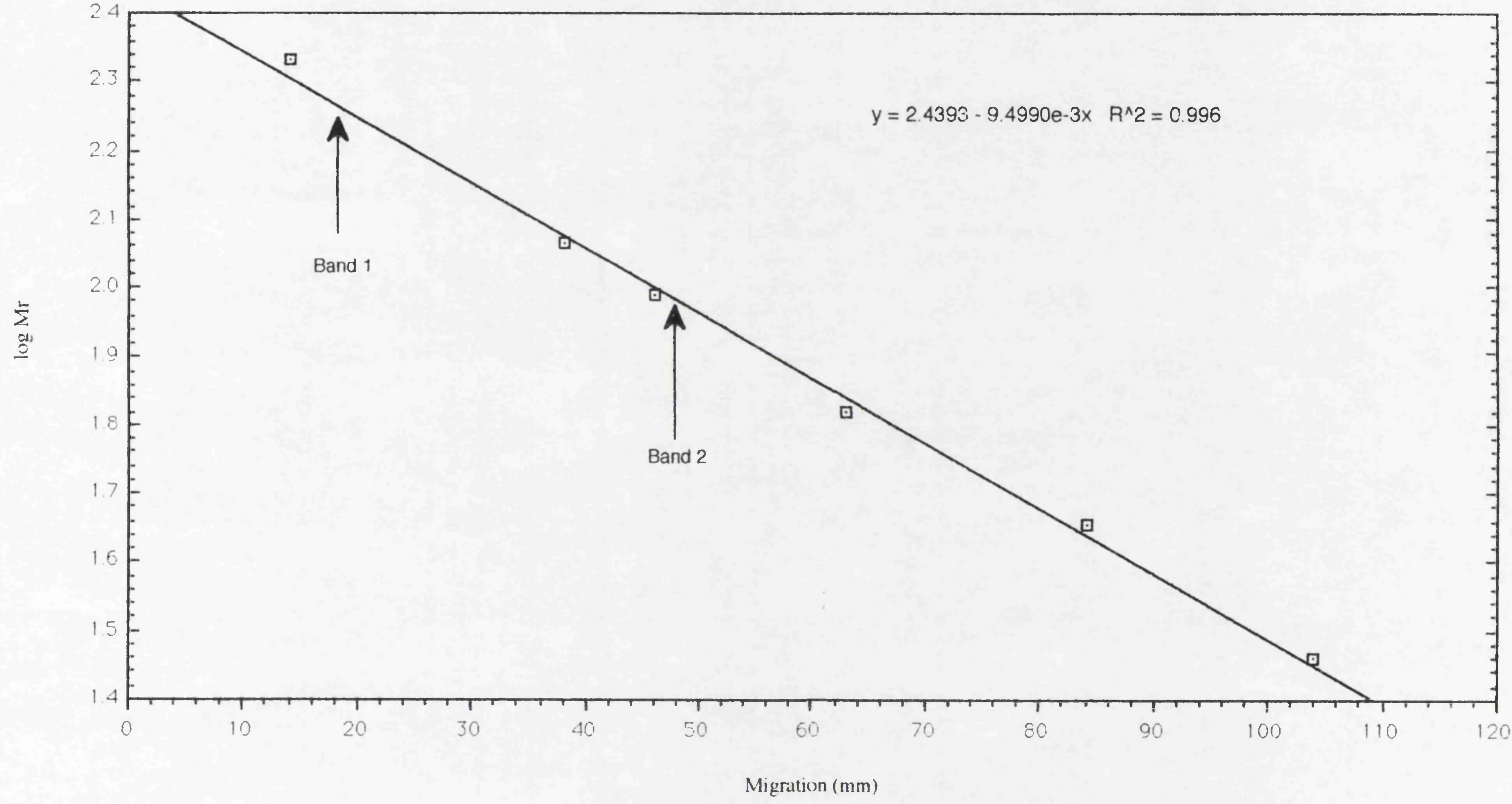


Table 5.7. List of Protein Standards Used During Gel Filtration Chromatography.

Protein	Molecular Weight (kDa)
Blue Dextran	2,000,000
Thyroglobulin	669,000
Apoferritin	443,000
β-Amylase	200,000
Alcohol Dehydrogense	150,000
Bovine Serum Albumin	66,000
Carbonic Anhydrase	29,000

Table 5.8. Correlation Of Predicted Oligomeric Form of the 42kDa Repressor With Observed Moleculer Weight

Band No	Observed Mr	Mul timer
1	182kDa	Tetramer
2	95.5kDa	Dimer

tetramer, but not for the hexameric or dimeric forms of the repressor (Figure 5.11). The decamer was not tested for binding activity due to its extremely low abundance.

5.5. Chemical Cross-linking of the 42kDa Repressor Using Dimethylsuberimidate

Cross-linking of the ØC31 42kDa monomer repressor protein was performed using dimethylsuberimidate (Lumsden and Coggins, 1977). Purified 42kDa repressor was mixed with 10mM dimethylsuberimidate in 100mM Triethanolamine.HCl pH8 containing salt at 50mM concentration (Section 2.9.8). Repressor protein was added at a concentration of $375\mu g.ml^{-1}$ and the cross-linking reaction performed at room temperature for 1 hour. The reaction was terminated by the addition of sodium bicarbonate in excess. The complexes formed were analyzed on a 5% continuous native gel in 50mM NaPi buffer pH7 (White, 1989). As a standard, 100µg of aldolase was also crosslinked with dimethylsuberimidate (Lumsden and Coggins, 1977).

From the polyacrylamide gel, two bands are observed which when compared to size standards (Figure 5.14) possess molecular weights of 182 and 95.5kDa respectively (Table 5.8) which would be predicted to correspond to a tetramer and dimer (Figure 5.13). Very small quantities of putative trimeric and pentameric complexes are observed on a western blot of the cross-linked gel data (Figure 5.13). The aldolase control does not appear to have cross-linked with a very high efficiency; the reason for this failure is presumed to be because of Tris buffer in the commercial preparation of aldolase. Tris interferes with cross-linking by dimethylsuberimidate (Sigma specification) because of the free amine groups it possesses which are able to interact with the dimethylsuberimidate and reduce the efficiency of cross-linking (Coggins, 1978).

5.6. Discussion and Conclusions

The experiments described above demonstrate that the 42kDa repressor protein produced from the *c* locus of the streptomycete bacteriophage ØC31, heterologously expressed in *Escherichia coli* has been purified in an active form. The purification has enabled antibody to be produced against the 42kDa repressor which cross-reacts with the 74 and 54kDa repressor proteins and has demonstrated all three repressor proteins are produced in a *Streptomyces coelicolor* lysogen (Smith and Owen, appendix). The 42kDa repressor has been purified to greater than 90% purity using ammonium sulphate precipitation, anion-exchange chromatography and heparin-agarose chromatography.

Analysis of fractions for DNA binding activity by the 42kDa repressor, it becomes evident that a proportion of the activity for the *c*P1 promoter fragment has not bound to the DEAE-Sephacel column, but has eluted in the washing step. This represents a major

loss of repressor protein and is almost certainly due to the binding capacity of the DEAE-Sephacel column being exceeded. DEAE-Sephacel has a binding capacity of 8.2mg.ml^{-1} of ferritin (440kDa) and 38mg.ml^{-1} a-lactalbumin (14kDa). From analysis of the data collected it can be calculated that 54mg of protein bound to the DEAE-sephacel matrix under the conditions specified. This represented a loss of 18mg of protein which equates to a 25% loss of total protein. This can be rectified in future by using a DEAE-Sephacel column with a larger volume of matrix. Heparin-agarose chromatography proved very effective as a final purification step raising the purification factor of the 42kDa repressor from 1.19 following DEAE-sephacel chromatography to 12.55.

A purification table is presented (Table 5.5) but the following caveats should be noted. The activity was measured using the band-shift assay using protein from stages during the purification where the salt and DNA concentration varies. This method relies upon a shift in the relative mobility of a labelled fragment of DNA upon binding of a specific DNA binding protein. The binding can be influenced by a number of environmental factors including temperature, ionic concentration, competitor DNA and pH. Binding reactions took place at 24°C in sodium phosphate buffer at pH6.6 and each binding reaction had 250ng of cold pUC19 non-specific competitor added to it. The DNA content inherent in each of the purification stages is unknown; although in subsequent binding experiments it was found that up to 10μg of cold pUC19 could be added to a 42kDa repressor: DNA complex before any competition effects were observed. The salt concentration of each of the purification stages was not controlled; this could significantly alter the DNA binding characteristics of the repressor. For example, the affinity of *lac* repressor to *lac* operator DNA *in vitro* changes by a factor of 20 as the concentration of NaCl changes from 0.1M to 0.2M (Winter and von Hippel 1981, Mossing and Record 1985, Schleif 1988).

In summary, the 42kDa repressor protein has been purified to a state whereby it constitutes greater than 90% of the total protein using a purification scheme consisting of three steps. The protein has been demonstrated to be active and will bind the cP1^{H3} promoter fragment. That a purification table with only a 3.5-fold increase in purity is presented is misleading, it should be borne in mind that the 42kDa repressor is substantially overexpressed in *E.coli* and represents by far the major protein product assuming probably greater than 5% of the total protein of the cell. When analysed on a denaturing gel, the 42kDa protein represents the only major protein band and possesses a specific activity of 4028U.mg^{-1}.5fmoles^{-1} cP1^{H3}.

Biochemical studies of the 42kDa repressor were performed using dimethylsuberimidate cross-linking and gel filtration chromatography and demonstrated that the protein most probably exists as a tetramer, although higher oligomeric forms were observed using gel filtration chromatography. Chemical cross-linking is less likely to distinguish very large protein multimers because of limitations in the size of protein able to enter the gel matrix. To overcome this drawback, the newly developed high density agarose matrices may provide a solution. The presence of the higher oligomeric forms of the ØC31 42kDa repressor may be the result of the overexpression of the protein in *E.coli*. The T7 gene 18 protein when overproduced is observed on a gel filtration column as an octamer with a molecular weight of approximately 300kDa (White and Richardson, 1987). The majority of activity of the overproduced protein resides with the octamer, no activity is observed for either the observed tetrameric or dimeric forms of the protein (White and Richardson, 1987). In contrast, however, analysis of T7 infected cells for gene 18 protein activity demonstrates that the majority of activity elutes with dimeric and tetrameric forms of the protein (Hamada *et al*, 1984). Thus, it is not clear what the active form of T7 gene 18 protein is in the cell. It is possible that overproduction of the protein has resulted in aggregation. Alternatively, the octamer may be processed to active dimeric or octameric forms in the cell. In the case of the 42kDa repressor it is possible the hexameric and decameric forms may be a result of some similar sort of protein aggregation due to their gross overproduction. The active form of the 42kDa repressor protein is predicted to be a tetramer.

The size of the repressor proteins from ØC31 is one of the few examples of a relatively large repressor compared to other prokaryotic repressor proteins. The tetrameric size of the 42kDa monomer is 168kDa. In comparison, the *deo*R repressor from *E.coli* is an octamer of 220kDa, the *lac* repressor from *E.coli* is a tetramer of 150kDa (Chakerian *et al*, 1991) and *ahr*C from *B.subtilis* is a hexamer of 102kDa (PG Stockley, personal communication) . The *lac* repressor and the *deo* repressor have a role in DNA looping. The oligomerization of the *deo* repressor to form an octamer allows it the potential to contact three operators simultaneously via a double loop although this has not been observed as yet (Mortenson, Dandanell andHammer, 1989). The *ahr*C gene product has also recently been observed to participate in DNA looping (P.G. Stockley, personal communication). As the 42kDa protein exists as a tetramer it is possible that the 42kDa repressor may form looped complexes. A similar mechanism is observed in bacteriophage P22 whereby DNA looping occurs to regulate tightly gene expression (Valenzuala and Ptashne, 1989).

In summary, biochemical studies have demonstrated that the 42kDa repressor most probably exists in solution as a tetramer of 168kDa, which has been shown to possess DNA binding activity. Detailed studies of the interaction of the 42kDa repressor protein with DNA are presented in Chapter 6 using the band shift assay and DNaseI footprinting. The isolation of the 42kDa repressor should simplify the characterization of DNA binding by complexes containing both the 54 and 42kDa repressor proteins, and later with the addition of the 74kDa repressor.

Chapter 6

Characterization of DNA Binding by the ØC31 42kDa Repressor Protein

6.1 Introduction

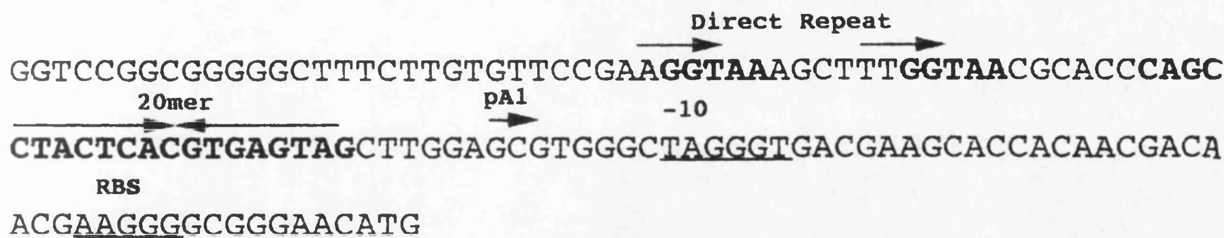
The band-shift assay was used to measure DNA-binding by the 42kDa repressor to fragments from the repressor gene locus and from the early region. Some evidence for auto-regulation had been observed in previous experiments, e.g. heat induction of a ØC31 c^{ts} lysogen of *Streptomyces coelicolor* results in increased transcription from *c*P1 (Sinclair and Bibb, 1989), and the repression of a *c*P1-*lac*Z fusion in *E.coli* by the presence of compatible plasmids expressing the 42kDa, or both the 42 and 54kDa proteins. Interestingly, a conserved 20 base pair highly conserved motif was identified throughout the phage genome and in particular, identical copies are present upstream of the -10 sequences in *c*P1 and *c*P2 (section 1.7.2.; C. Ingham and M.C.M Smith, personal communication). This motif, named the "20mer" is believed to be a protein binding site and possibly a site for general regulation within the phage as it appears to be associated exclusively with promoters and/or terminators (C.Ingham and M.C.M Smith, personal communication).

6.2. *c*P1 Promoter

Expression from the *c* gene is thought to be autoregulated (section 1.7) and therefore, was chosen initially as the substrate to test DNA binding by the 42kDa repressor. The 320bp *Hin*dIII-*Pvu*I fragment was cloned into the polylinker of pUC19 (p19H3-300; M.C.M Smith, personal communication) as was the larger 475bp *Nhe*I-*Sma*I (pCO107)(Figure 6.1).

The polymerase chain reaction (PCR) was used to produce DNA fragments from p19H3-300 and pCO107. In both reactions oligo. 473 (Figure 6.2; section 2.7.14) was used as a primer internal to the *c* gene primer and either the universal primer or the reverse primer of pUC19 used as the complementary primer depending on the orientation of the clone. To produce labelled DNA fragments, a single primer was end-labelled using T4 polynucleotide kinase (Section 2.8.1).

Figure 6.1. Diagrammatic Representation and Sequence of *c*P1 Promoter Region



*c*P1

Terminator
Direct Repeat'
20mer
-10
ATG

NheI
AvaII
HindIII
20mer
ATG
TaqI
PvuI
SmaI
1127
1168
1200
1338
1521
1600
Primer 473

*c*P1-H3

*c*P1-AT

*c*P1-Nhe

Figure 6.2. Diagrammatic Representation of the Fragments of the *c*P1 Promoter DNA Used in DNA Binding Studies

Lane 1 2 3 4 5 6 7 8 9 10

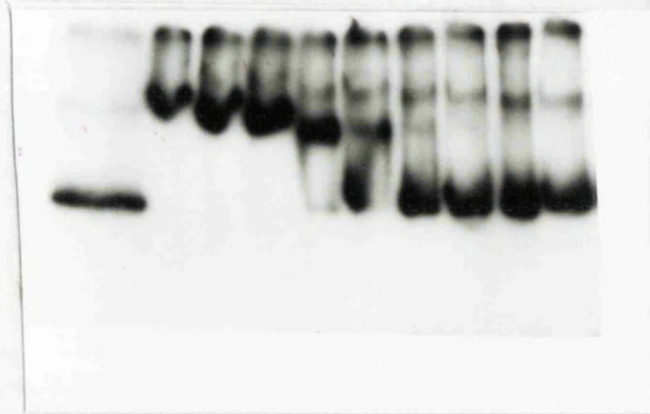


Figure 6.3. DNA Binding Assay of the 42kDa Repressor to the cP1^{H3} Promoter. DNA binding was performed using the band-shift assay.

Lane	Repressor (nM)	Lane	Repressor nM)
1	No Protein	6	119
2	684	7	89
3	298	8	60
4	208	9	45
5	149	10	21

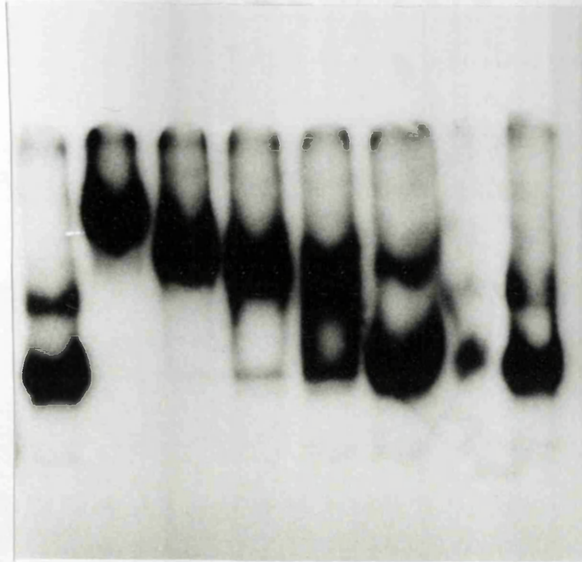


Figure 6.4. Affinity of the 42kDa Repressor Protein for the $cP1^{Nhe}$ Promoter Fragment.

Lane	Repressor (nM)	Lane	Repressor (nM)
1	No Protein	5	161
2	655	6	119
3	321	7	Not Loaded
4	208	8	44

Figure 6.5. Diagrammatic Representation and Sequence of *c*P2 Promoter Region

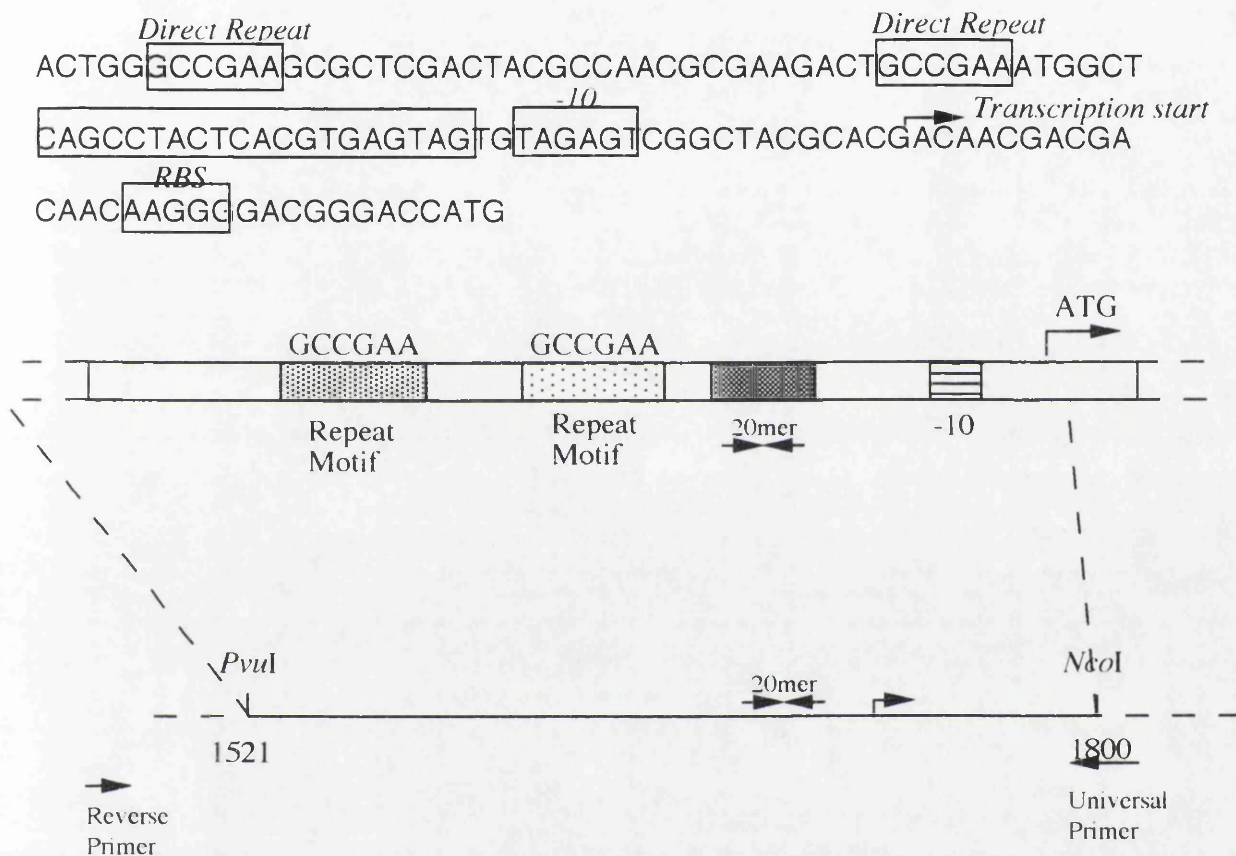


Figure 6.6. Diagrammatic Representation of the Fragments of the *c*P2 promoter DNA Used in DNA Binding Studies

6.2.1. Binding of the 42kDa Repressor Protein to the *c*P1 Promoter

Two DNA fragments containing the cP1 regulatory region were tested for DNA binding. The fragment *c*P1^{H3} includes the promoter *c*P1 and the '20mer' site (Figure 6.1). The other *c*P1 fragment used (*c*P1^{Nhe}) contains the promoter sequence, the '20mer', a direct repeat which overlaps the *Hin*dIII restriction site and the terminator (t^k) (Figure 6.1).

The band-shift assay was used to demonstrate the affinity of the 42kDa repressor for *c*P1^{H3} and cP1^{Nhe} promoter fragments (Figure 6.3 and 6.4). Serial dilutions of the 42kDa repressor protein were incubated with 6.6fmoles of labelled *c*P1^{H3} and 12.6fmoles cP1^{Nhe} using standard conditions (section 2.15.1). It can be noted that binding is first observed with 90nM of 42kDa repressor monomer (Figure 6.3). As the amount of protein is increased, virtually all free DNA is formed into a protein:DNA complex at 215nM repressor monomer. As more repressor is added the complex seems to be retarded further possibly due to non-specific interactions between repressor proteins. A second labelled band is observed in all lanes, which is believed to be an aberrant PCR product from the single end-labelled primer. This band does not appear to be retarded even at 208nM repressor monomer.

The band-shift experiment was repeated for *c*P1^{Nhe} using approximately 12.6fmoles of labelled probe (Figure 6.4). DNA binding is first observed at a concentration of 115nM repressor monomer and is complete at 325nM of the 42kDa monomer. At higher repressor concentrations complexes migrate less. The calculated dissociation constants for the 42kDa repressor to both fragments are presented in section 6.5.

6.3. *c*P2 Promoter

A blunt-ended 340bp *Pvu*I-*Nco*I fragment encompassing the *c*P2 promoter region was cloned into the polylinker of pUC19 using *Bam*HI linkers (p19B300; M.C.M Smith, personal communication). The *c*P2 fragment was amplified by the PCR using the pUC reverse and universal primers from either side of the cloned insert to generate a DNA fragment of 380bp (Figure 6.5). Within the *c*P2 promoter region is a sequence with good identity to a streptomycete -10 sequence downsteam of a predicted '20mer' motif (section 1.7.2), whilst upstream of the '20mer' is a directly-repeated sequence (Figure 6.6).

6.3.1. Binding of the 42kDa Repressor to the *c*P2 Promoter Fragment

Binding of the 42kDa repressor to 10fmoles of the *c*P2 promoter fragment was analysed in the band-shift assay (Figure 6.7). The formation of a DNA:protein complex was initially

Lane 1 2 3 4 5 6 7 8 9

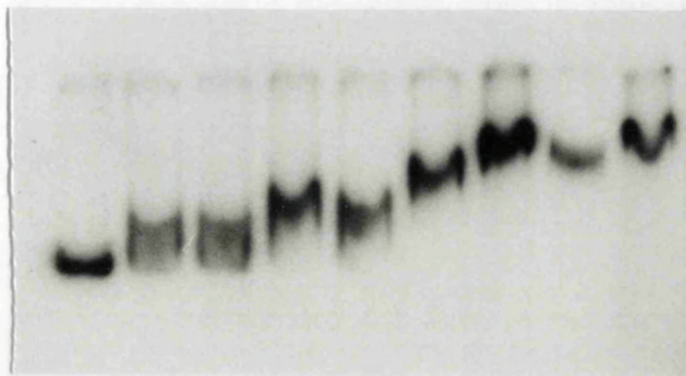


Figure 6.7. Affinity of the 42kDa Repressor Protein for the *c*P2 Promoter.

Lane	Repressor (nM)	Lane	Repressor (nM)
1	No Protein	6	161
2	45	7	208
3	59	8	321
4	89	9	982
5	119		

observed at a protein concentration of 60nM 42kDa repressor monomer. At 240nM monomer, no free DNA probe was visible and binding was assumed to be complete. The mobility of the retarded complexes is reduced further by adding high concentrations of protein. The reason for this is unclear, but a number of possible reasons exist and are discussed in section 6.11.

6.4. *c*P3 Promoter

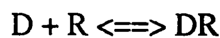
The putative *c*P3 promoter of the repressor gene lies just upstream of the start of the 42kDa repressor putative transcription start-point (Figure 1.5). The *c*P3 promoter was excised from the *c* gene on a 230bp *Ava*II-*Sal*I restriction digest and end-labelled using T4 polynucleotide kinase after pre-treatment with Calf Intestinal Alkaline Phosphatase (CIP) (section 2.8.2). The *c*P3 promoter contains half a conserved'20mer' motif found in *c*P1 and *c*P2 and two possible -10 sequences (Figure 1.5).

6.4.1. Binding of the 42kDa Repressor to the *c*P3 Promoter Fragment

An end-labelled 230bp fragment encompassing the putative *c*P3 promoter was used in a DNA-binding assay. The 42kDa repressor binds *c*P3 at 238nM repressor monomer (Figure 6.8) and is complete when no free probe is observed at 654nM 42kDa monomer. Calculations of the K_D for the 42kDa repressor with the *c*P3 fragment are described in section 6.5.

6.5 Calculation of the Dissociation Constant for the 42kDa Repressor With Promoter Fragments Isolated From the Repressor Gene

The dissociation constant (K_D) for each of the 42kDa repressor: promoter complexes was calculated as a relative measure of the affinity of the 42kDa repressor for the different binding sites. Binding of a protein to DNA sets up an equilibrium, free DNA and free protein are in equilibrium with a DNA:protein complex



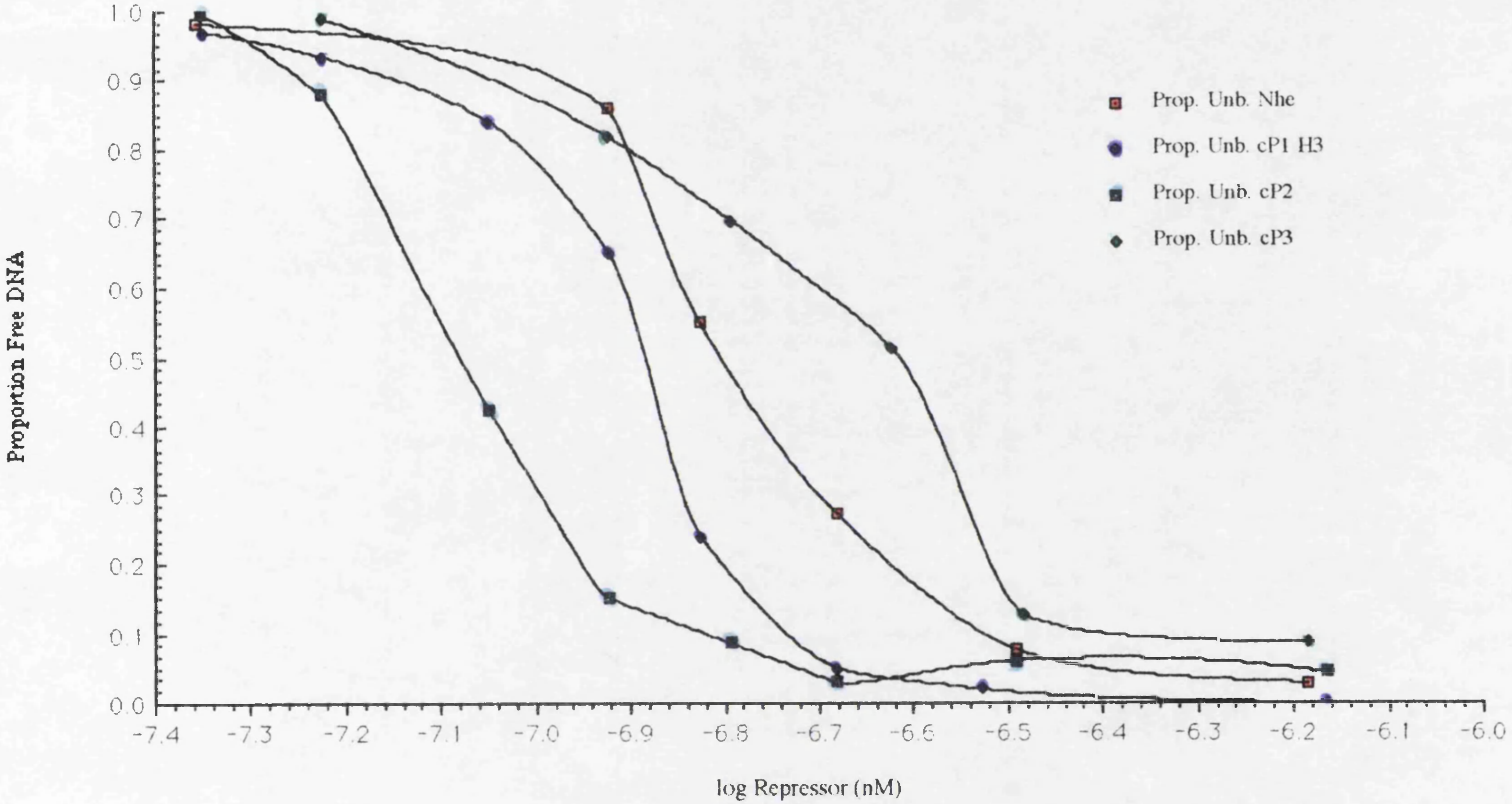
$$D + R \Longleftrightarrow DR$$

Lane 1 2 3 4 5 6 7 8

Figure 6.8. Affinity of the 42kDa Repressor for the *c*P3 Promoter..

Lane	Repressor (nM)	Lane	Repressor (nM)
1	No Protein	5	238
2	60	6	327
3	119	7	654
4	161	8	982

Figure 6.9. Binding Curves of Binding of 42kDa Repressor to c Gene DNA



Lane 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

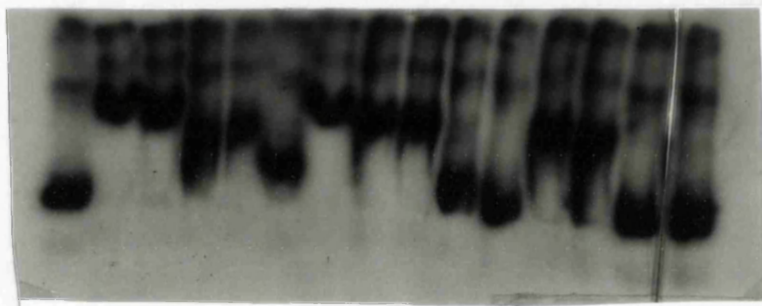


Figure 6.10. Competition for 42kDa Repressor Binding Incubated in the Presence of pUC19, p19H3-300 and p19B300.

Lane	Competitor	fmoles
1	No Protein	-
2	pUC19	169
3	pUC19	338
4	pUC19	507
5	pUC19	676
6	pUC19	1014
7	p19H3-300	57
8	p19H3-300	114
9	p19H3-300	228
10	p19H3-300	456
11	p19H3-300	912
12	p19B300	47
13	p19B300	94
14	p19B300	188
15	p19B300	376

Therefore, the dissociation constant, K_D is

$$K_D = \frac{(D)(R)}{(DR)} \quad (1)$$

where D= DNA Binding Site
R= Repressor
DR= Repressor:DNA Complex
$(R)_o$ = Repressor Conc. at t=0

When the repressor is in excess, then (R) approaches $(R)_o$. Thus, when 50% of DNA is unbound:

$$(D) = (DR), \text{ so } K_D = (R) \sim (R)_o \quad (2)$$

The dissociation constant is approximately equal to the repressor concentration at which 50% of the DNA is unbound. This data may be plotted from the equation derived from equation (1).

$$\frac{(D)}{(DR)} = \frac{(K_D)}{(R)} \quad (3)$$

A plot of the proportion of total DNA unbound against repressor concentration enables one to derive the K_D.

The data were derived from band shift assays described in section 6.4. Bands corresponding to free and bound DNA were isolated from a dried gel and Cerenkov counted for two minutes. The K_D derived from this experiment assumes that the repressor protein is a monomer and is 100% active, and as such is a gross overestimate. It is believed from data presented in section 5.4 that the active form of the 42kDa repressor is tetrameric, however, other forms do exist, which constitute a relatively large proportion of total protein. The data was plotted using the proportion of free DNA against log protein concentration (Figure 6.9; Carey, 1991).

From the graphs it is apparent that each of the four fragments used in the band shift assay had varying affinities for the 42kDa repressor (Figure 6.9). The calculated affinities for each of the fragments show that the *c*P2 promoter possesses the greatest affinity for the 42kDa repressor with a K_D of 85nM, however, it should be noted that retardations produced by the cP2 fragment do not form a stable bandshift, but a smear typical of dissociated protein-DNA complexes at low repressor concentrations presumably because of the large size of the *c*P2 PCR product destabilizing the complex during electrophoresis (Table 6.1). Both forms of the *c*P1 promoter bound the 42kDa repressor; with an affinity for $cP1^{H3}$ at 130nM and $cP1^{Nhe}$ at 155nM (Table 6.1). The *c*P3 fragment bound with relatively low affinity for the 42kDa repressor at 240nM protein.

Figure 6.11. Competition for 42kDa Repressor Binding to cP1-H3

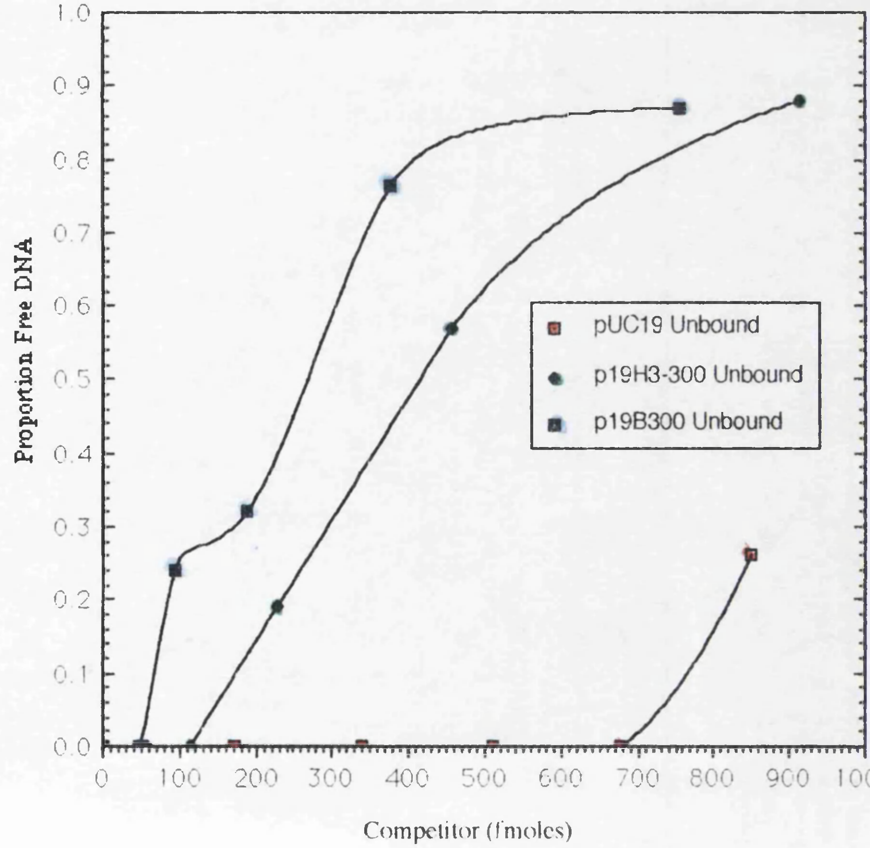
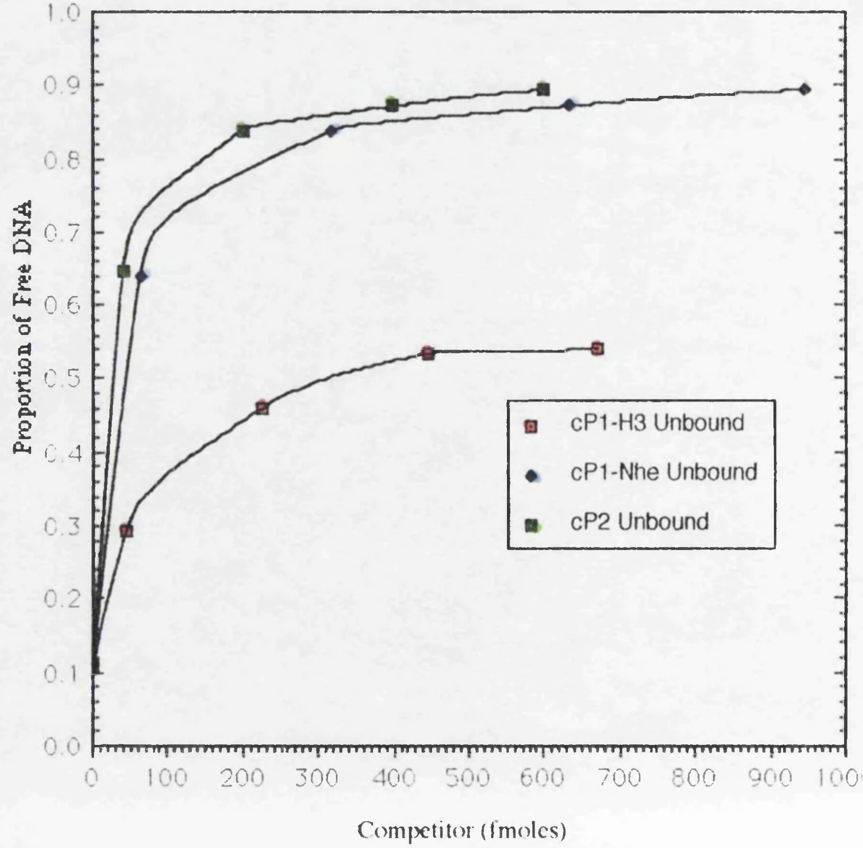


Figure 6.13. Competition for 42kDa Repressor Binding to cP1-Nhe



Lane 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

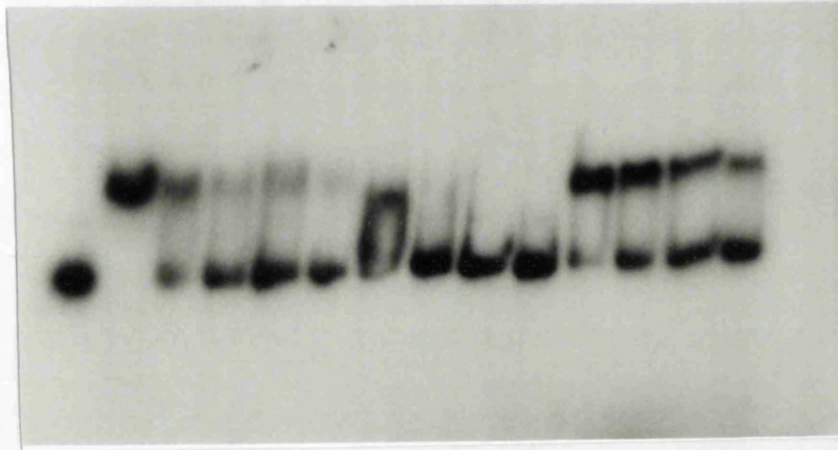


Figure 6.12. Competition for 42kDa Repressor Binding in the Presence of Linear Fragments of *c*P1-H3, *c*P1-Nhe and *c*P2.

Lane	Competitor	Competitor (fmoles)	Lane	Competitor	Competitor (fmoles)
1	No Protein	-	9	*c*P2	500
2	No Competitor		10	*c*P2	750
3	*c*P1-Nhe	50	11	*c*P1-H3	30
4	*c*P1-Nhe	250	12	*c*P1-H3	150
5	*c*P1-Nhe	500	13	*c*P1-H3	300
6	*c*P1-Nhe	750	14	*c*P1-H3	450
7	*c*P2	50			
8	*c*P2	250			

6.6. Specificity of Binding by the 42kDa Repressor Protein to *c*P1 and *c*P2

To investigate the specificity of binding of the 42kDa protein for the promoter fragments, DNA binding was competed away with unlabelled DNA encoding the *c* gene promoters. It should be noted that in addition to the competitor DNA, 150-350-fold weight excess of cold pUC18/19 non-specific DNA was added routinely to eliminate the majority of non-specific protein-DNA interactions. The order in which each of the components of the competition experiments were added was invariant as competitor DNA was added first to the reaction (both pUC19 and c gene promoters), followed by the repressor. This was left for approximately 3-4 minutes before the addition of labelled probe DNA and incubation at room temperature for 15 minutes.

6.6.1. Specificity of Binding of *c*P1^{H3} with the 42kDa Repressor

325nM 42kDa repressor monomer was added to 6.6fmoles of the *c*P1^{H3} fragment in the presence of competitor DNA's . Three different competitors were used; pUC19 as a non-specific competitor control, p19H3-300 which contains the *c*P1^{H3} fragment and p19B300 which contains the *c*P2 fragment. All competitors were added in the form of supercoiled plasmids. Addition of pUC19 to 1014pmoles competed binding yielding *c*P1^{H3} DNA with a mobility very similar to that of free DNA.

456fmoles (69 fold excess) of supercoiled p19H3-300 competed out binding such that greater than 55% of labelled DNA probe was free whilst binding was virtually abolished with the addition of 912fmoles p19H3-300. This competition assay suggests that an excess of 42kDa repressor was added (i.e. greater than was needed for 100% binding to the *c*P1^{H3} fragment) and that binding of the 42kDa repressor was specific for the *c*P1^{H3} promoter fragment.

Competition of cP1^{H3}-42kDa protein complexes with p19B300 proved to be more effective than p19H3-300. Addition of 188fmoles of p19B300 virtually competed out binding of the *c*P1^{H3} promoter. The *c*P2 and *c*P1^{H3} promoter fragments are more effective competitors for binding by the 42kDa protein than pUC19 (Figure 6.11). By comparing the number of pmoles required to compete 50% binding the *c*P2 promoter appears to have a 2-fold greater affinity for the 42kDa repressor than the *c*P1^{H3} promoter (Figure 6.11).

6.6.2. Specificty of Binding of the 42kDa Repressor to the *c*P1^{Nhe} Promoter Fragment

Known quantities of the PCR fragments of the *c*P1^{Nhe}, *c*P1^{H3} ; and *c*P2 promoter regions (Figure 6.2 and 6.4) were added as described in section 6.6 to a DNA:protein complex

containing 240nM of the 42kDa repressor monomer and 12fmoles of the labelled *c*P1^{Nhe} fragment (Figure 6.12).Addition of *c*P2 promoter fragment was found to be a more effective competitor than either *c*P1^{Nhe} and *c*P1^{H3}, whilst *c*P1^{Nhe} was more effective than *c*P1^{H3}. These results show that 50% of binding is competed away by the addition of 3-fold molar excess *c*P2, a 4-fold molar excess of *c*P1^{Nhe} and a 13-fold molar excess of *c*P1^{H3} (Figure 6.13).

6.7. DNaseI Footprinting of the *c*P1^{Nhe} Fragment With the 42kDa Repressor Protein

An *Ava*II-*Taq*I fragment was labelled at the *Ava*II 5' end with T4 polynucleotide kinase (section 2.8.2). The footprinted region contains all the features of the *c*P1^{Nhe} fragment used in the band-shift assay including the '20mer' motif, a putative direct repeat and the terminator $\{t_k\}$(Figure 6.1). Unfortunately the footprinting data does not provide a clear result. Some protection was observed on the coding strand (Figure 6.14). Possible sites of DNaseI protection and cutting enhancement are marked on Figure 6.15.

6.8. The 42kDa Repressor Binds to a Site in the Early Region

The early lytic genes are normally repressed during lysogeny, presumably by one or more of the repressor proteins. The region chosen for study was a 32bp fragment (CIFC) which lies near to a region which is inverted in a virulent mutant of ØC31 (M.C.M Smith, personal communication). Our working hypothesis states that the rearrangement in the virulent mutant has resulted in the constitutive expression of an activator of lytic promoters. Hence the rearranged region may well contain repressor binding sites. Moreover, this region has been sequenced (Chater *et al*, unpublished data) and found to contain two putative '20mer' motifs bis ected by a transcriptional terminator (C.J. Ingham and M.C.M Smith, personal communication). In addition to these motifs promoter activity has been found downstream, proximal to the 3' end of the fragment used in this study (C.J. Ingham and M.C.M Smith, personal communication).

This region was used in DNA binding studies using the gel retardation assay and the 42kDa repressor protein to demonstrate if DNA binding activity by the 42kDa repressor was associated with the '20mer' motif.

6.8.1. The 42kDa Repressor Binds to the CIFC Early Region

The 328bp labelled CIFC fragment was synthesised by the PCR using the CI and FC oligonucleotides as primers (Figures 6.16 and 6.17). The CIFC fragment was incubated with increasing quantities of the purified 42kDa repressor protein. As little as 30nM of purified repressor protein was found to be enough to demonstrate binding (Figure 6.18).

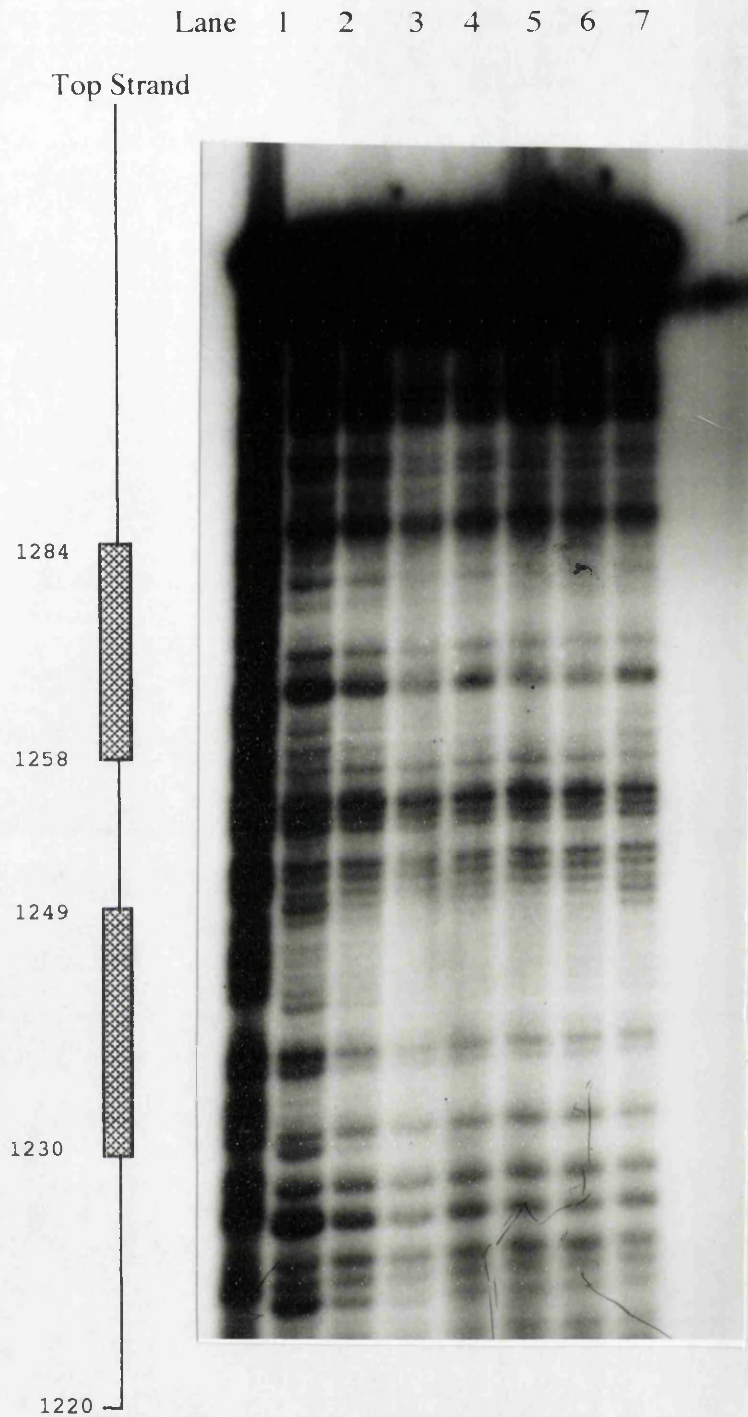
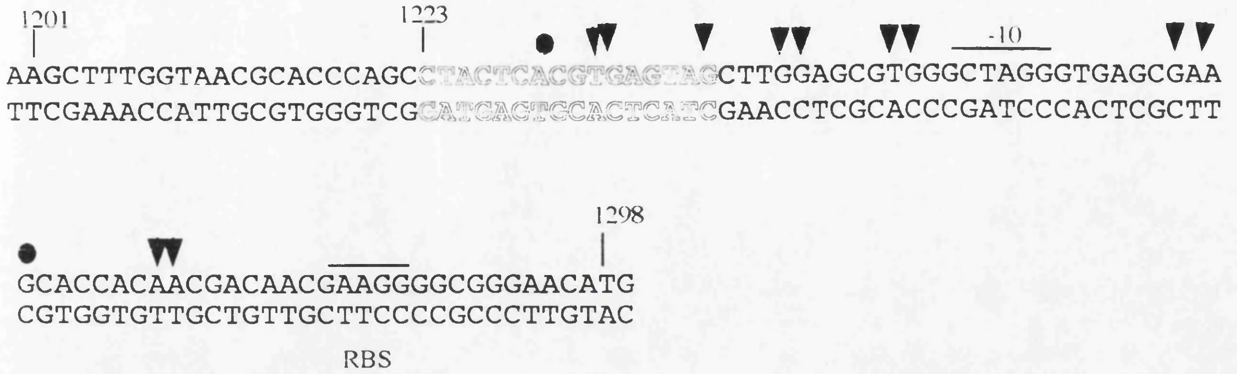


Figure 6.14. DNaseI Footprinting of the 42kDa Repressor Bound to the *c*P1 Promoter Shaded Areas Designate Regions Protected from DNaseI Cleavage.

Lane	Protein (ng)
1	A+G Ladder
2	No Protein
3	100
4	200
5	275
6	550
7	825

Figure 6.15. Diagrammatic Representation of DNaseI Footprint Data of 42kDa Repressor and cP1 Promoter



DNaseI Protected Band

Reduced Cleavage

The amount of unbound DNA decreased until no free probe was observed in the presence of 238nM of the repressor. Upon the addition of 238nM of repressor a putative second complex appeared on the band-shift assay, although not clear in the bandshift assay shown, was clearly visible on a shorter exposure. To investigate this fragment more thoroughly, each half of the CIFC fragment was isolated by PCR, labelled and analyzed in band-shift assays. These fragments, termed CIGC and CHFC refer to the PCR primers used to generate each fragment (Figure 6.17).

6.8.2. Binding of the 42kDa Repressor with Each Half of the CIFC Early DNA Fragment

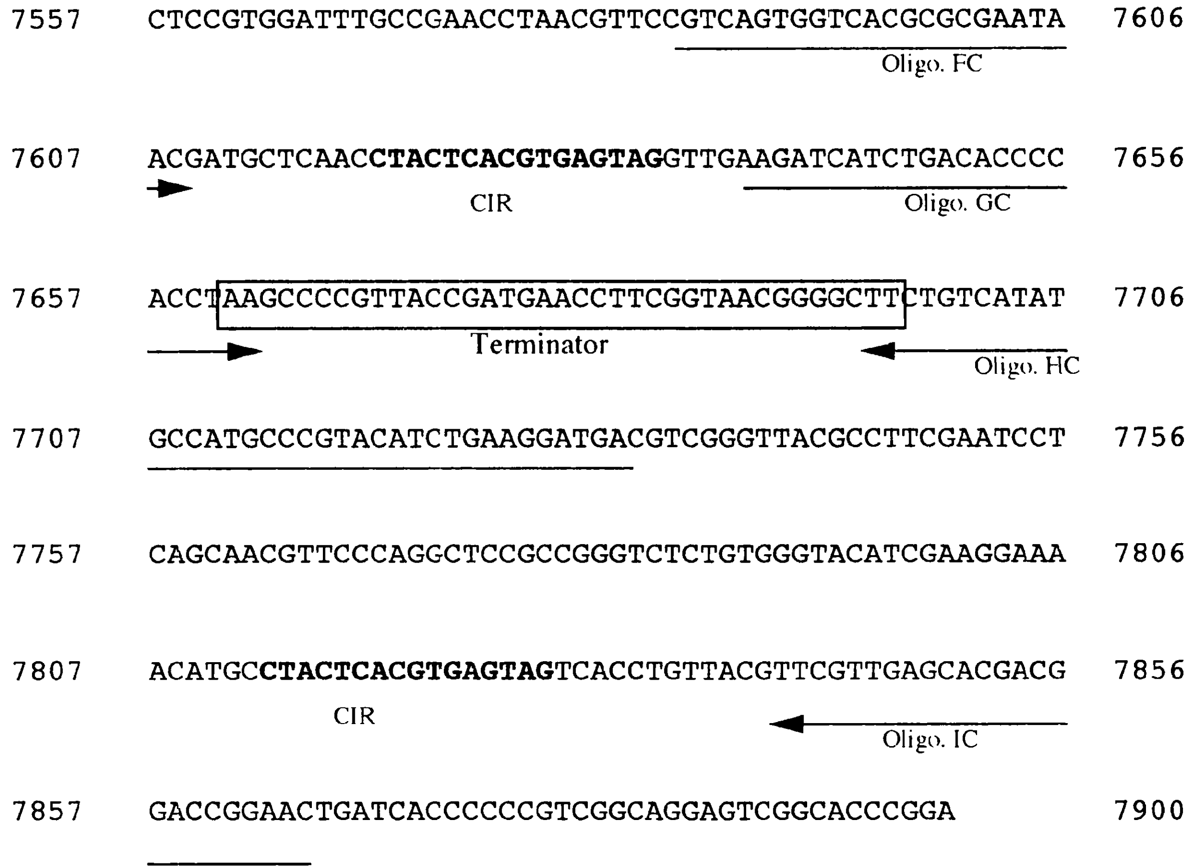
The CIGC fragment was produced by the PCR between the CI and GC primers to produce a fragment of 225 base pairs in length. Similarly, the CHFC fragment was produced by the PCR between the CH and FC complementary oligonucleotides and is 138 base pairs in length.

Figure 6.19. demonstrates the binding of the 42kDa repressor for the CIGC fragment from the early region of ØC31. As the protein concentration was increased, binding to the promoter was observed initially at 120nM repressor. At 161nM repressor, about 90% of free DNA was bound in a complex with the repressor protein. The identical experiment was repeated for the CHFC region, which contains the complementary half of the CIFC region to the CIGC fragment (Figure 6.20). The experiment demonstrates that binding is only observed in the presence of 327nM of repressor (Figure 6.20). About 90% of the CHFC fragment is bound in the presence of 982nM of the 42kDa protein. These data imply that both the CIGC and CHFC fragments contain a binding site for the 42kDa repressor, one of which possesses a greater affinity for the protein than the other.

6.9. Calculation of the Dissociation Constant for the CIFC, CIGC and CHFC DNA Fragments

The free DNA and bound DNA was cut out of the filter, counted and used to plot a graph of free DNA against repressor concentration (Figure 6.21). The dissociation constant for

Figure 6.16 Early Region Intergenic Sequence Used in DNA Binding Studies



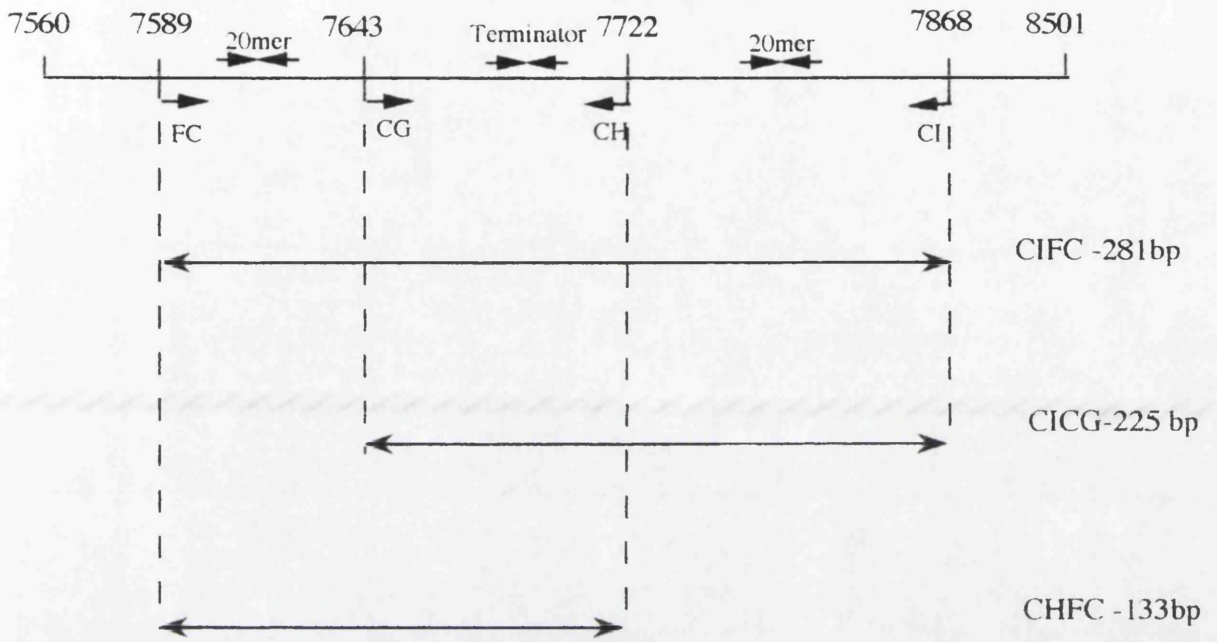


Figure 6.17. Early Promoter Sequences Used for DNA Binding Studies

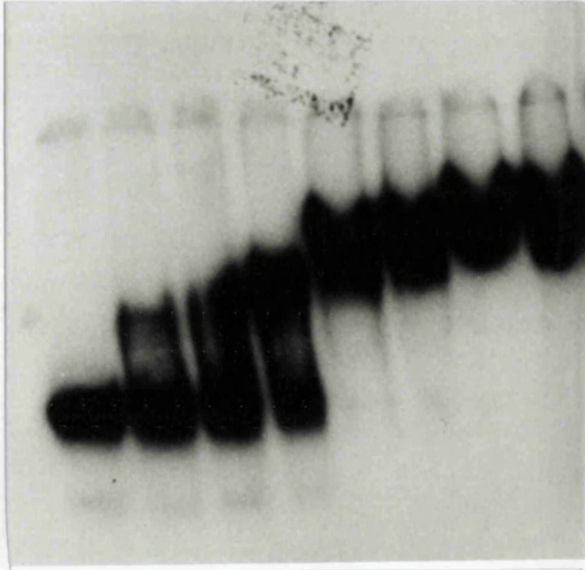


Figure 6.18. Band-Shift Assay Demonstrating Affinity of the 42kDa Repressor for the CIFC Early Gene Region.

Lane	Repressor (nM)	Lane	Repressor(nM)
1	No Protein	5	240
2	30	6	327
3	60	7	654
4	120	8	981

Lane 8 7 6 5 4 3 2 1

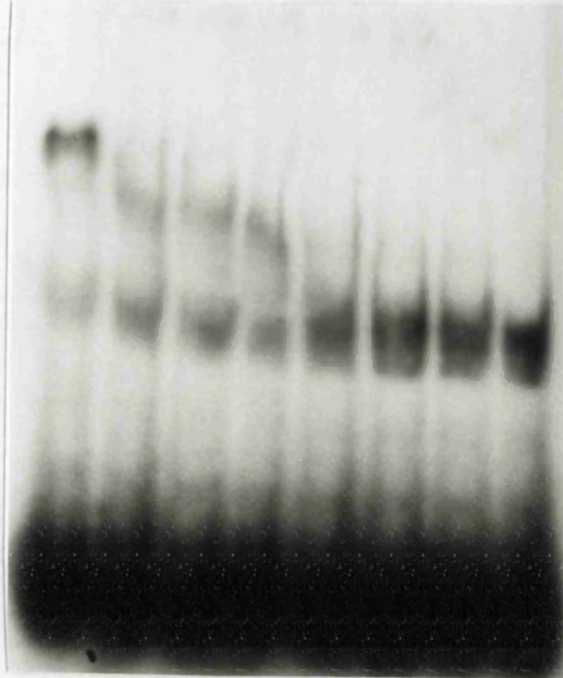


Figure 6.19 Band-Shift Assay Demonstrating Affinity of the 42kDa Repressor for the CIGC Early Gene Region.

Lane	Repressor (nM)	Lane	Repressor (nM)
1	No Protein	5	161
2	30	6	240
3	60	7	322
4	120	8	480

Lane 1 2 3 4 5 6 7 8

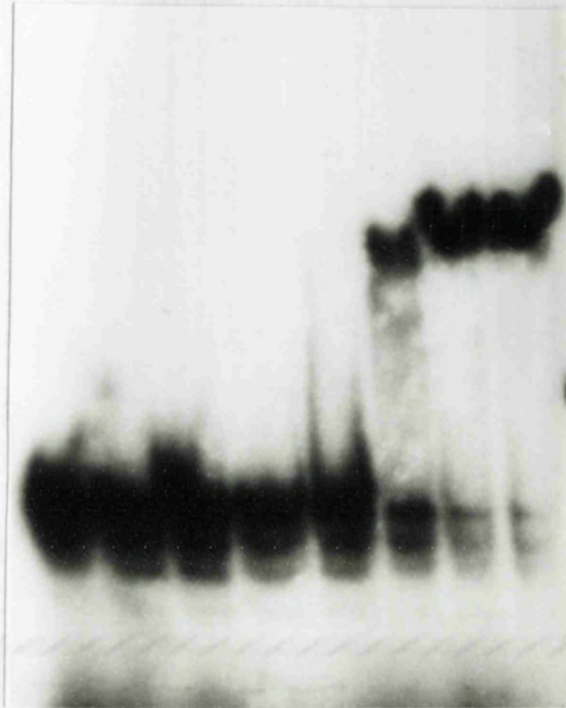


Figure 6.20. Band-Shift Demonstrating Affinity of 42kDa Repressor for the CHFC Early Gene Region.

Lane	Repressor (nM)	Lane	Repressor (nM)
1	No Protein	5	327
2	60	6	654
3	120	7	981
4	240	8	1308

the three fragments from the early region of ØC31 were determined (Table 3.1). The K_D stated for each of the fragments assumes that the 42kDa repressor is a monomer and is 100% active. From these data it can be observed that the CIFC fragment possesses a K_D of 112nM, the CIGC fragment a K_D of 398nM of the 42kDa repressor monomer and the CHFC fragment a dissociation constant of 794nM monomeric repressor. These values are relatively large and indicate that each half of the CIFC fragment in isolation have relatively low affinities for the 42kDa repressor.

6.10. Competition for the 42kDa Repressor Binding to the CIFC Early Region

A competition experiment was performed using labelled CIFC DNA and the 42kDa repressor to investigate the specificity of 42kDa repressor binding. Binding was competed with unlabelled PCR products of the CIFC, CIGC and CHFC regions (Figure 6.22).

It is apparent that CIFC and CIGC are more effective competitors than the CHFC fragment (Figure 6.23). Addition of up to 30-fold excess CHFC fragment did not appear to perturb binding of the 42kDa repressor to the CIFC region. The addition of 4-fold excess of the cold CIFC or 15-fold excess of the CIGC fragments compete out binding by 50%. These data correlated well with the dissociation constant data in that the CIFC fragment was found to bind the 42kDa repressor far more effectively than the CIGC fragment, which is itself far more effective than the CHFC fragment.

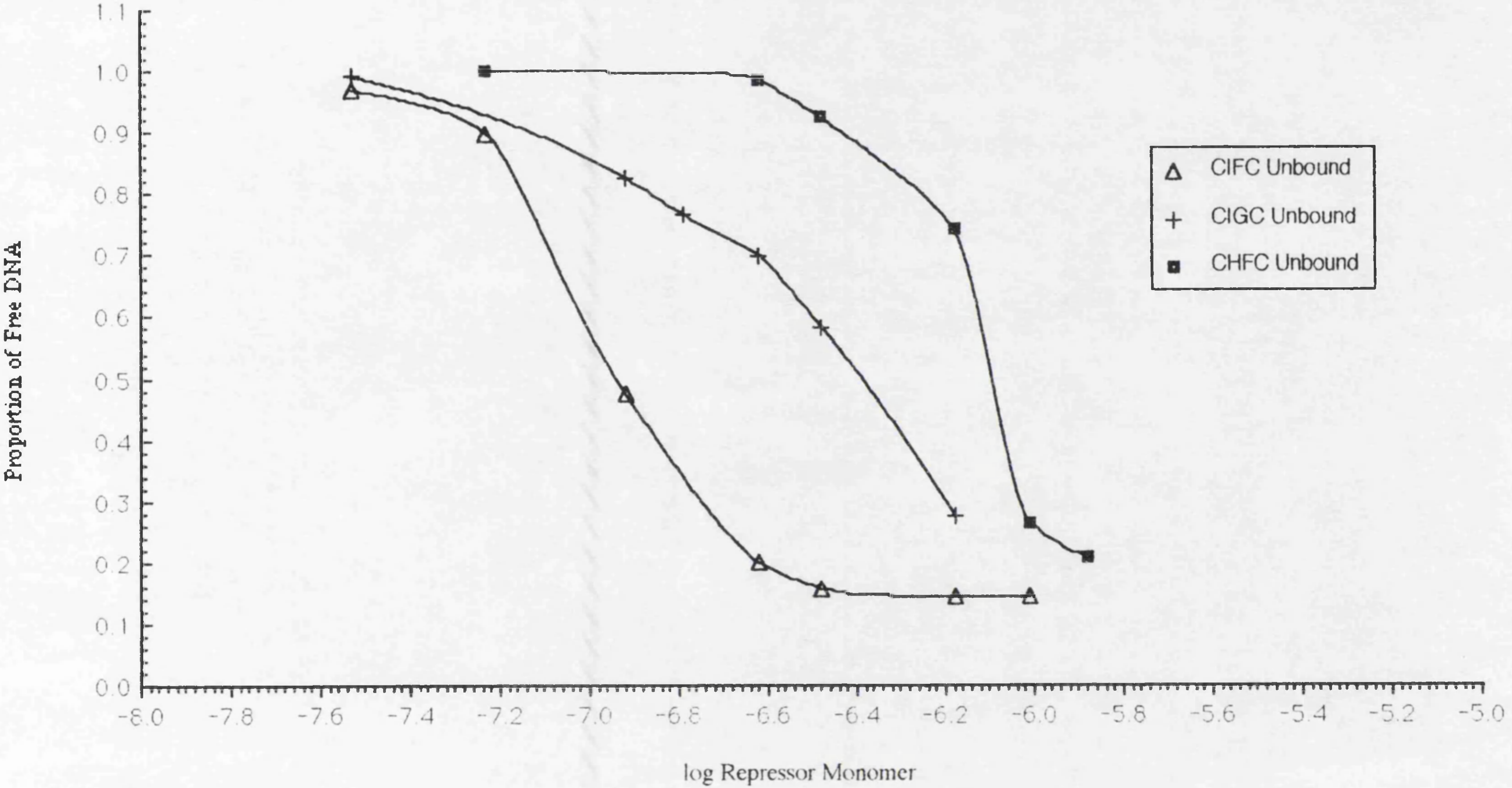
6.11. Conclusions and Discussion

Binding by the purified 42kDa repressor protein to promoter regions from the ØC31 *c* gene and to a putative regulatory region within the early region was investigated. DNA binding was demonstrated using the band shift assay and by preliminary *in vitro* DNaseI footprinting.

Bandshifts were only observed when the pH of the buffer system was lowered to a slightly acidic pH of 6.6. These observations suggest that charge of the 42kDa protein plays an important role in the mobility of the complexes, however, it may not be the only factor involved. The Lac repressor of *E.coli* has an acidic pI but retards complexes seemingly independent of charge. In this case the relatively large mass of the Lac repressor (150kDa) is thought to be influential (Fried and Crothers, 1981; Lane *et al*, 1992).

Seven different regions have been investigated for their ability to act as binding sites for the 42kDa repressor. The most thorough investigation has been of the interaction of the

Figure 6.21. Affinity of 42kDa Repressor for Early Gene Region



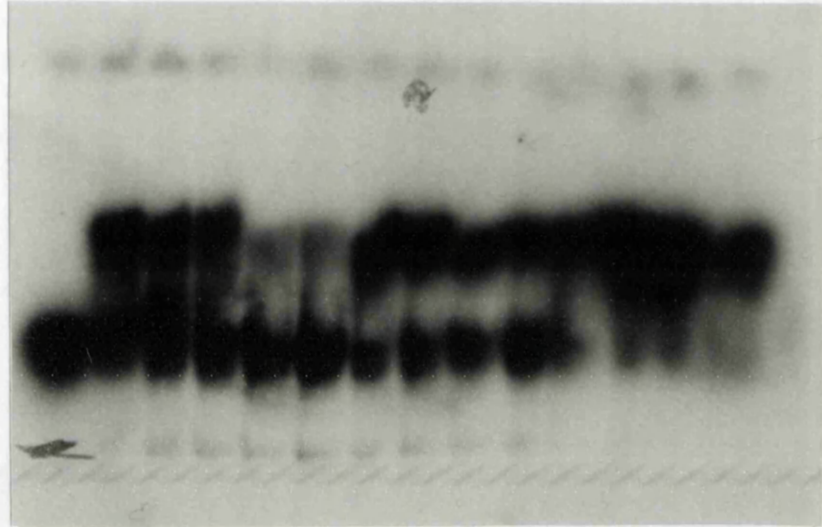


Figure 6.22. Band-Shift Assay Demonstrating the Specificty of Binding of the 42kDa Repressor for the CIFC Fragment.
CIFC DNA (11.5fmoles) was incubated with 119nM of the 42kDa repressor monomer. Cold PCR fragments of regions from the early gene were incubated with the protein-DNA complex.

Lane	Competitor	Competitor (fmoles)	Lane	Competitor	Competitor (fmoles)
1	No Protein	-	8	CIGC	150
2	No Competitor	-	9	CIGC	300
3	CIFC	46	10	CIGC	450
4	CIFC	230	11	CHFC	38
5	CIFC	460	12	CHFC	190
6	CIFC	690	13	CHFC	380
7	CIGC	30	14	CHFC	570

42kDa repressor with the *c*P1 promoter. The sequence and features distinctive of each of the promoters have been described in Chapter 1. Overall, the promoters with the greatest affinity for the 42kDa repressor appear to be the *c*P2 and the CIFC region. The dissociation constant (K_d) for the *c*P2 and CIFC regions are 85 and 112nM respectively when considered for a monomer binding with 100% activity. From the calculation of the dissociation constant one is able to construct a hierarchy of promoters with the greatest affinity. From these data, it is evident that region CHFC and *c*P3 possess the least affinity for the 42kDa repressor (Table 6.1).

The ability of various fragments to compete for 42kDa repressor binding supported the relative affinities. Cold fragments of each of the promoters were used at increasing concentrations to compete a protein-DNA complex comprising of the 42kDa repressor and the *c*P1^{H3}, *c*P1^{Nhe} or CIFC DNA's. Competition with known amounts of specific DNA fragments enabled an estimate to be made of the relative affinity of a protein for a specific region of DNA compared to a non-specific site.

Competition for 42kDa repressor with labelled *c*P1^{H3} and increasing pUC19 showed that up to 850fmoles of pUC19 could be added without affecting the affinity of the 42kDa repressor for *c*P1^{H3} DNA. Some change in mobility of complexes were observed, possibly due to competetion of non-specific binding. Free DNA was not observed although a complex with mobility very close to free DNA was. Up to 1pmole of pUC19 could be added before the enhanced mobility DNA became visible (Figure 6.11). Competition with p19H3-300 and p19B300 proved a great deal more effective than competition with pUC19. Addition of 456fmoles of p19H3-300 competes out over 50% of binding to free DNA, whilst the addition of 376fmoles p19B300 eliminates virtually 80% of bound complex. From the graph of this experiment it can be estimated that 265fmoles p19B300 and 410fmoles p19H3-300 are required to compete out binding by 50%. This preliminary data implied that the *c*P2 promoter possessed a binding site with greater affinity for the 42kDa repressor than the *c*P1^{H3} promoter. However, the following caveat should be borne in mind when considering this data. No distinct complex was observed for the *c*P2 promoter in the presence of low amounts of the 42kDa repressor; a smear between a complex of slightly less mobility and the free DNA was observed. This was probably caused by the relatively large *c*P2 promoter DNA fragment possibly destabilising the protein-DNA complex when migrating through the gel.

By comparison, a competition experiment using *c*P1^{Nhe}-42kDa protein complexes and adding linear fragments of *c*P1^{Nhe}, *c*P1^{H3} and *c*P2 promoter competitor DNA demonstrated that *c*P2 possesses the greatest affinity for the 42kDa repressor. A total of 28fmoles of *c*P2 fragment was required to compete out binding of the 42kDa repressor to

cp1	1215	CCAGC**CTACTCACGTGAGTAG**CTTGGAGCGTGGCTAGGGTGAGCGAA***GCACCACAACGACAACGA***AGGGGCGGGAAC
cp2	1790	TCAGC**CTACTCACGTGAGTAG**TGTAGAGTCGGCTAC***GCACGACAACGACGACAACAA***GGGGACGGGCACC
cp3	2116	TGAGGTTCCGGGCC**TGAGTA**TGACGGTCAGTCAGAAGAGCTACACGCCCGTTGAGGG***CGACGACAAG***GGAGAGAAC
CHFC	7614	TCAAC**CTACTCACGTGAGTAG**GTTGAAGATCATCT***GACACCCCAC***CTAAGCCCCGTTACCGATGAACCTTC
CIGC	7808	CATGC**CTACTCACGTGAGTAG**TCACCTGTTACGTTCGTTGA***GCACGAC***GGACCGGAA

Figure 6.23. Alignments of Putative 42kDa Repressor Binding Sites.
Bold letters indicate bases forming part of the 20mer motif, italicised letters relate to a moderately conserved motif, which possibly plays a role during 42kDa repressor binding.

Table 6.1. Dissociation Constant of the 42kDa Repressor for the c Gene Promoter Fragments

Promoter	Repressor Concentration at 50% Binding	Kd
cP1Nhe	170ng	155nM
cP1H3	270ng	130nM
cP2	84ng	85nM
cP3	370ng	240nM
CIFC		112nM
CIGC		398nM
CHFC		794nM

The Kd calculated from this data asssumes that the 42kDa repressor is a monomer and is 100% active. The protein concentration was measured using the Bradford assay.

*c*P1^{Nhe}, compared to 43fmoles of *c*P1^{Nhe} itself and 310fmoles of the *c*P1^{H3} fragment. The *c*P2 fragment is 1.5 times better than the *c*P1^{Nhe} fragment at competing for binding of the 42kDa protein and is consistent with the observed difference in the dissociation constant between *c*P1^{Nhe} and *c*P2 (1.9 fold)(Table 6.1). What is rather less consistent is the difference between *c*P1^{Nhe} and *c*P1^{H3} , and *c*P2 and *c*P1^{H3} (Table 6.1). The competition assay values are 3-fold greater than those observed in the prediction of the dissociation constant. The cP1^{Nhe} competition assay also provides different data to that observed in the cP1^{H3} competition experiment, particularly for cP2 which is approximately 9-fold less adept at competing the labelled cP1^{Nhe} promoter fragment than the cold cP1^{Nhe} fragment. The competition assay with cP1^{H3} used supercoiled plasmid DNA as a substrate rather than linear promoter fragments. This could possibly be to blame for the decreased affinity for supercoiled DNA than linear DNA by the 42kDa repressor. Although this is considered unlikely given that under physiological conditions the repressor is likely to have supercoiled phage DNA integrated into the *Streptomyces* genome as its natural binding substrate. Another alternative is that excess 42kDa repressor was added in the cP1^{H3} competition experiment so distorting the assay. A shoulder of 100% binding is observed on the graph possibly indicating that saturating levels of the 42kDa repressor were present. The large excess of non-specific pUC19 plasmid DNA may also have played a significant role in distorting this data.

The competition experiment using *c*P1^{H3} as a labelled substrate demonstrated that binding of the 42kDa repressor to *c*P1 DNA is specific and that non-specific DNA such as pUC19 was unable to compete out binding even at relatively high concentrations (5.6pM). The competition assay using*c*P1^{Nhe} as a labelled target provides a demonstration of the relative affinities of the different *c*P1 and *c*P2 promoter DNA fragments for binding the repressor. These data relate quite well to the dissociation constant data and demonstrate that the *c*P2 promoter is the best binding substrate in the repressor gene, followed in order by the *c*P1^{Nhe} , *c*P1^{H3}and *c*P3 promoters (Table 6.1).

Preliminary DNaseI footprinting data suggested that the region of DNA between the *Hin*dIII and *Nhe*I sites may not be contacted by the repressor, but that the true binding site lies downstream and partially overlaps the putative 20mer motif. However, these data should be taken as a preliminary indication as full protection was not observed.

Binding experiments shown in Figures 6.9 and 6.10 show that at higher repressor concentrations, non-specific DNA interactions are made. These may be made by binding of repressor molecules to non-specific stretches of DNA, or due to protein-protein interactions and aggregation. This has also been observed for the Trp repressor, where it was found that at slightly higher repressor concentrations than were needed for complete binding, more than one repressor bound non-specifically (Carey, 1988). It was observed

that more repressors were present in stoichiometrical experiments at higher protein concentrations, than actual possible binding sites on the DNA indicating that protein-protein interactions may play a role in binding of the Trp repressor *in vivo* (Carey, 1988).

The data presented above allow a prediction to be made regarding the activity of the repressor protein. It should be remembered that all protein concentration values quoted in this thesis are those for a repressor monomer that possess 100% activity. Data presented in Chapter 5 suggested that physiological binding form of the repressor is a tetramer.

Binding has been demonstrated to an inter-genic region within the 'early' region of the phage genome. The exact function of this region is not clear, but it does possess a putative transcriptional terminator originally mapped in high resolution S1 transcription studies (C. Ingham, personal communication). Promoter activity has also been detected from this region using a kanamycin resistance assay (C. Ingham, personal communication). The region lies at the boundary of a 7kb fragment of the phage genome which becomes inverted in a virulent mutant of ØC31(M.C.M Smith, personal communication). In addition, the region possesses two '20mer' motifs.

The dissociation constant has been calculated for this entire region, and for each individual half (Figure 6.16). These data demonstrate that the entire CIFC region possesses approximately the same affinity for the 42kDa repressor as the *c*P2 promoter. The DNA fragment was divided into two using the PCR. Each half of the CIFC fragment was found to bind the 42kDa repressor but each possessed a lower affinity for the repressor than the entire CIFC fragment. The CIGC fragment possessed a K_D of 407nM, roughly equivalent to that of the $cP1^{H3}$ promoter, whilst CHFC possessed a much lower affinity and a dissociation constant of approximately 794nM. These data were confirmed in a competition assay, whereby binding of the CIFC fragment by the 42kDa repressor was competed out effectively by cold CIFC. In contrast, only partial competition was observed with the CIGC fragment and no competition was observed at the concentrations used for the CHFC fragment. This is discussed further in Chapter 7 where recent data has confirmed these observations. Thus, from these data we can conclude that the 42kDa repressor binds to possibly two regulatory elements within this part of the early region of the phage genome. These regions have been putatively identified and are presented in Figure 6.23.

The binding site of the 42kDa repressor was preliminarily identified using DNaseI footprinting of the $cP1^{Nhe}$ fragment. This experiment demonstrated that a region downstream of the putative 20mer extending for 35base pairs, overlapping the predicted -10 promoter sequence was probably protected in the presence of the 42kDa repressor.Alignments of all the promoter regions shown to bind the 42kDa repressor

demonstrate that there is at least partial sequence conservation of the region observed to be bound by the repressor in the DNaseI footprinting assay between all these regions (Figure 6.22).

Footprinting demonstrates that the 20mer motif plays at least a partial role in binding and that sites in the 3' half of the inverted-repeat are protected from DNaseI. These data taken together with the band-shift data and the alignments lead to a possible conclusion that part of the binding site for the 42kDa protein is the 20mer, but the 20mer on its own provides relatively weak affinity for the 42kDa repressor (c.f. CHFC K_D=794nM) and also that regions which do not possess an entire 20mer motif also have the ability to bind the 42kDa repressor (*c*P3).

It should be noted that the experiments described in this chapter describe the binding of a single repressor to regions of promoter DNA. The situation *in vivo* is likely to be different given that at least 3 putative repressor proteins are produced from the *c* gene, which are likely to interact with each other when binding to DNA. This interaction is likely to increase the affinity of the 42kDa repressor for its binding site. The inter-play between each of the repressor proteins is vital to our understanding of the regulation of gene expression in ØC31 and will be an important feature of future work. The study of the 42kDa repressor has demonstrated that it possesses the ability to bind DNA with high affinity and that a putative binding motif has been identified based on DNaseI footprinting data and sequence analysis of regions shown to bind the repressor.

Chapter 7

General Discussion

Streptomyces lividans containing the 42kDa repressor protein expressed from plasmid pMS201 is not fully immune to superinfection by ØC31. In contrast a strain containing plasmid pMS231, that expresses both the 54 and 42kDa proteins is immune. Consequently it has been suggested that the 54 and 42kDa proteins are the true repressors for lysogeny. Repression by the 42kDa protein alone has been demonstrated in *E.coli* where expression of *lacZ* in a *c*p1-*lacZ* fusion is repressed in the presence of compatible plasmids expressing either the 42kDa protein alone or the 42 and 54kDa proteins. On the basis of these data, DNA-binding by the purified 42kDa protein alone was investigated, particularly using promoter *c*p1 as a target. The 42kDa repressor protein was overexpressed in *E.coli* , purified to near homogeneity and partially characterized biochemically. The pure 42kDa protein binds to at least five DNA-binding sites as tested using the DNA bandshift assay.

It is probable that the 54 and 42kDa repressors interact to allow ØC31 to undergo lysogeny. This interaction could either take the form of homo-oligomers of the 54 and 42kDa repressors co-operating via protein-protein interactions to regulate transcription, or alternatively they could combine to form a hetero-oligomer. Some tentative evidence supports the possible existence of a hetero-oligomer. During purification trials of the 54 and 42kDa proteins on a DNA-cellulose column at a relatively high salt concentration, greater than 95% of protein loaded onto the column eluted in the flow-through. However, analysis of fractions throughout the elution profile demonstrated that approximately equal levels of the 54 and 42kDa repressor protein were present in three fractions eluting at 300mM salt, even though unequal levels were present in the original extract. These extracts were derived by heterologous expression in *E.coli* and may not be representative of the situation in *Streptomyces* species.

The gene organisation of the ØC31 repressor locus tempts one to speculate on mechanisms by which the genes are co-ordinately regulated. It is known from Western blotting studies of *S.coelicolor* A3(2) infected with ØC31that relatively large, approximately equal quantities of the 42 and 54kDa repressor proteins are produced, whereas only small amounts of the 74kDa repressor are observed. One such mechanism of gene regulation could involve DNA looping *c.f.* arabinose operon in *E.coli* (Dunn et al, 1984). Given the multiplicities of CIR's in ØC31 (17-

20 at the last count!), repressor-DNA interactions may allow formation of a complex array of repression loops.

Several examples of enforced structural change as the result of protein binding have now been documented. The earliest studies of DNA looping came from studies of the arabinose (*ara*) operon in which distantly-spaced sites are brought into close proximity through the action of DNA looping (Dunn, 1984; Hahn *et al'* 1984; Hahn *et al*, 1986). Regulation of mRNA production for the proteins in this operon involves DNA recognition at multiple operator sites by the AraC protein, with its negative or positive regulatory effect depending on the specific sites occupied. This, in turn, is influenced by the action of the sugar molecule, arabinose. The structural genes *araBAD* are positively regulated by AraC stimulating transcription more than 100-fold in the presence of arabinose. AraC autoregulates its own expression independently of the presence of the sugar by binding to two distantly-spaced operators, about 210bp apart. DNA looping occurs between protein molecules bound at these distant sites to allow protein:protein interactions. This isolates the promoter for *ara*C, resulting in reduced expression of the AraC protein product. These data have been demonstrated by footprinting studies. In addition it has been shown that the integration of DNA fragments prevent loop formation by disrupting the phase effect of the protein:protein interaction (Dunn *et al.* (1984). The formation of the loop has been shown to be dependent on supercoiling (Hahn, Dunn and Schleif, 1986; Lobell and Schleif, 1990).

Other prokaryotic regulatory systems have been shown to utilise DNA looping; including the the lactose operon (Flashner and Gralla, 1988; Sasse-Dwight and Gralla, 1990;Gralla, 1989) and the *deo* operon (Dandanell *et al*, 1987). The *deo* operon is especially interesting, as the Deo repressor is an octamer and in principle is capable of interacting with four operators simultaneously. Such an interaction is predicted to result in the formation of double-loops and has been observed by electron microscopy (Amouyal *et al*, 1989). DNA looping plays a very important part in the control of eukaryotic genes through the interaction of enhancers with promoters. Numerous examples have been demonstrated and include the steroid hormone receptors (Theveny *et al*, 1987), the yeast RAP-1 protein (Hofman *et al*, 1989),and a large number of other transcription factors (Muller, Gerstener and Schaffner, 1988).

Looping could occur between repressors bound at the CIR's upstream of both *c*P1 and *c*P2, repressing transcription from the *c*p1 promoter. The repressor proteins are

quite large by comparison with other prokaryotic repressor proteins. The 42kDa repressor probably forms a tetramer of approximately 170kDa. DNA-binding proteins of such a large size are frequently able to participate in DNA looping(for examples see above).

Within the 'early' region of ØC31, a region designated CIFC was found to contain two CIR's, either side of a terminator. Work presented in earlier chapters has demonstrated that the 42kDa repressor binds to the CHFC (upstream) and CIGC (downstream) halves of the CIFC region. It binds to the CHFC fragment poorly but relatively well to the CIGC fragment. Neither fragment, on their own, match the affinity of the complete CIFC region for the 42kDa repressor in a bandshift assay. Promoter activity has recently been found associated with the CIGC fragment (C.J. Ingham and M.C.M. Smith, personal communication). It is possible that the CIR in the CIGC half co-operates with the CIR in the other half via some DNA looping mechanism. This in turn may increase the affinity of the 42kDa repressor for binding the region and aid repression of transcription. Between the two CIR's is the terminator. The role of the CIR's in transcription termination, or the role of the terminator in transcriptional repression/activation is currently under investigation (C.J. Ingham and M.C.M. Smith, personal communication).

The footprinting data gained during the project, implies that the 42kDa repressor binds downstream of the CIR, using the right-hand half of the CIR as a target sequence. Some recent footprinting data of the 42kDa repressor with the *c*P2 promoter have confirmed the footprinting data using the *c*P1 promoter (S.E.Wilson and M.C.M Smith, personal communication). If these data are correct one could speculate on a role for the left-hand half of the CIR in binding the 54kDa repressor. Binding of the 54kDa repressor could be dependent on binding of the 42kDa repressor to the downstream half of the CIR, co-operatively interacting with the 54kDa repressor to allow binding, or *vice versa.*

Western blotting of *S.coelicolor* A3(2) lysogens, induced and uninduced, has demonstrated that very little of the 74kDa is produced by comparison with the 54 and 42kDa proteins (Smith and Owen, appendix). It has been suggested that the 74kDa repressor functions as an antirepressor, but the mechanism by which this is accomplished is not clear. There is some evidence for an activator protein in ØC31, as the induction of ØC31 lysogens leads to an increase in mRNA levels (Clayton and Bibb, 1991; Ingham *et al,* in press). It is likely that the 74kDa protein acts to stimulate expression from the activator promoter. This could be accomplished by the

74kDa repressor inhibiting the 54 and 42kDa repressors from binding, most probably through an increased affinity for its DNA recognition site which may possibly overlap the 54 and 42kDa protein binding sites. Alternatively the 74kDa repressor may interact with host-encoded proteins, so preventing transcriptional repression, or possibly stimulating transcription of 'late' promoters by presenting the host protein in a context that is recognised by RNA polymerase.

The CIR sites are so highly conserved and numerous within ØC31 that their presence must be essential. Footprinting data have demonstrated that at least the 42kDa repressor bind the CIR and it is possible that the 54 and 74kDa repressor proteins do so also. The CIR sequences seem to provide a focus for gene regulation within the phage. It is possible to imagine the 42kDa, and even the 54kDa repressor protein forming a complex to regulate transcription initiation, and possibly aid termination. The distinctive structure of the CIR is such that binding using 'indirect readout' could be speculated upon. It could then co-operatively bind the 54kDa repressor to form a tightly bound complex. Some of the CIR's are adjacent to putative terminator motifs, although no evidence is available for the involvement of the repressor proteins in termination. Alternatively the 54:42 repressor complex may recruit host-encoded proteins to regulate gene expression by occlusion of RNA polymerase or the titrating out of specific σ factors required for transcription of these promoters. Certainly the large size of the repressor proteins with a capacity to make such protein:protein interactions make such a scheme plausible.

Bibliography

Ackers, G. K., Johnson, A. and Shea M. (1982). Quantitative model for gene regulation by λ phage repressor. Proceedings of the National Academy of Sciences USA **79**:1129-1133.

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

Adhya, S. (1989). Multipartite genetic control elements: communication by DNA loop. Annual Review of Genetics. **23**: 227-250.

Aggarwal, A.K., Rodgers, D.W., Drottar, M., Ptashne, M., and Harrison, S.C. (1988). Recognition of a DNA operator by the repressor of phage 434: A view at high resolution. Science **242** : 899-907.

Alazard, R., Betermier, M. and Chandler, M. (1992). *Escherichia coli* integration host factor stabilises bacteriophage Mu repressor interactions with operator DNA *in vitro.* Molecular Microbiology **6**:1707-1714.

Alberti, S., Oehler, S., v.Wilcken-Bergmann, B., Kramer, H., and Muller-Hill, B. (1991). Dimer-to-tetramer assembly of lac repressor involves a leucine heptad repeat. New Biologist **3**:57-62.

Amouyal, M., and Buc, H. (1987). Journal of Molecular Biology **195**:795-808.

Amouyal, M., Mortensen, L., Buc, H., and Hammer, K. (1989). Single and double-loop formation when DeoR repressor binds to its natural operator sites. Cell **58**:545-551.

Anderson, J. E. , Ptashne, M. and Harrison, S. C. (1985). A phage repressor-operator complex at 7 A resolution. Nature **316**:596-601.

Anderson, J. E. , Ptashne, M. and Harrison, S. C. (1987). Structure of the repressor-operator complex of bacteriophage 434. Nature **326**: 846-852.

Anderson, W.F., Ohlensorf, D.H., Takeda, Y., and Matthews, B.W. (1981) Structure of the Cro repressor from bacteriophage Lambda and its interaction with DNA. Nature **290**:754.

Anderson, W.F., Takeda, Y., Echols, H., and Matthews, B.W. (1979) The structure of a repressor: Crystallographic data for the Cro regulatory protein of bacteriophage Lambda.Journal of Molecular Biology **130**:507.

Anzai, H., Murakami, T., Imai, S., Satoh, A., Nagaoka, K., and Thompson, C.J. (1987). Transcriptional regulation of Bialaphos biosynthesis in *Streptomyces hygroscopicus*. Journal of Bacteriology **169**:3482-3488.

Babcock, M.J., and Kendrick, K.E. (1990). Transcriptional and translational features of a sporulation gene of *Streptomyces griseus.* Gene **95**:57-63.

Babcock, M.J., McGrew, M., and Schottel, J.L. (1992). Iden.tification of a protein-binding sequence involved in expression of an *esterase* gene from *Streptomyces scabies*.Journal of Bacteriology **174**:4287-4293.

Bauw, G., van Damme, J., Puype, M., Vanderkerckhove, J., Gesser, B., Ratz, G.P., Lauridsen, J.B., and Celis, J.E. (1989). Protein-electro-blotting and -microsequencing strategies in generating protein data bases from two dimensional gels. Proceedings of the National Academy of Science USA **86**:7701-7705.

Baylis, H.A., and Bibb, M.J. (1988). Organization of the ribosomal RNA genes in *Streptomyces coelicolor A3(2)*. Molecular and General Genetics **211**: 191-196.

Bayliss, H.A., and Bibb, M.J.(1988). Transcriptional analysis of the 16S rRNA gene of the rrnD gene set of *Streptomyces coelicolor* A3(2). Molecular Microbiology **2**:569-579.

Berdy, J. (1980). C.R.C Handbook of Antibiotic Compounds. Published by CRC Press, Florida.

Berg, D.E. (1989). Transposon Tn5. In 'Mobile DNA'. Berg, D.E., and Howe, M.M. (eds). Washington D.C. American Society for Microbiology, pp185-210.

Berg, J.M.(1990). Zinc Finger Domains: Hypotheses and Current Knowledge. Annual Review of Biophysics and Biophysical Chemistry **19**: 405-421.

Berg, O.G., and von Hippel, P.H. (1988). Selection of DNA binding sites by regulatory proteins. II. The binding specificity of cyclic AMP receptor protein to recognition sites. Journal of Molecular BIology **200**:709-723.

Bertani, G. (1951). Studies in lysogenesis. I. The mode of phage liberation by lysogenic *E.coli*.Journal of Bacteriology **62**:293-300.

Bibb, M.J., Findlay, P.R., and Johnson, M.W. (1984). The relationship between base composition and codon usage in bacterial genes and its use for simple and reliable identification of protein-coding sequences. Gene **30**:157-166.

Birnboim, H.C., and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant DNA. Nucleic Acids Research **7**:1513-1523.

Bradley, D.E. (1967). Ultrastructure of bacteriophages and bacteriocins. Bacteriological Reviews **31**:230-314.

Brandon, C., and Tooze, J. (1991). Introduction to Protein Structure. Garland Publishing Co., New York.

Brennan, R.G., and Matthews, B.W. (1989a). Structural basis of DNA-protein recognition. Trends in Biochemistry **14**: 286-290.

Brennan, R.G., and Matthews, B.W. (1989b). The helix-turn-helix DNA binding motif. Journal of Biological Chemistry **264**: 1903-1906.

Brennan, R.G., Roderick, S.L., Takeda, Y., and Matthews, B.W. (1990). Protein-DNA conformational changes in the crystal structure of a Lambda Cro-operator complex. Proceedings of the National Academy of Sciences USA **87**:8165-8169.

Brown, K.L., Wood, S., and Buttner, M.J. (1992). Isolation and characterisation of the major vegatative RNA polymerase of *Streptomyces coelicolor* A3(2); renaturation of a sigma subunit using GroEL. Molecular Microbiology **6**:1133-1139.

Breg, J.N., Opheusden, J.H., Burgering, M.J., Boelens, R., and Kaptein, R. (1990). Structure of Arc repressor in solution: evidence for a family of β sheet DNA-binding proteins. Nature **346** 586- 7.

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

Burgess, R.R., and Jendrisak, J.J. (1975). A procedure for the rapid, large scale purification of *E.coli* DNA-dependent RNA polymerase involving polymin-P precipitation and DNA-cellulose chromatography. Biochemistry **14**:4634-4638.

Burton, Z., Burgess, R.R., Lin, J., Moore, D., Holder, S., and Gross, C.A. (1981). The nucleotide sequence of the cloned *rpo*D gene for the RNA-polymerase sigma subunit from *Escherichia coli* K12. Nucleic Acids Research **9**:2889-2903.

Busby, S. (1992). Kinked, curved or bent but certainly not going straight. Current Biology **2**: 53-55.

Busby, S.J.W. Positive regulation in gene expression. In:Booth, I.R., and Higgins, C.F.(eds), Regulation of Gene Expression-25 Years On. The Society for General Microbiology Symposium 39, Cambridge University Press, Cambridge, U.K., 1986, pp.51-57.

Buttner, M.J., and Brown, N.L. (1985). RNA polymerase-DNA interactions in *Streptomyces: in vitro* studies of a *S.lividans* plasmid promoter with *S.coelicolor* RNA polymerase Journal of Molecular Biology **185**:177-188.

Buttner, M.J., and Lewis, C.G.(1992). Construction and characterisation of *Streptomyces coelicolor A3(2)* mutants that are multiply deficient in the nonessential *hrd*-encoded RNA polymerase sigma factors. Journal of Bacteriology **174**:5165-5167.

Buttner, M.J., Chater, K.F., and Bibb, M.J. (1990). Cloning, disruption and transcriptional analysis of three RNA polymerase sigma factor genes of *Streptomyces coelicolor* A3(2).Journal of Bacteriology **172**:3367-3378.

Buttner, M.J., Smith, A.M., and Bibb, M.J. (1988). At least three different RNA polymerase holoenzymes direct transcription of the agarase gene (dagA) of *Streptomyces coelicolor* A3(2).Cell **52**:599-607.

+

Carey, J. (1988). Gel retardation at low pH resolves trp repressor-DNA complexes for quantitative study. Proceedings of the National Academy of Science USA **85**:975-979.

Carey, J. (1991). Gel Retardation. In 'Methods in Enzymology' vol.**208** Academic Press. pp103-117.

Carey, J., Lewis, D.E., Lavoie, T.A., and Yang, J. (1991). How does Trp repressor bind its operator? Journal of Biological Chemistry **266**:24509-24513.

Cashel, M., and Rudd, K.E. (1987). The stringent response. In "*Escherichia coli* and *Salmonella typhimurium:* Cellular and Molecular Biology". Neidhart, F.C. (Ed-in-chief). Washington D.C.: American Society for Microbiology, pp.1410-1438.

Chadwick, P., Pirrotta, V., Steinberg, R., Hopkins, N., and Ptashne, M. (1971). The Lambda and 434 phage repressors. Cold Spring Harbor Symposium on Quantitative Biology pp35:28

Chakerian, A.E., and Matthews, K.S. (1992). Effect of Lac repressor oligomerization on regulatory outcome. Molecular Microbiology **6**: 963-968.

Chater, K.F. (1984). Morphological and physiological differentiation in *Streptomyces.* In "Microbial Development" pp89-115 Eds R.Losick, L. Shapiro Cold Spring Harbor Laboratories.

+ Caballero, J.L., Malpartida, F., and Hopwood, D.A. (1991). Transcriptional organisation and regulation of an antibiotic export complex in the producing *Streptomyces* culture. Molecular and General Genetics 230: 401-12.

Chater, K.F. (1986). *Streptomyces* phages and their application to *Streptomyces* genetics. In 'The Bacteria' Vol. IX. Queener, S.E., and Day, L.E. (eds). Orlando: Academic Press, pp119-158.

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

Chater, K.F., Bruton, C.J., Plaskitt, K.A., Buttner, M.J., Mendez, C., and Helmann, J.D. (1989). The developmental fate of *S.coelicolor* hyphae depends upon a gene product homologous with the motility sigma factor of *B.subtilis.* Cell **59**:133-143.

Chinenova, T.A., and Lomovskaya, N.D. (1975). Temperature sensitive mutants of actinophage øC31 of *Streptomyces coelicolor* A3(2). Genetika (Moscow) **11**:132-141.

Chinenova, T.A., Mkrtumian, N.M., and Lomovskaya, N.D. (1982). Genetic study of a novel phage resistance character in *Streptomyces coelicolor* A3(2). Genetika (Moscow) **18**:1945-1952.

Christie, G., *et al.*, (1986). Regulation of bacteriophage P2 late-gene expression: the *ogr* gene. Proc. Natl. Acad. Sci. USA **83**:32238-3242.

Clayton, T.M. (1990). Ph.D Thesis. University of East Anglia.

Clayton, T.M., and Bibb, M.J. (1990). Induction of a ØC31 prophage inhibits rRNA transcription in *Streptomyces coelicolor* A3(2). Molecular Microbiology **4**:2179-2185.

Collado-Vides, J., Magasanik, B., and Gralla, J.D. (1991). Control Site Location and Transcriptional Regulation in *Escherichia coli.* Microbiological Reviews **55**: 371-394.

Craig N., and Nash, H. (1984). *E. coli* integration host factor binds to specific sites in DNA. Cell **39**:707-716.

Dai, D., and Ishiguro, E.E. (1991) Complementation of the *lyt*D1 mutation of *E.coli* by either the *cI* or *cro* gene of bacteriophage Lambda. Journal of Bacteriology **173**:893-895.

Dammann, T., and Wohlleben, W. (1992). A metalloprotease gene from *Streptomyces coelicolor*'Muller' and its transcriptional activator, a member of the LysR family. Molecular Microbiology **6**:2267-2278.

Dandanell, G., and Hammer, K. (1991). *deo*P1 promoter and operator mutants in *Escherichia coli:* isolation and characterization. Molecular Microbiology **5**:2371-2376.

Davis, B.J. (1964). Annals of the New York Academy of Science. **121**:4 04.

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

De Lorenzo, V., Wee, S., Herrero, M., and Neilands, J.B. (1987). Operator sequences of the Aerobactin operon of plasmid ColV-K30 binding the ferric uptake regulation (Fur) repressor. Journal of Bacteriology **169***:* 2624-2630.

Devereux, J., Haberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Research **12**:387-395.

Diaz, L.A., Hardisson, C., and Rodicio, M.R. (1989). Isolation and characterization of actinophage infecting *Streptomyces* species and their interaction with host restriction-modification systems. Journal of General Microbiology **135:** 1847-1856.

Diekman, S. (1989). Definitions and Nomenclature of nucleic acid structure parameters. EMBO Journal **8:** 1-4.

Dodd, I.B., and Egan, J.B. (1987). Systematic Method for the Detection of Potential Lambda Cro-like DNA-binding Regions in Proteins. Journal of Molecular Biology **194:** 557-564.

Dorman, C.J., et al. ;. (1988). DNA supercoiling and the anaerobic and growth phase regulation of *ton* B gene expression. Journal of Bacteriology **170:** 2816-2826.

+

Dunn, T.M., et al. (1984). An operator at 280 base pairs that is required for repression of *ara*BAD operon promoter: Addition of helical turns between the operator and promoter cylindrically hinders expression. Proceedings of the National Academy of Sciences USA **81:** 123-125.

Echols, H. (1986). Bacteriophage Lambda development: temporal switches and the choice of lysis or lysogeny. Trends in Genetics January, 26-30.

Errington, J., and Illing, N. (1992). Establishment of cell-specific transcription during sporulation in *Bacillus subtilis*. Molecular Microbiology **6**:689-695.

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

Flashner, Y., and Gralla, J.D. (1988). DNA Dynamic Flexibility and Protein Recognition: Differential Stimulation by Bacterial Histone-like Protein HU. Cell **54:** 713-721.

Folkmanis, A., Takeda, Y., Simuth, J., Gussin, G., and Echols, H. (1976). Purification and properties of a DNA-binding protein with characteristics expected for the Cro protein of bacteriophage Lambda, a repressor essential for lytic growth. Proceedings of the National Academy of Sciences USA **73**:2249-2253.

Fornwald, J.A., Schmidt, F.J., Adams, C.W., Rosenberg, M., and Brawner, M.E. (1987). Two promoters, one inducible and one constitutive, control transcription of the *Streptomyces lividans galactose* operon. Proceedings of the National Academy of Sciences USA **84**:2130-2134.

Freemont, P.S., Lane, A.N., and Sanderson, M.R. (1991). Structural Aspects of Protein-DNA Recognition. Biochemical Journal **278:** 1-23.

Fried, M., and Crothers, D.M. (1981). Equilibria and kinetics of the Lac repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleics Acid Research **9**: 6505-6525

Galas, D.J., and Schmitz, A. (1978). DNAase footprinting: a simple method for the detection of protein-DNA binding specificity. Nucleic Acids Research **5**:3157-3170.

Garner, M.M., and Revzin, A. (1981). A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the *Escherichia coli lactose* operon regulatory system. Nucleic Acids Research **9**: 3047-3060.

Gartenburg, M.R., and Crothers, D.M. (1988). DNA sequence determinants of CAP-induced bending and protein binding affinity. Nature **333:**824-829.

Gibson, T. (1984). PhD Thesis. University of Cambridge.

+ Doyle, D., McDowall, K.J., Butler, M.J., and Hunter, I.S. (1991). Characterisation of an oxytetracycline resistance gene *otr*A, of *Streptomyces rimosus*: nucleotide sequence, deduced function and promoter analysis. Molecular Microbiology **5**:2923-33.

Geistlich, M., Losick, R., Turner, J.R., and Nagaraja Rao, R. (1992). Characterization of a novel regulatory gene governing the expression of a polyketide synthase gene in *Streptomyces ambofaciens*. Molecular Microbiology **6**:2019-2029.

Giladi, H., Gottesman, M., and Oppenheim, A.B. (1990). Integration host factor stimulates the phage Lambda pL promoter. Journal of Molecular Biology **213**:109-121.

Giladi, H., Koby, S., Gottesman, M., Oppenheim, A. (1992). Supercoiling, Integration host factor and a dual promoter system, participate in the control of the bacteriophage Lambda pL promoter. Journal of Molecular Biology **224**:937-948.

Gill, D.R., and Salmond, G.P.C. (1990). The identification of the *Escherichia coli ftsY* gene product: an unusual protein. Molecular Microbiology **4**:575-583.

Goldenberg, D.P. (1989). Analysis of protein conformation by gel electrophoresis. In "Protein Structure-a practical approach", Creighton, T.E. (Ed.) IRL Press, Oxford, Britain. pp225-250.

Gralla, J.D. (1989). Bacterial gene regulation from distant DNA sites. Cell **57** : 193-195

Grossman, A.D., Erickson, J.W., and Gross, C.A. (1984). The *htpR* gene product of *Escherichia coli* is a sigma factor for heat-shock promoters. Cell **38**:383-390.

Guilfoile, P.G., and Hutchinson, C.R. (1992). The *Streptomyces glaucescens* TcmR protein represses transcription of the divergently oriented *tcmR* and *tcmA* genes by binding to an intergenic operator region. Journal of Bacteriology **174**:3659-3666.

Gussin, G.N., Johnson, A.D., Pabo, C.O., and Sauer, R.T. (1983). Repressor and cro protein: structure, function, and role in lysogenization. In "Lambda II" Hendrix et al. (Eds.). Cold Spring Harbor, New York.

Hager, D.A., and Burgess, R.R. (1980). Elution of proteins from sodium dodecyl siulfate polyacrylamide gels, removal of sodium dodecyl sulfate and renaturation of enzymatic activity: Results with sigma subunit of *E.coli* RNA polymerase, wheat germ topoisomerase, and other enzymes. Analytical Biochemistry **109**:76-86.

Hahn, S., Dunn, T., and Schleif, R. (1984). Upstream repression and Crp stimulation of the *Escherichia coli* L-arabinose operon. Journal of Molecular Biology **180**:61-72.

Haldenwang, W.G., and Losick, R. (1980). Novel RNA-polymerase sigma-factor from *Bacillus subtilis*. Proceedings of the National Academy of Sciences USA **77**:7000-7004.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. Journal of Molecular Biology **166**: 557-580.

Hara, O., and Beppu, T. (1982). Journal of Antibiotics **35**: 349-356.

Harlow, E., and Lane, D. (1988). Antibodies: A Laboratory Manual. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Hariharan, N., Kelley, D.E., and Perry, R.P. (1991). ∂, a transcription factor that binds to downstream elements in several polymerase II promoters, is a functionally versatile zinc finger protein. Proceedings of the National Academy of Science USA **88**: 9799-9803.

Harrison, S.C. (1991). A structural taxonomy of DNA-binding domains. Nature 353:715-719.

Harrison, S.C., and Aggrawal, A.K. (1990). DNA Recognition by Proteins with the Helix-Turn-Helix Motif. Annual Review of Biochemistry **59**: 933-969.

Heinzel, T., and Schuster, H. Abstracts from the "Molecular Genetics of Bacteria and Phages" Meeting at Cold Spring Harbor, New York. 1991.

Heinzel, T., Velleman, M., and Schuster, H. (1990). Journal of Biological Chemistry **265**:17928-17934.

Heisig, A., Riedel, H-D., Dobrinski, B., Lurz, R., and Schuster, H. (1989). Organization of the immunity region *imm*I of bacteriophage P1 and synthesis of the P1 antirepressor. Journal of Molecular Biology **209**:525-538.

Helmann, J.D. (1991). Altenative sigma factors and the regulation of flagellar gene expression. Molecular Microbiology **5**:2875-2882.

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

Helmann, J.D., Marquez, L.M., and Chamberlin, M.J. (1988). Cloning, sequencing and disruption of the *Bacillus subtilis sigma* gene. Journal of Bacteriology **170**:1568-1574.

Hendrickson, W., and Schleif, R. (1985). A dimer of AraC protein contacts three adjacent major groove regions of the *ara*I DNA site. Proceedings of the National Academy of Science. USA. **82**: 3129-3133.

Herskowitz, I., and Hagen, D. (1980). The lysis-lysogeny dexision of phage lambda: explicit programming and responsiveness. Annual Reviews of Genetics **14**:399-445.

Higgins, C.F. et al. (1988). A physiological role for DNA supercoiling in the osmotic regulation of gene expression in *S.typhimurium* and *E.coli* Cell **152**: 569-584

Hirschman, J., Wong, P.K., Sei, K., Keener, J., and Kustu, S. (1985). Products of nitrogen regulatory genes *ntr*A and *ntr*C of enteric bacteria activate *gln*A transcription in vitro- evidence that the *ntr*C gene product is a sigma factor. Proceedings of the National Academy of Sciences USA **82**:7525-7529.

Ho, Y-S., Wulff, D., and Rosenberg, M. (1986). Protein-nucleic acid interactions involved in transcription activation by the phage lambda regulatory protein cII. In "Regulation of Gene Expression-25 Years On" Eds: Booth, I.R., and Higgins, C.F. pp.79-103. Symposium 39 of the Society for General Microbiology. Cambridge University Press. Cambridge.

Hochschild, A. and Ptashne, M. (1986). Homologous interactions of Lambda repressor and Lambda Cro with the Lambda operator. Cell **44**: 925-933

Hochschild, A. and Ptashne, M. (1988). Interaction at a distance between Lambda repressors disrupts gene activation. Nature **336**: 353-357.

Hochschild, A., Irwin, N., and Ptashne, M. (1983). Repressor structure and the mechanism of positive control. Cell **32**:319

Holmes, D.S., and Quigley, M. (1981). A rapid boiling method for the preparation of bacterial plasmids. Analytical Biochemistry **114**:193-197.

Hindle, Z., and Smith, C.P. "Regulation of Glycerol Operon Gene Expression by the Glycerol Repressor". Poster presentation at 3rd Meeting on "Biology of the Actinomycetes"; August 16-20 1991.

Holt, T.G., Chang, C., Laurent-Winter, C., Murukami, T., Garrels, J.I., Davies, J.E., and Thompson, C.J. (1992). Global changes in gene expression related to antibiotic synthesis in *Streptomyces hygroscopicus*. Molecular Microbiology **6**:969-980.

Hoopes, B., and McClure, W. (1985). A cII-dependent promoter is located within the Q gene of bacteriophage Lambda. Proceedings of the National Academy of Science USA **82**:3134-3138.

Hopwood, D.A., Kieser, T. (1990). The *Streptomyces* Genome. In "The Bacterial Chromosome". Drlica, K., and Riley, M. (ed.). ASM, Washington D.C. pp147-162.

Horinouchi, S., and Beppu, T. (1988). Regulation of secondary metabolism in *Streptomyces*. In "Biology of Actinomyces 88", p71-76. Eds:Okami, Japan Scientific Societies Press.

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

Hoyt, M.A., Knight, D.M., Das, A., Miller, H.I., and Echols, H. (1982). Control of phage Lambda development by stability and synthesis of cII protein: Role of the viral *c*III and host *hfl*A, *him*A and *him*D genes. Cell **31**:565.

Hsiang, M., Cole, R., Takeda Y., Echols, H. (1977). Amino acid sequence of Cro regulatory protein of bacteriophage Lambda. Nature **270**:275-277

Hsiao, Y. (1991). Rapid Method for Double-stranded Sequencing. Nucleic Acids Research **19**:2787

Hunt, T.P., and Magasanik, B. (1985). Transcription of *gln*A by purified *Escherichia coli* components- RNA polymerase and the products of *gln*F, *gln*G and *gln*L. Proceedings of the National Academy of Sciences USA **82**:8453-8457.

Janssen, G.R., and Bibb, M.J. (1991). Tandem promoters, *tsr*p1 and *tsr*p2, direct transcription of the thiostrepton resistance gene (*tsr*) of *Streptomyces azureus*: Transcriptional initiation from *tsr*p2 occurs after deletion of the -35 region. Molecular and General Genetics **221**:339-346.

Jaurin, B., and Cohen, S.N. (1985). *Streptomyces* contain *Escherichia coli*-type A+T rich promoters having novel structural features. Gene **39**: 191-201.

Joachimiak, A., Kelley, R.L., Gunsalus, R.P., Yanofsky, C., and Sigler, P.B. (1983). Purification and characterization of trp aporepressor. Proceedings of the National Academy of Sciences. USA **80**:668-672.

Johns, E.W. In 'The HMG Chromosomal Proteins', New York., Academic Press. 1982.

Johnson, A.D., Pabo, C.O., and Sauer, R.T. (1980). Bacteriophage Lambda Repressor and cro Protein: Interactions with Operator DNA. In "Methods in Enzymology vol.**65** pp839-856. Academic Press

Jose-Gama, M., Toussaint, A., and Higgins, N.P. (1992). Stabilization of bacteriophage Mu repressor-operator complexes by the *Escherichia coli* integration host factor protein. Molecular Microbiology **6**:1715-1722.

Kassavetis, G.A., and Geiduschek, E.P. (1984). Defining a bacteriophage T4 late promoter- Bacteriophage T4 gene protein suffices for directing late promoter recognition. Proceedings of the National Academy of Sciences USA **81**:5101-5105.

Ingham, C.J., and Smith, M.C.M. (1992). Transcription map of the early region of the *Streptomyces* bacteriophage ØC31. Gene **122**:77-84.

Jacob, F., and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. Journal of Molecular Biology **3**:318.

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

Kieser, T., and Hopwood, D.A. (1991). Genetic manipulation of *Streptomyces*- Integrating vectors and gene replacement. In "Methods in Enzymology" Vol. **204.** Academic Press.,pp430-458.

King, J., Lenk, E.V., and Botstein, D. (1973). Mechanism of head assembly and DNA encapsulation in *Salmonella* phage P22. II. Morphogenetic pathway. Journal of Molecular Biology **80**:867.

Knight, K.L., and Sauer, R.T. (1989). DNA-binding specificity of the Arc and Mnt repressors is determined by a short region of N-terminal residues. Proceedings of the National Academy of Sciences USA **86**:797-801.

Knight, K.L., and Sauer, R.T. (1992). Biochemical and genetic analysis of operator contacts made by residues with the Beta sheet DNA binding motif of Mnt repressor. EMBO Journal **11**:215-223.

Koudelka, G. B., Harbury, P., Harrison, S. C. and Ptashne, M. (1988). DNA twisting and the affinity of bacteriophage 434 operator for bacteriaphage 434 repressor. Proceedings of the. National Academy of Sciences **85**: 4633-4637

Kourilsky, P. (1973). Lysogenization of bacteriophage lambda. I. Multiple infection and the lysogenic response. Molecular and General Genetics **122**: 183.

Kourilsky, P. (1974). Lysogenization of bacteriophage lambda. II. Identification of genes involved in the multiplicity dependent processes. Biochimie **56**:11

Kozmik, Z., Urbanek, P., and Paces, V. (1990). Albumin improves the formation and detection of some specific protein-DNA complexes in the mobility shift assay. Nucleic Acids Research **18**:2198

Kuhnke, G., Theres, C., Fritz, H-J., and Ehring, R. (1989). RNA polymerase and Gal repressor bind simultaneously and with DNA bending to the control region of the *E.coli galactose* operon. EMBO Journal **8**:1247-1255.

Kuhstoss, S., Richardson, M.A., and Rao, R.N. (1991). Plasmid cloning vectors that integrate site-specifically in *Streptomyces* spp. Gene **97**:143-146.

Kur, J., Hasan, N., and Szybalski, W. (1989). Physical and biological consequences of interactions between integration host factor (IHF) and coliphage Lambda late P'r promoter and its mutants. Gene **81**:1-15.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227**:680-685.

Lane, D., Prentki, P., and Chandler, M. (1992). Use of gel retardation to analyze protein-nucleic acid interactions. Microbiological Reviews **56** :509-528.

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

Levine, M. (1972). Replication and lysogeny with phage P22 in *Salmonella typhimuriam.* Current Topics in Microbiology and Immunology **58**:135.

Levine, M., Truesdell, S., Ramakrishnan, T., and Bronson, M.J. (1975). Dual control of lysogeny by bacteriophage P22: An antirepressor and its controlling elememts. Journal of Molecular Biology **91**:421.

Lim, A. and Sauer R. (1989). Alternative packing arrangements in the hydrophobic core of Lambda repressor. Nature **339**: 31-36.

Lobell, R.B., and Schleif, R.F. (1990). DNA looping and unlooping by the AraC protein. Science **250**:528-532.

Lomovskaya, N.D., Chater, K.F., and Mkrtumian, N.D. (1980). Genetics and Molecular Biology of *Streptomyces* Bacteriophages. Microbiological Reviews **44**:206-229.

Losick, R., and Pero, J. (1981). Cascades of sigma factors. Cell **25**:582-584.

Lumsden, J., and Coggins, J.R. (1977). The sub-unit structure of the Arom multienzyme complex of *Neurospora crassa*. Biochemical Journal **161**: 599-607.

Magasanik, B. (1989). Gene Regulation from Sites Near and Far. New Biologist **1**:247-251.

Majumdar, A., Rudikoff, S., and Adhya, S. (1987). Purification and properties of Gal repressor:pL-galR fusion in pFC31 plasmid vector. Journal of Biological Chemistry **262**:2326-2331.

Matsudaira, P. (1987). Sequence from Picomole Quantities of Proteins Electroblotted onto Polyvinylidene Difluoride Membranes. Journal of Biological Chemistry **262**: 10035-10038.

Matthews, K.S. (1992). DNA Looping. Microbiological Reviews. **56**: 125-136.

Maxam, A.M., and Gilbert, W. (1977). A new method for sequencing DNA. Proceedings of the National Academy of Science USA **74**:560-564.

McKay, D.B., and Steitz, T.A. (1981). Strucuture of catabolite gene activator protein at 2.9A resolution suggests binding to left-handed DNA. Nature **290**:744-749.

Menzel, R., and Gellert, M. (1983). Regulation of the genes for *E.coli* DNA gyrase: homeostatic control of DNA supercoiling. Cell **34**:105-113.

Meyer, B.J., Maurer, R., and Ptashne, M. (1980). Gene regulation at the right operator (OR) of bacteriophage Lambda. II. OR1, OR2, and OR3: Their roles in mediating the effects of repressor and cro. Journal of Molecular Biology **139**:163.

Miller, G., and Feiss, M. (1988). The bacteriophage Lambda cohesive end site-Isolation of spacing substitution mutations that result in dependence on *Escherichia coli* integration host factor. Molecular and General Genetics **212**:157-165.

Miller, H.I., Kikuchi, A., Nash, H.A., Weisberg, R.A., and Friedman, D.I. (1979). Site-specific recombination of bacteriophage Lambda: The role of host gene products. Cold Spring Harbor Symposium on Quantitative Biology **43**:1121.

Mirel. D.B., and Chamberlin, M.J. (1989). The *Bacillus subtilis* flagellin gene (*hag*) is transcribed by the sigma-28 form of RNA polymerase. Journal of Bacteriology **171**:3095-3101.

Miyake, K., et al. (1989). Detection and properties of A-factor binding protein from *Streptomyces griseus*. Journal of Bacteriology **171**: 4298-4302.

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

Mkrtumian, N.M., and Lomovskaya, N.D. (1972). A mutation affecting the ability of temperate actinophage ØC31 of *Streptomyces coelicolor* to lyse and lysogenize. Genetika **8**:135-141.

Morrissey, J.H. (1981). Silver stain for proteins in polyacrylamide gels: A modified procedure with enhanced uniform sensitivity. Analytical Biochemistry **117**:307-310.

Mortensen, L., Dandanell, G., and Hammer, K. (1989). Purification and characterisation of the deoR repressor of *Escherichia coli*. EMBO Journal **8**:325-331.

Muller, M.M., Gerster, T., and Schaffner, W. (1988). Enhancer sequences and the regulation of gene transcription. European Journal of Biochemistry **176**:485-495.

Nam, K., Blasi, U., Zagotta, M.T., and Young, R. (1990). Conservation of a dual-start motif in P22 lysis gene regulation. Journal of Bacteriology **172**:204-211.

Novikova, N.L., Kapitonova, O.N., and Lomovskaya, N.D. (1973). Prophage thermoinduction in germinating spores of *Streptomyces coelicolor* A3(2) (ØC31 ct1). Mikrobiologiya **42**:713-718.

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

Ohlendorf, D. H., Anderson, W. F., Fisher, R. G., Takeda, Y . and Matthews, B. W. (1982). The molecular basis of DNA protein recognition inferred from the structure of Cro repressor. Nature (London) **298:** 330:718-732

Ohtsubo, H., Nyman, K., Doroszkiewicz, W., and Ohtsubo, E. (1981). Multiple copies of iso-insertion sequences of IS1 in *Shigella dysenteriae* chromosome. Nature **292**: 640-643.

Ornstein, L. (1964). Annals of the New York Academy of Science **121**:321

Osborne. B., Stovall, S.R., and Baumstark, B.R. (1989). The *c*I genes of P1 and P7. Nucleic Acids Research **17**:7671-7680.

Pabo, C.O., and Lewis, M. (1982). The operator binding domain of Lambda repressor-structure and DNA recognition. Nature **298**:443.

Pabo, C.O., and Sauer, R.T. (1984). Protein-DNA Recognition. Annual Review of Biochemistry **53**: 293-321.

Pabo, C.O., Sauer, R.T., Sturtevant, J.M., Ptashne, M. (1979). The Lambda repressor contains two domains. Proceedings of the National Academy of Sciences USA **76**:1608.

Prentki, P., Chandler, M. and Galas, D. (1987). *Escherichia coli* integration host factor bends the DNA at the ends of IS1 and in an insertion hotspot with multiple IHF binding sites. EMBO Journal **6**:2479-2487.

Ptashne, M. (1986). A Genetic Switch. Cell Press and Blackwell Scientific Publications, Cambridge, MA.

Owen, C.E., and Chapman, P. (1991). Protein Sequencing- The Glasgow Way. Steptomycete Laboratory Publication.

Reichardt, L.F. (1975). Control of bacteriophage Lambda repressor synthesis: Regulation of the maintenance pathway of the cI and cro products. Journal of Molecular Biology **93**:289.

Reichardt, L.F., and Kaiser, A.D. (1971). Control of Lambda repressor synthesis. Proceedings of the National Academy of Sciences USA **68**:2185.

Roberts, T. M. (1977). Sequence of cro gene of bacteriophage Lambda. Nature **270**:274-275

Rodriguez, A., Caso, J.L., Hardisson, C., and Suarez, J.E. (1986). Characteristics of the developmental cycle of actinophage ØC31. Journal of General Microbiology **132**:1695-1701.

Rosenberg, M., Court, D., Shimatake, H., Brady, C., and Wulff, D.L. (1978). The relationship between function and DNA sequence in an intercistronic regulatory region in phage Lambda. Nature **272**:414-423.

Saha S., Lundqvist B., and Haggard-Lungquist, E. (1987). Autoregulation of bascteriophage P2 repressor. EMBO Journal **6**:809-814.

Saikumar P, Murali R and Reddy, E. (1990). Role of tryptophan repeats and flanking amino acids in Myb-DNA interactions. Proceedings of the National Academy of Science USA **87**:8452-8456.

Sanger, F., Coulson, A.R., Hong, G.F., Hill, G.F., and Peterson, G.B. (1982). The nucleotide sequence of bacteriophage Lambda DNA. Journal of Molecular Biology **162**:729-775.

Sasse-Dwight S., and Gralla J. (1990). Role of eukaryotic-type functional domains found in the prokaryotic enhancer receptor factor σ54. Cell **62**:945-954.

Sauer, R.T. (1978). DNA sequence of the bacteriophage Lambda *cI* gene. Nature **276**:301-302.

Sauer, R.T., Krovatin, W., DeAnda, J., Youderian, P., and Susskind, M.M. (1983). Primary strucuture of the *immI* immunity region of bacteriophage P22. Journal of Molecular Biology **168**:699.

Sauer, R.T., Yocum, R., Doolittle, R., Lewis, M., and Pabo, C.O. (1982). Homology among DNA-binding proteins suggests use of a conserved super-secondary structure. Nature **298**:447.

Schaefer, T.S., and Hays, J.B. (1990). The *bof* gene of bacteriophage P1: DNA sequence and evidence for roles in regulation of phage cI and *ref* genes. Journal of Bacteriology **172**:3269-3277.

Schagger, H., and von Jagow, G. (1987). Tricine-Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis for the Separation of Proteins in the Range from 1 to 100kDa. Analytical Biochemistry **166**:368-379.

Schevitz, R.W., Otwinowski, Z., Joachimiak, A., Lawson, C.L., and Sigler, P.B. (1985). The 3-dimensional structure of Trp repressor. Nature **317**:782-786.

Schinkel, A.H., Koerkamp, M.J.A.G., Teunissen, A.W.R.H., and Tabak, H.F. (1988). RNA-polymerase induces DNA-bending at yeast mitochondrial promoters. Nucleic Acids Research **16**:9147-9163.

Schleif, R. (1988). DNA binding by proteins. Science **241**: 1182-1187

Schleif, R. (1988). DNA Looping. Science **240**: 127-128.

Schlein, C.H. (1990). Solubility as a function of protein structure and solvent components. Bio/technology **8**:308-317.

Sinclair, R., and Bibb, M.J. (1988). The repressor gene (*c*) of the *Streptomyces* temperate phage ØC31: nucleotide sequence, analysis and functional cloning. Molecular and General Genetics **213:** 269-277.

Sinclair, R.B., and Bibb, M.J. (1989). Transcriptional analysis of the repressor gene of the temperate *Streptomyces* phage ØC31. Gene **85:** 275-282.

Sladkova, I.A., Vasil'chenko, L.G., Lomovskaya, N.D., and Mkrtumian, N.M. (1970). Physical mapping of *Streptomyces* A3(2) actinophages. I. Location of *c* region of actinophage øC31. Mol. Biol. (Moscow) **14**:916-921.

Smith, C.P., and Chatre, K.F. (1988). Strucutre and Regulation of Controlling Sequences for the *Streptomyces coelicolor* glycerol operon. Journal of Molecular Biology **204**:569-580.

Smith, M.C.M., Ingham, C.J., Owen, C.E., and Wood, N.T. (1992). Gene expression in the *Streptomyces* phage ØC31. Gene **115** : 43-48.

Sogaard_Anderson, L., Pedersen, H., and Valentin-Hansen, P. (1991). A novel function of the cAMP-CRP complex in *Escherichia coli:* cAMP-CRP functions as an adaptor for the CytR repressor in the *deo* operon. Molecular Microbiology **5**:969-975.

Staacke, D., Walter, B., Kisters-Woike, B., Wilcken-Bergmann, B., and Muller-Hill, B. (1990). How Trp repressor binds its operator. EMBO Journal **9**:1963-1967.

Stark, W.M., Grindley, N.D.F., Hatfull, G.F., and Boocock, M.R. (1991). Resolvase catalysed reactions between Res sites differing in the central dinucleotides if subsite I. EMBO Journal **10**:3541-3548.

Stragier, P., and Losick, R. (1990). Cascades of sigma factors revisited. Molecular Microbiology **4**:1801-1806.

Stragier, P., and Patte, J.C. (1983). Regulation of diaminopimelate decarboxylase synthesis in *E.coli*. Journal of Molecular Biology **168**:333-350.

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

Strohl, W.R. (1992). Compilation and analysis of DNA sequences associated with apparent streptomycete promoters. Nucleic Acids Research **20**:961-974.

Struhl, K. (1989). Helix-turn-helix, zinc finger, and leucine zipper motifs for eukaryotic transcriptional regulatory proteins. Trends in Biochemistry **14**:137-140.

Suarez, J.E., Caso, J.L., Rodriguez, A., and Hardisson, C. (1984). Structural characteristics of the *Streptomyces* bacteriophage ØC31. FEMS Microbiology Letters **22**:113-117.

Susskind, M., and Youderian, P. (1982). Transcription in vitro of the bacteriophage P22 antirepressor gene. Journal of Molecular Biology **154**:427.

Susskind, M.M., and Youderian, P. (1983). Bacteriophage P22 antirepressor and its control. In Lambda II., Hendrix et al, (Eds)., Cold Spring Harbor, New York. pp347-363.

Takeda, Y., Sarai, A., and Rivera, V.M. (1989). Analysis of the sequence-specific interactions between Cro repressor and operator DNA by systematic base substitution experiments. Proceedings of the National Academy of Sciences USA **86**:439-443.

Tanaka, K., Shiina, T., and Takahashi, H. (1988). Multiple principle sigma factor homologs in eubacteria: idenitfication of the "rpoD" box. Science **242**:1040-1042.

Tanaka, K., Shiina, T., and Takahashi, H. (1991). Nucleotide sequence of genes *hrd*A, *hrd*C and *hrd*D from *Streptomyces coelicolor* A3(2) having similarity to *rpo*D genes. Molecular and General Genetics **229**:334-340.

Thompson, C.J. (1992). "Regulation of Secondary Metabolism" - 27th Meeting of the Society for General Microbiology, Bristol, January 1992.

Thompson, J.F., and Landy, A. (1988). Empirical estimations of protein induced DNA bending angles- Applications to Lambda site-specific recombination complexes. Nucleic Acids Research **16**:9687-9705.

Towbin, H., Staehelin, T., and Gordon, J. (1979). Proceedings of the National Academy of Science USA **76**:4350.

Travers, A.A. (1991). To Bend or.....? Current Biology **1**: 171-172.

Valenzuala, D., and Ptashne, M. (1989). P22 repressor mutants deficient in co-operative binding and DNA loop formation. EMBO Journal **8**:4345-4350.

Vasil'chenko, L.G., Mkrtumian, N.M., and Lomovskaya, N.D. (1981). Genetic mapping and characterization of deletion mutants of actinophage ØC31 of *Streptomyces coelicolor* A3(2) incapable of lysogenization. Genetika **17**:1967-1974.

Velleman, M., and Schuster, H. Abstracts from the "Molecular Genetics of Bacteria and Phages" Meeting, Cold Spring Harbor, New York. 1991

Velleman, M., Dreisekelmann, B., and Schuster, H. (1987). Multiple repressor binding sites in the genome of bacteriophage P1. Proceedings of the National Academy of Sciences USA **84**:5570-5574.

Velleman, M., Heirich, M., Gunther, A., and Schuster, H. (1990). Journal of Biological Chemistry **265**:18511-18517.

von Hippel, P.H., Bear, D.G., Morgan, W.D., and McSwiggen, J.A. (1984). Protein-nucleic acid interactions in transcription: A molecular analysis. Annual Review of Biochemistry **62**: 389-445.

Westpheling, J., and Brawner, M. (1989). Two transcribing activities are involved in expression of the *Streptomyces* galactose operon. Journal of Bacteriology **171**:1355-1361.

Westpheling, J., Ranes, M., and Losick, R. (1985). RNA polymerase heterogeneity in *Streptomyces coelicolor*. Nature **313**:22-27.

White, P.J., Young, J., Hunter, I.S., Nimmo, H., and Coggins, J.R. (1990). Purification of 3-dehydroquinase from *Streptomyces coelicolor* A3(2). Biochemical Journal **265**:735-738.

Woese, C.R. (1987). Bacterial evolution. Microbiological Reviews **51**:221-271.

Wu, A.M., Ghosh, S., Echols, H., and Spiegelman, W.G. (1972). Repression of the cI protein of phage Lambda: in vitro inhibition of RNA synthesis. Journal of Molecular Biology **67**:407.

Wulff, D.L., and Rosenberg, M. (1983). Establishment of repressor synthesis. In "The Bacteriophage Lambda" vol. 2. Eds: Hendrix, J., Roberts, J., Stahl, F., and Weissman, R. pp. 53-73. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

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

Yarmolinsky, M.B., and Sternberg, N. "Bacteriophage P1" In The Bacteriophages Volume 1 (Ed. R.Calendar) Plenum Press, New York, 1988.

Youderian, P., Chadwick, S.J., and Susskind, M.M. (1982). Autogenous regulation by the bacteriophage P22 *arc* gene product. Journal of Molecular Biology **154**:449.