STUDIES ON THE CONTROL OF LATE GENE TRANSCRIPTION IN COLIPHAGE 186

A Thesis submitted for

the degree of

Doctor of Philosophy

at the

University of Adelaide

by

JUSTIN ANDREW DIBBENS (B.Sc. Hons)

Department of Biochemistry

June, 1990

FOR MY MOTHER

TABLE OF CONTENTS

TABLE OF CONTENTS (cont.)

TABLE OF CONTENTS (cont.)

TABLE OF CONTENTS (cont.)

TABLE OF CONTENTS (cont.)

TABLE OF CONTENTS (cont.)

APPENDICES:

SUMMARY

STUDIES ON THE CONTROL OF LATE GENE TRANSCRIPTION IN COLIPHAGE 186.

This thesis describes studies carried out in an investigation of the control of late gene transcription in temperate coliphage 186. The specific aims of this study were to identify the phage functions involved in activating late gene transcription and to investigate the mechanism of control of late gene transcription.

The location and characterization of a 186 late promoter was chosen as the first step in investigating the control of late transcription in 186. Polarity and marker rescue studies had indicated that the *Bam*HI-*Eco*RI (10.0% to 13.3%) region was likely to encode the promoter for the transcription unit encoding late genes *VUTSRQ* (Hocking and Egan, 1982c; Finnegan and Egan, 1979). In this work the DNA sequence of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region was determined. Computer assisted analysis of the DNA sequence led to the prediction that two divergent genes were located on this restriction fragment. The rightward gene was identified as the *V* gene and the leftward gene was designated *CP12*.

The opposite directions of *V* and *CP12* indicated the likelihood that divergent transcription initiated between these genes. Extensive homology was found in this region with the corresponding late promoters of the related phage P2, and led to the prediction that gene *V* was transcribed from a rightward late promoter p_V, while *CP12* was predicted to be transcribed from a leftward late promoter p_{12}. Primer extension analysis of *in vivo* RNA confirmed the *in vivo* activities of these promoters and demonstrated that the activity of both promoters was dependent upon the *B* gene product, consistent with the requirement of the *B* gene product for late gene transcription as demonstrated by Finnegan and Egan (1981).

An assay system for transcription from a 186 late promoter was established by placing the *galK* gene under control of the late promoter p_V on a multicopy plasmid. Consistent with the need for phage functions for activation, p_V showed no activity in the absence of phage functions *in vivo*. Late gene transcription has been shown to be dependent on the functions of gene *B* and replication gene *A* (Finnegan and Egan, 1981). In this work, the *B* gene product was shown to be the only phage function required directly

for activation of p_V transcription and to be required at high concentration for activation. Phage replication had previously been concluded to provide a template topology necessary for B protein to activate late transcription (Kalionis *et al.*, 1986b), however late transcription was shown to occur in the absence of phage replication and ruled out the need of a specific template toplogy. Studies indicated that the role of replication in activation of late gene transcription was to increase *B* gene dosage and to potentially derepress *B* gene transcription. These studies also indicated a protocol for the isolation of host mutants unable to activate late gene transcription.

Kalionis *et al.* (1986b) concluded that the *B* gene product autogenously controlled its own transcription in the lysogenic state, either directly or indirectly. Control of *B* gene transcription was investigated since there was no evidence of activation of late transcription in a lysogen, even though B protein was present and active in repressing its own transcription in this state. In this work the control of *B* gene transcription was shown not to be direct, but to be indirect and require the involvement of the *c*I gene product for repression of transcription from the *B* promoter. The inability of the prophage *B* gene to activate late transcription was concluded to be a direct consequence of repression of *B* gene transcription and not due to the *c*I gene product interfering with *B*-dependent activation directly.

The *c*I gene product is normally involved in repression of lytic transcription by binding to a primary operator site overlapping the early lytic promoter (Lamont *et al.*, 1988). However, the presumptive operator site for the repressor of *B* gene transcription (Kalionis *et al.*, 1986b) shows no homology to the primary CI operator. Since the autogenous control of *B* gene transcription was indirect in the lysogenic state, this indicated that either the expression of the *c*I gene was under the control of B protein or that B protein modified the activity of the CI protein. Expression of the *c*I gene does not appear to be under *B* gene control and indicates that the control of *B* gene transcription is likely to be by a mechanism in which the B protein alters the DNA-binding specificty of the CI protein so that it represses transcription from the *B* promoter.

The work presented in this thesis provides the basis for further studies in elucidating the mechanism of late promoter activation and control of *B* gene transcription in coliphage 186.

STATEMENT

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university. To the best of my knowledge and belief, this thesis contains material not previously published or written by another person except where due reference is made in the text.

I consent to this thesis being made available for photocopying and loan.

Justin A. Dibbens

ACKNOWLEDGEMENTS

The author wishes to thank the following persons who have been associated at some time with this thesis.

Pofessor W.H. Elliot for his permission to work in the Biochemistry Department.

My supervisor Dr. Barry Egan for his advice and enthusiasm throughout the course of my Ph. D., his constructive criticism of this thesis, and for sharing his philosophy on life.

Members of Lab 19, both past and present for their friendship, advice and providing unusual entertainment. In particular I would like to thank Ian Dodd for his enthusiasm, good advice and help with the computing.

Anna Heath for typing the bulk of this thesis and for cheerfully accepting the continued revisions.

Brandt Clifford for his preparation of the photographs.

Jon Anargyros and Janis Soltys for their excellent technical assistance.

Rosslyn Foster, Mary Milnthorp and Judy Whittle for providing a continuous supply of clean glassware.

To all my friends who have been so supportive. In particular, Felicity Johnson and Simon Heath for their friendship and support, and for helping me to cope with the fickle nature of Science.

Finally, my family for their support, unfailing encouragement and understanding, without which this thesis would not have been possible.

ABBREVIATIONS

min	-	minute
A_{600}	-	absorbance at 600 nm
A_{650}	-	absorbance at 650 nm
DNA	-	deoxyribonucleic acid
RNA	-	ribonucleic acid
kd	-	kilodalton
b	-	base
bp	-	base-pair
kb	-	kilobase (1,000 base-pairs)
kPa	-	kilopascal
Ci	-	Curie
uCi	-	microCurie
mg	-	milligram
ml	-	millilitre
mA	-	milliamp
M	-	Molar
mM	-	millimolar
uM	-	micromolar
mmol	-	millimole
ng	-	nanogram
ug	-	microgram
ul	-	microlitre
nm	-	nanometre
u	-	micron

ABBREVIATIONS (cont.)

V	-	volts
Kcal	-	Kilocalorie
EtBr	-	ethidium bromide
Tris	-	Tris (hydroxymethyl) aminomethane
EDTA	-	ethylenediamine tetra acetate
BCIG	-	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
IPTG	-	Isopropyl-B-D-thiogalactopyranoside
gm	-	gram
rpm	-	revolutions per minute
g	-	gravitational force
wt	-	wild-type
UV	-	ultraviolet
cpm	-	counts per minute
U	-	units
pfu	-	plaque-forming units
cfu	-	colony-forming units
rbs	-	ribosome-binding site
aa	-	amino acids
RF	-	replicative-form
EM	-	electron microscope
moa	-	mutiplicity of addition

Other abbreviations are described in Chapter 2.

CHAPTER 1

INTRODUCTION

CHAPTER 1. INTRODUCTION.

1.1 CONTROL OF PHAGE GENE EXPRESSION.

Bacteriophages follow a strictly coordinated developmental plan during lytic development. This involves the concerted expression of a number of different functions at various stages in bacteriophage growth. Additionally, temperate phages have the capacity to either follow lytic or lysogenic development and must ensure that once the decision to adopt one of the alternative states has been made, functions detrimental to the maintainence of that state are not expressed. Temperate phages present ideal systems for the study of how gene expression is regulated in prokaryotes and this is reflected in that studies of temperate phage λ have played a pivotal role in the elucidation of the mechanisms of how prokaryotic gene expression is controlled.

Further, many temperate phages are unrelated in nucleotide and genetic composition but demonstrate similar developmental programmes. It is therefore of much interest to determine the mechanisms by which unrelated phages control expression of their genes in order to pursue similar developmental programmes.

1.2 COLIPHAGE 186.

Coliphage 186 is a temperate phage of the P2-related group of phage (Bertani and Six, 1988), which includes phages P2, 186, P4, 299, 18 and W. This group of phages is distinct from the lamboid group of phages (λ, 434, 424, P22) in morphology, nucleotide composition and genetic organization (Bertani and Bertani, 1971; Skalka and Hanson, 1972; Hocking and Egan, 1982a; Szybalski and Szybalski, 1979).

Phage 186 has several properties which distinguish it from the prototype phage of the family, P2 (Bertani and Six, 1988). These include: the ability of 186 to be efficiently induced from the prophage state under conditions which do not induce a P2 prophage, such as treatment with UV irradiation, nalidixic acid and mitomycin C (Woods and Egan, 1974; Bertani, 1968); the lack of interference by 186 lysogens of phage λ growth as seen with P2 lysogens (Woods and Egan, 1974; Lindahl *et al.*, 1970); and the requirement of the host genes *dnaA* and *dnaC* for 186 replication (Hooper and Egan, 1981).

Bacteriophage 186 has a double stranded, non-permuted DNA genome of approximately 30 kb in length and a molecular weight of 19.6 x 10^6 daltons (Wang, 1967; Wang and Schwartz, 1967; Chattoraj *et al.*, 1973; Younghusband *et al.*, 1975). The DNA possesses complementary cohesive ends (*cos*) 19 base pairs in length (Baldwin *et al.*, 1966; Wang *et al.*, 1973; Murray and Murray, 1973).

1.3 GENETIC ORGANIZATION OF THE 186 AND P2 GENOMES.

1.3.1 Genetic content and organization of phage 186.

A linear genetic map of 186 was constructed by a novel method using marker rescue frequencies (Hocking and Egan, 1982a) and the physical mapping of insertion and deletion mutants, localized by electron microscopy (Younghusband *et al.*, 1975). A physical restriction map has been constructed (Saint and Egan, 1979) and the genetic content of cloned restriction fragments determined by marker rescue (Finnegan and Egan, 1979), allowing the physical and genetic maps to be aligned with respect to each other. In addition, the DNA strands of 186 have been oriented with respect to the physical and genetic maps (Kalionis and Egan, 1981).

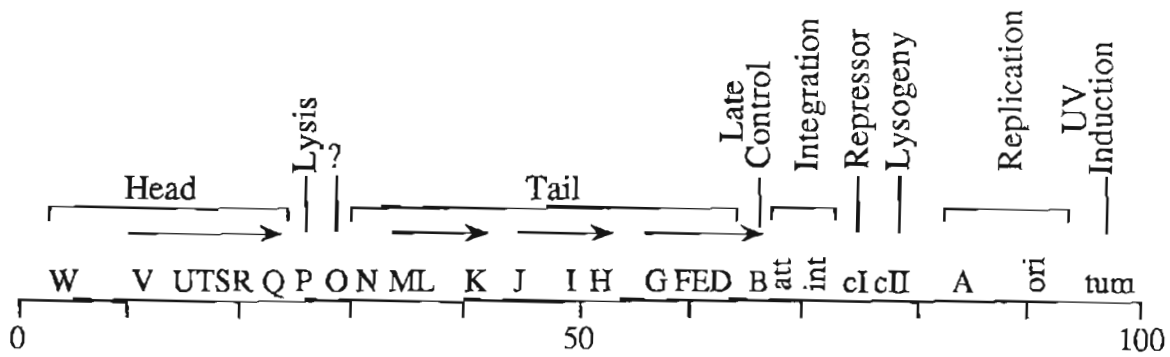
The physical and genetic map of 186 is shown in Figure 1.1(a) and is displayed as a linear structure terminating with the cohesive ends. Twenty two genes essential for lytic development have been identified (Hocking and Egan, 1982c,d) and all but two are located in the region 0% to 65.5% of the 186 genome (Finnegan and Egan, 1979). The genes located in the left two-thirds (0% to 65.5%) of the phage chromosome are involved in phage morphogenesis and host cell lysis (Hocking and Egan, 1982c) and are transcribed late in lytic infection (Finnegan and Egan, 1981). Genes involved in phage tail formation are represented by genes *N* to *D* and genes involved in phage head formation by genes *W* to *Q*. Gene *P* is involved in host cell lysis and the function of gene *O* is not known. The late genes of 186 have been divided into four groups on the basis of the polar effects of amber mutations on the expression of distal genes (Figure 1.1(a); Hocking and Egan, 1982c) and this indicates that the late region of 186 is organized into at least four transcription units.

Figure 1.1 Physical and genetic maps of phages 186 and P2.

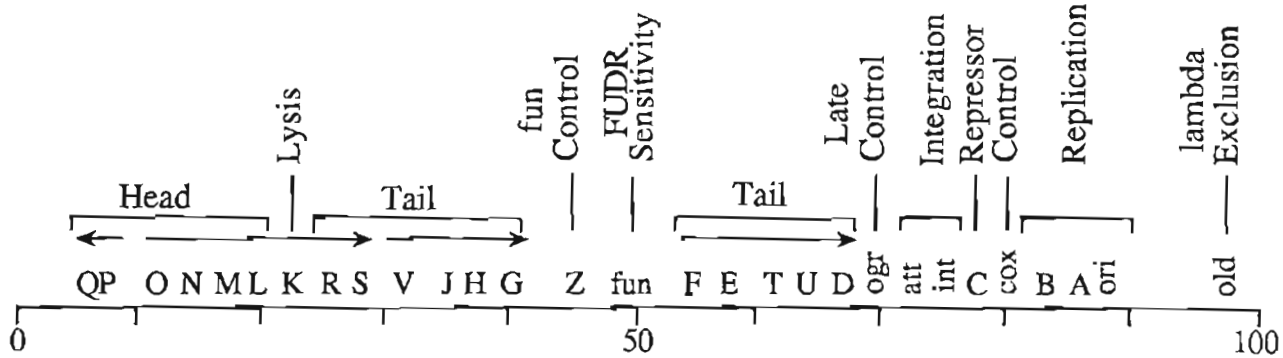
(a) Physical and genetic map of phage 186 as determined by Hocking and Egan (1982a). Functions of the genes are shown above the Figure, and were determined by Hocking and Egan (1982c,d). The physical mapping was described by Finnegan and Egan (1979). The *c*I gene codes for the maintainence repressor (Baldwin *et al.*, 1966; Woods and Egan, 1974). The *int* gene is required for integration and excision (Bradley *et al.*, 1975; J.B. Egan, personal communication). The physical positions of the *c*I and *int* genes were determined by Younghusband *et al.* (1975). The *c*II gene is required for the establishment of lysogeny (Huddleston, 1970) and mapped at 75.9% (Carter, 1985). The phage attachment site (*att*) has been located at 67.9% (Pell, 1987). The origin of replication (*ori*) was defined by EM studies at 92.9% (Chattoraj and Inman, 1973). The *tum* gene is involved in UV induction and is located at 95% (Lamont *et al.*, 1989). Transcription units are from polarity studies (Hocking and Egan, 1982c), and are represented by arrows.

(b) Physical and genetic map of phage P2. Functions of the genes are shown above the Figure. Gene order was determined by Lindahl (1969) and the positions of various genes located by heteroduplex mapping of deletions and insertions (Chattoraj and Inman, 1972, 1974; Bertani, 1975; Chattoraj *et al.*, 1975; Hyde and Bertani, 1975; Chattoraj and Bertani, 1980), marker rescue (Chattoraj, 1977; Westoo and Ljungquist, 1979, 1981; Ljungquist and Bertani, 1983). The *ogr* gene is involved in late gene control (Sunshine and Sauer, 1975; Birkeland and Lindqvist, 1986; Christie *et al.*, 1986). The *int* gene is involved in phage integration (Ljungquist and Bertani (1983). The *C* gene codes for the repressor of lytic transcription (Saha *et al.*, 1987a). The *cox* gene inhibits the synthesis of C protein (Saha *et al.*, 1987b; Saha *et al.*, 1989). Genes *B* and *A* are involved in phage replication, and the origin of replication (ori) is located at 89% (Schnos and Inman, 1971). The *old* gene is involved in λ exclusion (Lindahl *et al.*, 1970; Gibbs *et al.*, 1983) and its position located by deletion (Bertani, 1975).

(a)


Head
Lysis
?
Tail
Late Control
Integration
Repressor
Lysogeny
Replication
UV Induction
W V UTSR Q P O N ML K J I H G FED B att int cI cII A ori tum
0
50
100

(b)


Head
Lysis
Tail
fun Control
FUDR Sensitivity
Tail
Late Control
Integration
Repressor Control
Replication
lambda Exclusion
QP O N ML K R S V J H G Z fun F E T U D ogr att int C cox B A ori old
0
50
100

The region from 65.5% to 100% encodes the essential genes *B* and *A* (Hocking and Egan, 1982d), as well as a number of other non-essential genes, which include the genes *c*I, *int*, *c*II and *tum* (Figure 1.1(a)). The *B* gene encodes a function required for transcription of the genes involved in phage morphogenesis and host lysis (Finnegan and Egan, 1981) and gene *A* is required for phage replication (Hocking and Egan, 1982d). The *c*I gene encodes a function required for the maintainence of lysogeny (Baldwin *et al.*, 1966). The *int* gene is required for the integration of 186 DNA into the bacterial chromosome during the establishment of lysogeny and also for the excision of the 186 prophage from the bacterial chromosome during lysogenic induction (Bradley *et al.*, 1975; J.B. Egan, personal communication). The position of the *int* and *c*I genes was indicated by the physical mapping of two insertion mutants *ins*3 at 70.3% and *ins*1 at 73.5% (Younghusband *et al.*, 1975), which show the *int*$^-$ and *c*I$^-$ phenotypes respectively. The position of the phage *att* site was located by recombination mapping to between the *B* and *c*I genes (Hocking and Egan, 1982a) and by hybridization studies to the right of the *Xho*I site at 67.6% (Pell, 1987). Mutations in the *c*II gene define a second clear plaque complementation group (Huddleston, 1970; Lamont and Egan, manuscript in preparation) and the *c*II gene may encode a function required for the establishment of lysogeny (Huddleston, 1970; Carter, 1985), The *tum* gene encodes a function involved in UV induction and was identified by mutants of 186 able to form turbid plaques on mitomycin C (Lamont *et al.*, 1989). The 186 origin of replication (*ori*) has been mapped at 92.9% ± 1.8% (Chattoraj and Inman, 1973).

1.3.2 Genetic content and organization of phage P2.

Like 186, P2 DNA is double stranded, linear and non-permuted and terminates in single stranded, complementary ends (Bertani and Bertani, 1971; Inman *et al.*, 1971; Younghusband and Inman, 1974). P2 DNA has a molecular weight of 22 x 10^6 daltons and is approximately 33 kb in size (Bertani and Bertani, 1971). A total of 27 genes have been defined by mutation (Figure 1.1(b); Lindahl, 1969; Sunshine *et al.*, 1971). They can be grouped into three classes: those essential for lytic multiplication (genes designated by uppercase letters, excepting gene *C*), those involved in lysogeny (*int*, *C* and *cox*) and those of unknown or non-essential function (*old*, *Z* and *fun*). The order of the genes has been

established by three point crosses (Lindahl, 1969). The genes have been positioned on the chromosome by heteroduplex mapping of deletions and insertions (Chattoraj and Inman, 1972, 1974; Bertani, 1975; Chattoraj *et al.*, 1975; Hyde and Bertani, 1975; Chattoraj and Bertani, 1980), by marker rescue from cloned restriction fragments (Chattoraj, 1977; Westoo and Ljungquist, 1979, 1980; Ljungquist and Bertani, 1983) and by sequencing (Christie and Calendar, 1983, 1985; Ljungquist *et al.*, 1984; Haggard-Ljungquist *et al.*, 1987).

The late region of phage P2 encodes the genes involved in phage morphogenesis and host cell lysis and appears to be organized into four transcriptional units (Figure 1.1(b); Lindahl, 1971; Sunshine *et al.*, 1971). The head genes *P* and *Q* of P2 are transcribed in the opposite direction to the rest of the late genes (Lindqvist and Bovre, 1972; Geisselsoder *et al.*, 1973; Christie and Calendar, 1983), and therefore excludes the possibility that the late genes are expressed by the extension of a transcript from a single rightward promoter. Transcription of the late genes of P2 is dependent upon the functions of the *ogr* gene (Sunshine and Sauer, 1975) and the function of genes *A* and *B* (Geisselsoder *et al.*, 1973; Lengyel and Calendar, 1974; Funnell and Inman, 1982), which are both required for P2 replication in a wild-type host (Lindahl, 1974).

1.3.3 Homology between 186 and P2.

Phages 186 and P2 have a similar arrangement of genes with related functions (Figure 1.1; Hocking and Egan, 1982a,c; Lindahl, 1969; Sunshine *et al.*, 1971). The two phages show considerable homology in the region of their genomes encoding the structural (late) genes as demonstrated by electron heteroduplex mapping (Younghusband and Inman, 1974), which indicates about 50% sequence identity between 186 and P2 in the late regions (0% to 65.5%), whereas the early regions, encoding the regulatory functions, show little homology. The similarity in the late genes is reflected in the fact that viable P2-186 hybrid phage containing P2 late and 186 early genes have been isolated (Younghusband *et al.*, 1975; Bradley *et al.*, 1975; Hocking and Egan, 1982b). Hybrids 5 through 12 possess only the *B* gene of 186 for the activation of P2 late genes (Hocking and Egan, 1982b), demonstrating that the product of the late control gene *B* of 186 is functionally interchangeable with the product of the P2 late control gene *ogr*.

1.4 DNA SEQUENCE AND TRANSCRIPTIONAL ANALYSIS OF 186.

1.4.1 DNA sequence determination.

The DNA sequence spanning 65.5% to 2.3% of the 186 genome has been determined (Figure 1.2(a); Kalionis *et al.*, 1986a; Richardson *et al.*, 1989; Sivaprasad *et al.*, 1990; I. Lamont, personal communication). Evaluation of the DNA sequence by computer assisted analysis has enabled the identification of potential protein-coding regions (genes) on the basis of codon usage and the presence of ribosome binding sites, which are designated *CP* (for Computer Protein) followed by the chromosomal coordinate approximating the predicted initiation codon of the gene. The results of the computer assisted analysis of the DNA sequence spanning 65.5% to 2.3% are shown in Figure 1.2. Analysis of the DNA sequence from wild-type and mutants enabled the identification of the *D*, *B*, *int*, and *c*I genes (Figure 1.2(b); Kalionis *et al.*, 1986a,b), the *c*II gene (Figure 1.2(b); Carter, 1985) and the *tum* gene (Figure 1.2(d); Lamont *et al.*, 1989). Analysis of the DNA sequence of amber mutants in the replication gene *A*, revealed that the amber mutations mapped in two genes, *CP84* and *CP87* (Figure 1.2(c)), with gene *CP87* being the only phage gene necessary to form a minichromosome (Sivaprasad *et al.*, 1990).

1.4.2 Transcriptional analysis.

Computer assisted analysis has also allowed the prediction of promoters recognized by *E. coli* RNA polymerase, by the use of a weight matrix (Kalionis *et al.*, 1986a) based on the number of occurrences of each base at each position of known promoters (Hawley and McClure, 1983), and structures which fulfill the criteria for *rho*-independent terminators (Rosenberg and Court, 1979). This allowed the prediction of transcription units encoding identified genes and the computer predicted genes and the results of this analysis for the DNA sequence spanning 65.5% to 2.3% of the 186 chromosome are shown in Figure 1.2.

Analysis of the DNA sequence spanning 65.5% to 2.3% led to the prediction of a single leftward promoter p_L (2817 bp from the *Pst*I site at 65.5%; Kalionis *et al.*, 1986a). The transcript initiating from p_L is predicted to terminate at the *rho*-independent terminator t_L, to yield a transcript of 2180 bases. This transcript is predicted to encode genes *c*I, *int* and *CP69* (Figure 1.2(b)). A transcript of this size has been detected in a 186

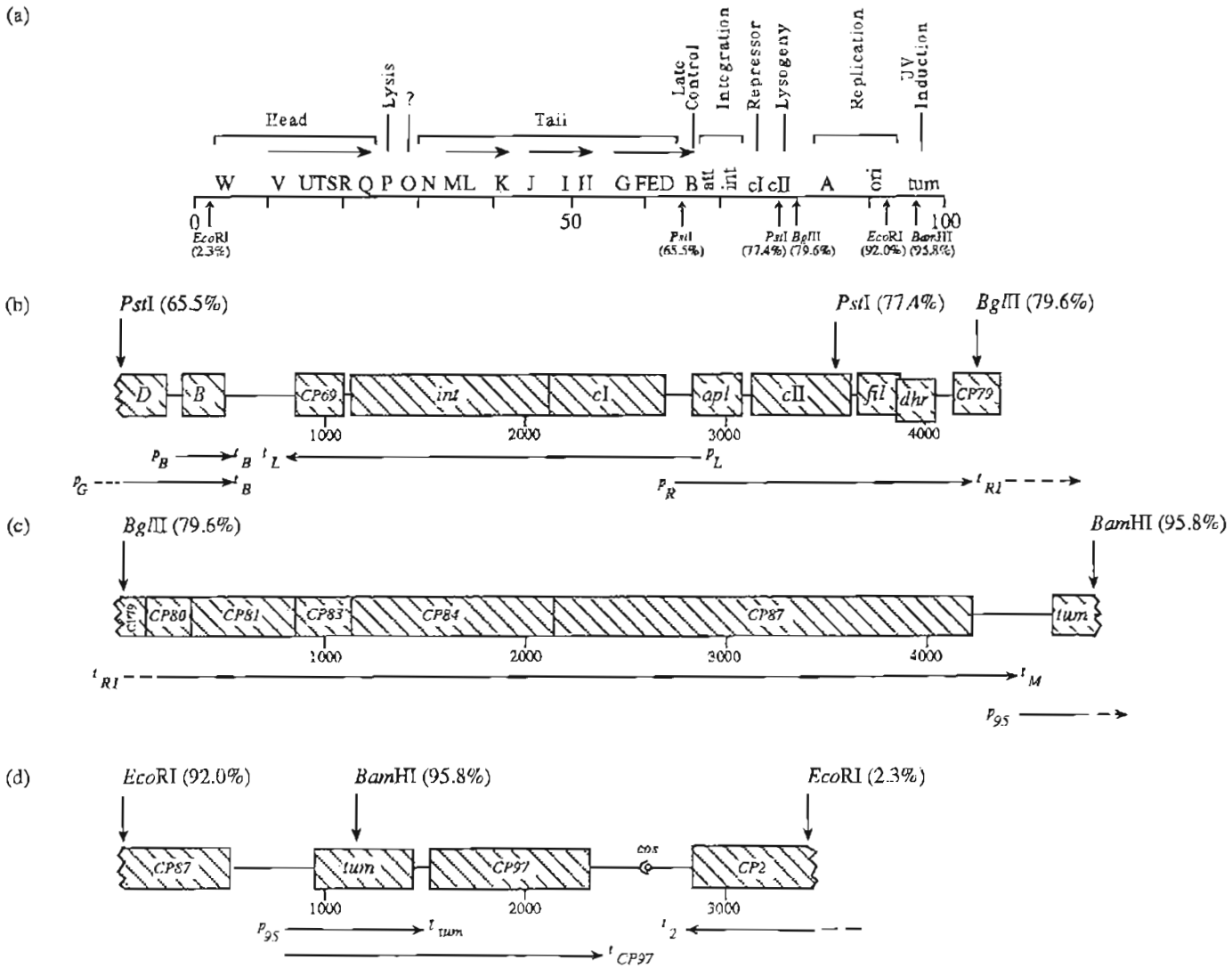
Figure 1.2 Diagrammatic representation of the DNA sequence spanning 65.5% to 2.3% of the 186 genome.

(a) Physical and genetic map of phage 186 (as described in Figure 1.1(a)) showing the restriction sites used in the determination of the DNA sequence spanning 65.5% to 2.3%.

(b) Representation of the predicted coding regions and transcriptional control signals as determined for the *Pst*I-*Pst*I (65.5% to 77.4%) region (Kalionis *et al.*, 1986a,b) and the *Pst*I-*Bgl*II (77.4% to 79.6%) region (Richardson *et al.*, 1989). Functions of the genes are described in Sections 1.3.1 and 1.4.2. Genes of unknown function are designated *CP* followed by the chromosomal coordinate approximating the predicted initiation codon of the gene. Promoters (*p*) and termination signals (*t*) are indicated. The *B* gene is predicted to be encoded on the transcript encoding late genes *GFED* late in lytic infection (Section 1.5.3(b); Kalionis *et al.*, 1986b).

(c) Representation of the predicted coding regions and transcriptional control signals as determined for the *Bgl*II-*BamH*I (79.6% to 95.8%) region (Sivaprasad *et al.*, 1990). The functions of the genes are described in Sections 1.3.1 and 1.4.2. Expression of genes *CP79* through to *CP87 in vivo* appears to be by antitermination of the *rho*-independent terminator t_{R1} (H. Richardson, 1987; R. Jarvinen, personal communication).

(d) Representation of the predicted coding regions and transcriptional control signals as determined for the *Eco*RI-*Eco*RI (92.0% to 2.3%) region (I. Lamont, personal communication). Functions of the *tum* gene and *CP97* are described in Sections 1.3.1 and 1.4.2, respectively. *CP2* is a leftward predicted coding region, presumably transcribed from a leftward late promoter and terminating at t_2. The position of the cohesive ends (*cos*) are shown.


(a)
Head
Lysis
?
Tail
Late Control
Integration
Repressor
Lysogeny
Replication
UV Induction
W V UTSR Q P O N ML K J I II G FED B att int cI cII A ori tum
0
50
100
EcoRI (2.3%)
PstI (65.5%)
PstI (77.4%) BglII (79.6%)
EcoRI (92.0%) BamHI (95.8%)
(b)
PstI (65.5%)
PstI (77.4%)
BglII (79.6%)
D
B
CP69
int
cI
apl
cII
fil
dhr
CP79
1000
2000
3000
4000
(c)
BglII (79.6%)
BamHI (95.8%)
CP79
CP80
CP81
CP83
CP84
CP87
tum
1000
2000
3000
4000
(d)
EcoRI (92.0%)
BamHI (95.8%)
EcoRI (2.3%)
CP87
tum
CP97
cos
CP2
1000
2000
3000

lysogen with a probe to this region (Kalionis *et al.*, 1986b) and the 5'-end determined by primer extension is consistent with initiation from p_L (Dodd *et al.*, 1990). This transcript is predicted to be the primary lysogenic transcript. The function of *CP69* is not known, but it is not required for integration, establishment or maintainence of lysogeny, or prophage induction (Crosby, 1985).

Three rightward promoters were predicted from the DNA sequence, p_B, p_R and p_{95}. The p_B promoter (263 bp from the *Pst*I site at 65.5%; Kalionis *et al.*, 1986a) is active *in vitro* (Pritchard and Egan, 1985) and is expected to give rise to a 290 base transcript encoding the *B* gene (Figure 1.2(b)). A transcript of this size has been identified *in vivo* (Kalionis *et al.*, 1986b), consistent with the position of the predicted *rho*-independent terminator t_B. The 5'-end of this transcript determined by primer extension analysis is consistent with initiation from p_B (Kalionis *et al.*, 1986b).

The p_R promoter (Figure 1.2(b); 2740 bp from the *Pst*I site at 65.5%; Kalionis *et al.*, 1986a) has been shown to be active *in vitro* (Pritchard and Egan, 1985) and the *in vivo* 5'-end has been determined by primer extension analysis (Dodd *et al.*, 1990). The p_R promoter is predicted to be the early lytic promoter (the first promoter expressed in lytic development) and to terminate transcription at the *rho*-independent terminator t_{R1} (Richardson *et al.*, 1989), consistent with the size of the *in vitro* transcript initiating from p_R (Pritchard and Egan, 1985). This early transcript is predicted to encode genes *apl*, cII, *fil* and *dhr* (Figure 1.2(b); Richardson *et al.*, 1989). There is evidence that the t_{R1} terminator is antiterminated *in vivo* (Richardson, 1987; R. Jarvinen, personal communication) and the transcript further encodes genes *CP79*, *CP80*, *CP81*, *CP84* and *CP87* (Figure 1.2(c)). Gene *apl* encodes a function which is involved in the maintainence of the lytic state by repression of the lysogenic promoter p_L (Dodd *et al.*, 1990). Genes *CP77* and *CP78* are non-esential. Expression of the *fil* gene results in an inhibition of *E. coli* cell division and therefore filamentation of the host, and expression of *dhr* results in an inhibition of *E. coli* DNA synthesis (Richardson and Egan, 1988). The function of *CP79* is not known. Genes *CP80*, *CP81*, *CP83* and *CP84* are non-essential (R. Jarvinen, personal communication). Gene *CP87* is the only phage gene required for the construction of a phage minichromosome (Sivaprasad *et al.*, 1990). Genes *CP80*, *CP81*, *CP83*, *CP84* and *CP87* are arranged such that the initiation codon of each gene overlaps the termination

codon of the previous gene, and thus are likely to be translationally coupled genes (Sivaprasad *et al.*, 1990; Normark *et al.*, 1983). This particular form of gene arrangement has been shown to be particularly important in the expression of *CP87*, which lacks a ribosome binding site, and depends upon the translation of *CP84* for expression (Sivaprasad *et al.*, 1990).

The p_R and p_L transcripts initially converge, then overlap and diverge (Figure 1.2(b)). DNA sequencing of 186 virulent mutants (able to form plaques in the presence of CI repressor) has located the CI binding site to a region overlapping the p_R promoter (Lamont *et al.*, 1988). Thus the binding of the repressor to the operator site is predicted to occlude the access of RNA polymerase to the p_R promoter and thus prevent expression of the 186 lytic genes.

Promoter p_{95} (Figure 1.2(d); 167 bp to the right of the *Pst*I site at 94.0%; Sivaprasad *et al.*, 1990) is under LexA control (Lamont *et al.*, 1989) and gives rise to two transcripts *in vitro* of 590 and 1540 bases (Pritchard and Egan, 1985) consistent with the positions of two *rho*-independent termination structures found in this region (I. Lamont, personal communication). The smaller transcript is predicted to encode only the *tum* gene, while the larger transcript is predicted to encode both *tum* and *CP97* (Figure 1.2(d)). The function of *CP97* is unknown but it may encode an immunity function (I. Lamont, personal communication).

1.5 CONTROL OF LATE GENE EXPRESSION IN PHAGE 186.

1.5.1 *In vivo* transcription studies of phage 186.

1.5.1(a) Temporal distribution.

Finnegan and Egan (1981) hybridized pulse-labelled RNA isolated at different times after prophage induction (by temperature increase) of wild-type or mutant lysogens, to cloned restriction fragments of known gene content (Finnegan and Egan, 1979) and determined the temporal distribution of RNA transcription from the 186 genome.

Transcription present in the prophage was from the region 67.9% to 74.9% (clone λpEC18; Figure 1.3(a)), which was expected since this region encoded the *c*I maintainence

Figure 1.3 Clone map of 186 as used in the *in vivo* transcription studies of Finnegan and Egan (1981).

(a) The Figure depicts the position and gene content of 186 restriction fragments cloned into pBR322 and used in the *in vivo* mRNA hybridization studies of Finnegan and Egan (1981). Chromosomal locations of restriction sites to the left of the *Pst*I site at 65.5% are as determined from the DNA sequence (Kalionis *et al.*, 1986a; Richardson *et al.*, 1989; Sivaprasad *et al.*, 1990).

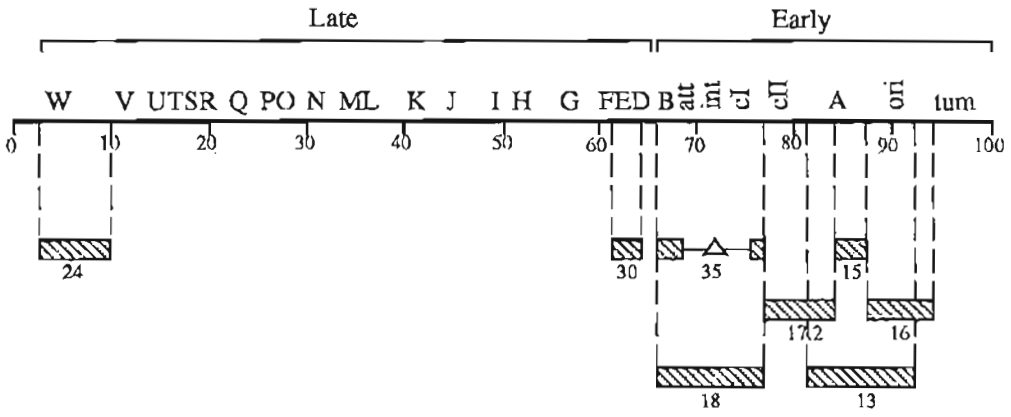
Plasmid pEC24 (Finnegan and Egan, 1979) carries the 2.3 kb *Eco*RI-*Bam*HI (2.3% to 10.0%) region and rescues alleles in gene *W* (Finnegan and Egan, 1979). Plasmid pEC30 (Finnegan, 1979) carries the *Pst*I-*Pst*I (61.0% to 64.0%) region and rescues alleles in gene *F* (Finnegan, 1979). Plasmids pEC24 and pEC30 were chosen as representative clones of the late region.

Plasmid pEC18 (Finnegan and Egan, 1979) carries the 3.6 kb *Pst*I-*Pst*I (65.5% to 77.4%) region. The gene content of this region as determined by Kalionis *et al.*, (1986a,b) is described in Section 1.4.2 and is shown in Figure 1.2(b). Plasmid pEC18 was found to be unstable (Finnegan and Egan, 1981) and the entire plasmid was subsequently cloned into the lambda vector λ762 and this clone (λpEC18) was used for *in vivo* mRNA hybridization studies. Plasmid pEC35 (Finnegan and Egan, 1981) carries the 1.8 kb *Pst*I-*Pst*I (65.5% to 77.4%) region isolated from 186 *del*1, which contains a deletion of the *int-c*I region spanning 67.9% to 74.0% (Kalionis *et al.*, 1986a). Plasmid pEC35 was assumed to relect transcription of the *B* gene.

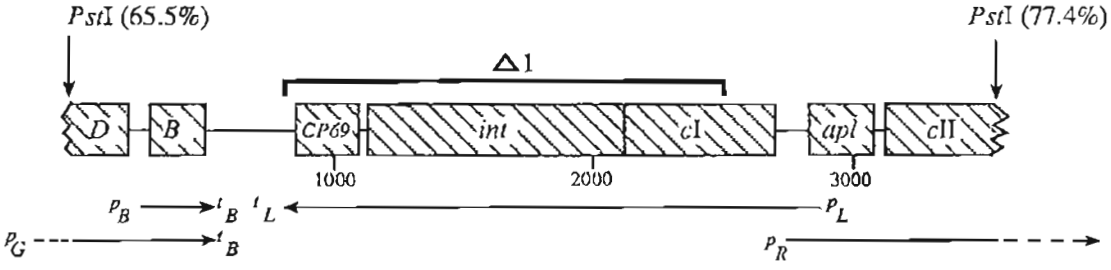
Plasmids pEC17.2, pEC13, pEC15 and pEC16 (Finnegan and Egan, 1979) span the early and middle regions of 186. Plasmid pEC17.2 carries the 2.2 kb *Pst*I-*Pst*I (77.4% to 84.6%) region. Plasmid pEC13 carries the 3.7 kb *Bgl*II-*Eco*RI (79.6% to 92.0%) region. Plasmid pEC15 carries the 0.9 kb *Pst*I-*Pst*I (84.6% to 87.5%) region. Plasmid pEC16 carries the 2.0 kb *Pst*I-*Pst*I (87.5% to 94.0%) region. Gene content of plasmids pEC17.2, pEC13, pEC15 and pEC16 is described in Section 1.4.2 and shown in Figure 1.2(b) and Figure 1.2(c).

(b) Diagrammatic representation of the genetic content of plasmid pEC35 and lambda clone λpEC18 as determined by Kalionis *et al.* (1986a), indicating the extent of the *del*1 deletion. Lambda clone λpEC18 contains the entire 3.6. kb *Pst*I-*Pst*I (65.5% to 77.4%) region. Plasmid pEC35 carries the equivalent restriction fragment from 186 *del*1. Functions of the genes is as described in Sections 1.3.1 and 1.4.2. The extent of the *del*1 deletion is indicated.

(a)


Late
Early
W V UTSR Q PO N ML K J I H G FED B att int cI cII A ori tum
0 10 20 30 40 50 60 70 80 90 100
24
30
35
15
17?2
16
18
13

(b)


PstI (65.5%)
PstI (77.4%)
Δ1
D
B
cro69
int
cI
apl
cII
1000
2000
3000
p_B
t_B
t_L
p_L
p_G
t_B
p_R

repressor and the *int* gene. Upon prophage induction, transcription was predominantly from the region spanning 65.5% to 100% (clones pEC17.2, pEC13, pEC15 and pEC16; Figure 1.3(a)). Transcription of this region reached a maximum at 30 minutes, and after this time the rate of transcription decreased markedly, suggesting a negative control of early gene transcription.

Transcription appearing late after prophage induction (20 to 25 minutes) was from the left two-thirds of the phage chromosome encoding the head, lysis and tail genes (Hocking and Egan, 1982c), as judged by hybridization to clones expected to be representative of the region 0% to 65.5% (clones pEC24 and pEC30; Figure 1.3(a)). Transcription of this region continued until cell lysis. In accordance with the fact that the region of the phage chromosome encoding the head, lysis and tail genes was transcribed late in infection, these genes were designated the late genes of 186.

1.5.1(b) Transcription after *Bam* prophage induction.

Induction of a prophage with a mutation in the *B* gene (*Bam*17) had two effects on transcription. Early gene transcription (65.5% to 100%; clones pEC17.2, pEC13, pEC15 and pEC16; Figure 1.3(a)) resembled that of the wild-type phage to 25 minutes. However, after this period the normal transition to late gene transcription (clones pEC24 and pEC30; Figure 1.3(a)) did not occur and it was concluded that transcription of the late region of 186 (0% to 65.5%) was dependent upon *B* gene function.

It was also found that transcription of the early region (65.5% to 100%) was at a higher level 35 minutes after heat induction, than that of wild-type. The higher level of transcription of this region from the *Bam* mutant compared with the wild-type led to the prediction that the *B* gene may be responsible for directly decreasing early gene transcription, or indirectly decreasing transcription of this region through another function under *B* control or by loss of phage templates due to DNA packaging and/or cell lysis.

1.5.1(c) Transcription after *Aam* prophage induction.

Induction of a 186 prophage with a defective replication gene (*Aam*30) resulted in transcription of the 65.5% to 100% region being reduced compared with the wild-type. This was predicted to be due to the decreased template number as a result of the mutation

in the replication gene. However, late gene transcription was essentially absent after induction of the *Aam* prophage and the appearence of late gene transcription was concluded also to be dependent upon the function of gene *A*.

Clone pEC35 (Figure 1.3(b)) is the *Pst*I fragment spanning 65.5% to 77.4% derived from 186 *del*1 DNA (Finnegan and Egan, 1981). This deletion removes 1835 bp from 67.9% to 74.0% (Kalionis *et al.*, 1986a), thus effectively deleting lysogenic transcription. The transcriptional activity of the pEC35 region (Figure 1.3(b)) was assumed to reflect transcription of the late control gene *B*, since lysogenic transcription was essentially absent. The transcriptional activity of this region decreased dramatically upon *Aam* prophage induction and led to the prediction that *B* gene transcription was dependent upon *A* gene function.

1.5.2 Identification of the 186 *B* gene and homology with the late control genes of phages P2 and P4.

1.5.2(a) Identification of the 186 *B* gene.

The late control gene *B* has been mapped to the right of gene *D* (Hocking and Egan, 1982b) and to the left of the *del*1 boundary at 67.9% (Kalionis *et al.*, 1986a). The DNA sequence spanning the *Pst*I-*Xho*I sites (65.5% to 67.6%) has been determined (Kalionis *et al.*, 1986b) and is shown in Figure 1.4. Computer assisted analysis of the DNA sequence spanning the *Pst*I-*Xho*I sites (65.5% to 67.6%) enabled the prediction of the potential gene *CP67* (Kalionis *et al.*, 1986a), which was identified as the *B* gene upon determination of the DNA sequence from *Bam*17 and *Bam*57 mutants (Kalionis *et al.*, 1986b). Translation of the DNA sequence demonstrated that the *B* gene coded for a slightly basic protein (9.7%) of 72 amino acids (Kalionis *et al.*, 1986a).

1.5.2(b) Homology between the late control proteins of 186, P2 and P4.

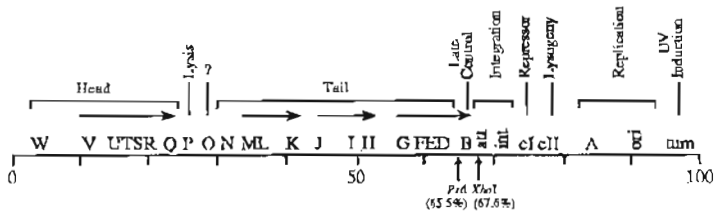
The late control gene *ogr* of P2 was identified by a dominant mutation that restored the ability of P2 to multiply in a host in which P2 late transcription was blocked (Sunshine and Sauer, 1975; Section 1.5.5). This mutation mapped to between the late tail gene *D* and the P2 attachment site (Sunshine and Sauer, 1975). DNA sequence analysis of the

Figure 1.4 Location and nucleotide sequence of the 186 *B* gene.

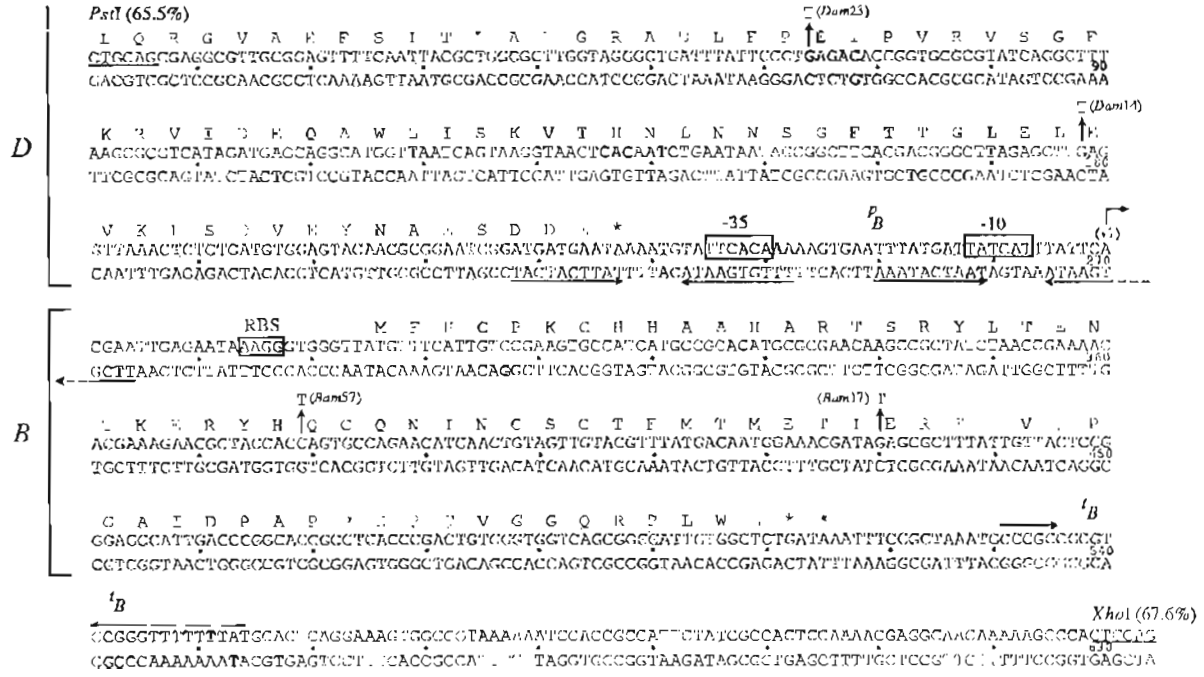
(a) Physical and genetic map of phage 186 as determined by Hocking and Egan (1982a), showing the location of the *Pst*I-*Xho*I (65.5% to 67.6%) region. Two of the four alleles of the *D* gene (*Dam*23 and *Dam*14) mapped to the right of the *Pst*I site at 65.5% (Finnegan and Egan, 1979).

(b) Nucleotide sequence of the *Pst*I-*Xho*I (65.5% to 67.6%) region. Predicted coding frames for the *D* and *B* genes are shown. The base changes associated with the amber mutations in the *D* gene (*Dam*23 and *Dam*14) and the *B* gene (*Bam*57 and *Bam*17) are indicated. Also indicated are the ribosome-binding site (RBS) for the *B* gene, the 5'-terminus of the *B* transcript (+1), the associated promoter (*p*) and termination signals (*t*), and the inverted repeat sequences overlapping the *B* promoter (represented by converging arrows), as identified by Kalionis *et al.* (1986b).

(a)


Head
Lysis
?
Tail
Late Control
Integration
Repressor
Lysogeny
Replication
UV Induction
W V U T S R Q P O N M L K J I H G F E D B att int cI cII A cro num
0
50
100
PstI XhoI
(65.5%) (67.6%)

(b)


PstI (65.5%)
D
B
RBS
T(Dam57)
(Dam17)
-35
-10
XhoI (67.6%)

mutations in the *ogr* gene, which overcame the host block to late transcription, identified the *ogr* gene (Birkeland and Lindqvist, 1986; Christie *et al.*, 1986). Translation of the DNA sequence of the *ogr* gene demonstrated that the P2 *ogr* gene also coded for a slightly basic protein (9.7%) of 72 amino acid residues.

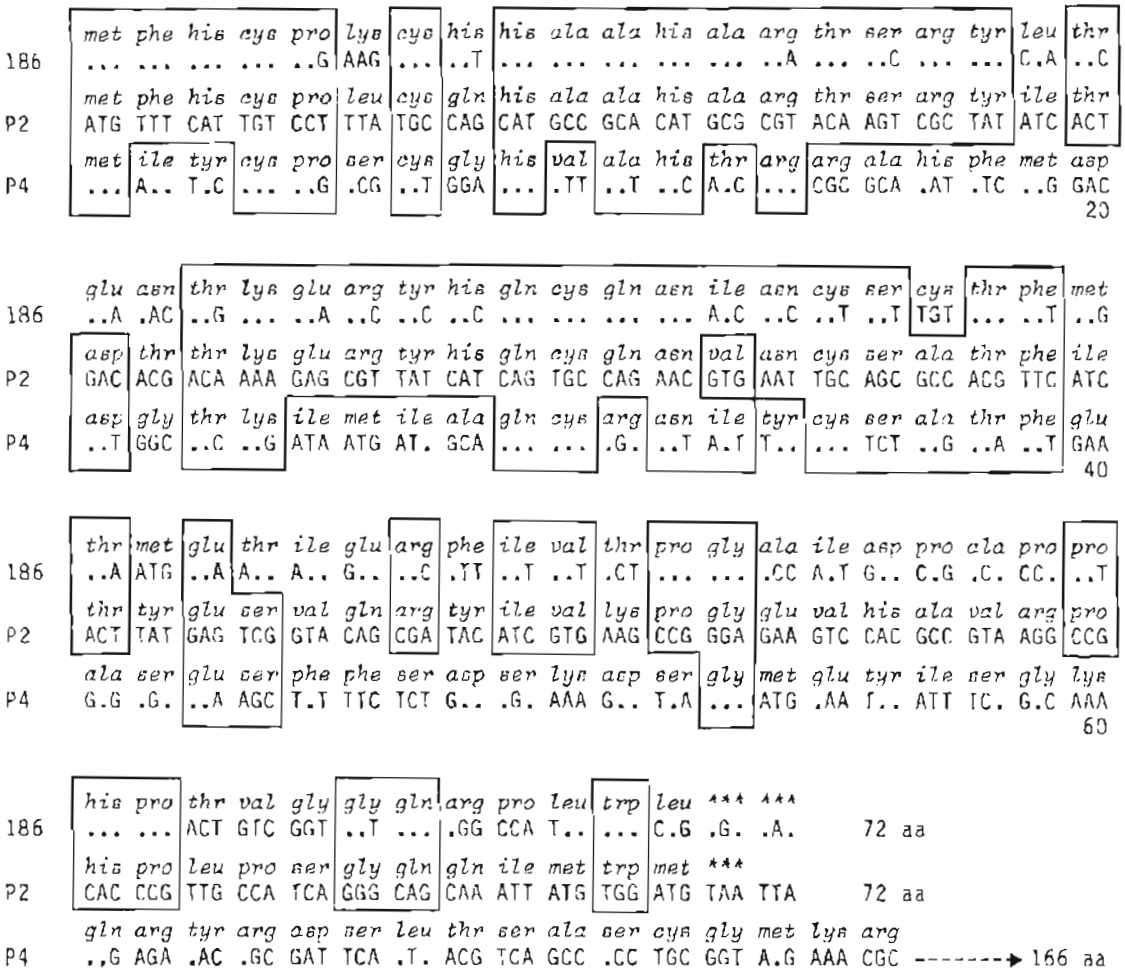
Extensive homology was found between the 186 B and P2 Ogr proteins (Figure 1.5; Kalionis *et al.*, 1986b). The 186 B protein and the P2 Ogr protein show 64% homology at the amino acid level, with 32 out of the 39 amino-terminal residues being identical (Kalionis *et al.*, 1986b). The homology is less for the remaining 33 carboxy-terminal amino acids with only 13 identical residues (Kalionis *et al.*, 1986b). The extensive homology was not surprising considering that both proteins are able to act at the same late promoters of P2, as demonstrated by the viability of hybrid P2-186 phages which only possess the *B* gene of 186 to activate the P2 late genes (Section 1.3.3; Hocking and Egan, 1982b).

The δ protein of P4 activates the same P2 promoters as the Ogr protein (Christie and Calendar, 1983) and can substitute for the 186 B protein, since it can activate the late genes of a 186 *Bam* helper (Sauer *et al.*, 1982). Consistent with the ability of P4 δ to substitute for 186 B, extensive homology was also found between the 186 B and P4 δ proteins (Figure 1.5; Kalionis *et al.*, 1986b). The P4 δ gene codes for a protein of 166 amino acids (Lin, 1984; Halling *et al.*, 1990), which is more basic (16.3%) than the B protein, and resembles a head to tail dimer of the B and Ogr proteins (Kalionis *et al.*, 1986b). The 186 B protein and the P4 δ protein have 18 identical residues in the first 39 amino-terminal residues, with only three identical in the remaining portion of the B protein (Figure 1.5; Kalionis *et al.*, 1986b). The larger size of the δ gene product may reflect its ability to activate late gene transcription of the helper in the absence of replication of the helper phage (Section 1.5.6; Six, 1975; Souza *et al.*, 1977; Sauer *et al.*, 1982), while activation of late gene transcription by the *B* and *ogr* gene products requires phage replication (Finnegan and Egan, 1981; Geisselsoder *et al.*, 1973; Lengyel and Calendar, 1974; Funnell and Inman, 1982).

Figure 1.5 DNA and amino acid homology of the 186 *B* gene with the P2 *ogr* gene and the P4 δ gene.

The Figure is reproduced from Kalionis *et al.*(1986b). Amino acid homology of the P2 *ogr* and the P4 δ genes to the 186 *B* gene is indicated by boxes and only nucleotides which differ from the DNA sequence of the 186 *B* gene are presented. For the purposes of this comparison, the numbering refers to the number of amino acid residues from the predicted initiation codon of each gene.

The 186 *B* gene sequence is as determined by Kalionis *et al.*, (1986b). The P4 δ sequence is from Lin (1984) and Halling *et al.* (1990). The P2 *ogr* sequence is from Birkeland and Lindqvist (1986) and Christie *et al.* (1986).


met phe his cys pro lys cys his his ala ala his ala arg thr ser arg tyr leu thr
186 ... ...G AAG ... ..T ... ...A ... ..C ... ... C.A ..C
met phe his cys pro leu cys gln his ala ala his ala arg thr ser arg tyr ile thr
P2 ATG TTT CAT TGT CCT TTA TGC CAG CAT GCC GCA CAT GCG CGT ACA AGT CGC TAT ATC ACT
met ile tyr cys pro ser cys gly his val ala his thr arg arg ala his phe met asp
P4 ... A.. T.C ... ..G .CG ..T GGA ... .TT ..T ..C A.C ... CGC GCA .AT .TC ..G GAC
20
glu asn thr lys glu arg tyr his gln cys gln asn ile asn cys ser cys thr phe met
186 ..A .AC ..G ... ..A ..C ..C ..C ... ... A.C ..C ..T ..T TGT ... ..T ..G
asp thr thr lys glu arg tyr his gln cys gln asn val asn cys ser ala thr phe ile
P2 GAC ACG ACA AAA GAG CGT TAT CAT CAG TGC CAG AAC GTG AAT TGC AGC GCC ACG TTC ATC
asp gly thr lys ile met ile ala gln cys arg asn ile tyr cys ser ala thr phe glu
P4 ..T GGC ..C ..G ATA ATG AT. GCA ... ... .G. ..T A.T T.. ... TCT ..G ..A ..T GAA
40
thr met glu thr ile glu arg phe ile val thr pro gly ala ile asp pro ala pro pro
186 ..A ATG ..A A.. A.. G.. ..C .TT ..T ..T .CT ... ... .CC A.T G.. C.G .C. CC. ..T
thr tyr glu ser val gln arg tyr ile val lys pro gly glu val his ala val arg pro
P2 ACT TAT GAG TCG GTA CAG CGA TAC ATC GTG AAG CCG GGA GAA GTC CAC GCC GTA AGG CCG
ala ser glu ser phe phe ser asp ser lys asp ser gly met glu tyr ile ser gly lys
P4 G.G .G. ..A AGC T.T TTC TCT G.. .G. AAA G.. T.A ... ATG .AA T.. ATT TC. G.C AAA
60
his pro thr val gly gly gln arg pro leu trp leu *** ***
186 ... ... ACT GTC GGT ..T ... .GG CCA T.. ... C.G .G. .A. 72 aa
his pro leu pro ser gly gln gln ile met trp met ***
P2 CAC CCG TTG CCA TCA GGG CAG CAA ATT ATG TGG ATG TAA TTA 72 aa
gln arg tyr arg asp ser leu thr ser ala ser cys gly met lys arg
P4 ..G AGA .AC .GC GAT TCA .T. ACG TCA GCC .CC TGC GGT A.G AAA CGC --------→ 166 aa

1.5.3 Transcriptional studies of the *B* gene.

1.5.3(a) *In vitro* transcription of the *B* gene.

The *B* gene transcript is generated *in vitro* by transcription of 186 DNA by *E. coli* RNA polymerase holoenzyme and is the most prominent *in vitro* transcript on a molar basis (Pritchard and Egan, 1985). The *B* gene was not expected to be transcribed *in vitro*, since transcription of the *B* gene was concluded to be dependent upon *A* gene function from *in vivo* transcription studies (Section 1.5.1(c); Finnegan and Egan, 1981). However, upon determination of the DNA sequence of the *B* gene (Figure 1.4; Kalionis *et al.*, 1986a), the promoter was found to have significant homology to the consensus sequence for *E. coli* promoters (Hawley and McClure, 1983) and was thus expected to be transcribed *in vitro* by unmodified *E. coli* RNA polymerase holoenzyme.

1.5.3(b) B gene transcription in the lysogenic state and after prophage induction.

Using Northern analysis, Kalionis *et al.* (1986b) found that the 290 base transcript encoding the *B* gene is present in the lysogenic state. This was consistent with the *B* promoter being active *in vitro* and the fact that there is no evidence of CI repressor control to the left of the *att* site (Hocking and Egan, 1982c). However, the level of transcription detected in the lysogen was considerably less than expected on the basis of *in vitro* studies (Pritchard and Egan, 1985), suggesting repressive control of *B* gene transcription. The level of transcript detected from a non-suppressing lysogen of a *Bam* mutant was found to be markedly increased with respect to the wild-type lysogen and Kalionis *et al.* (1986b) concluded that transcription of the *B* gene is repressed by the *B* protein itself (repressive autogenous control), either directly or indirectly.

After prophage induction, the level of the 290 base transcript encoding the *B* gene increased with time during the latent period, as determined by Northern analysis (Kalionis *et al.*, 1986b). Late in the latent period (35 minutes after prophage induction), the level of transcript was found to markedly increase, and this was considered to be most likely due to the increase in gene dosage as a result of phage replication (Kalionis *et al.*, 1986b).

The *B* gene also appeared to be encoded on a larger transcript of approximately 2 kb in size late in the latent period (Kalionis *et al* ., 1986b). No structure consistent with a

rho-independent terminator was found in the DNA sequence in the intergenic region between late gene *D* and the *B* gene, which indicated that rightward late transcription of gene *D* would continue into gene *B* and terminate at t_B (Kalionis *et al.*, 1986b). Polarity studies had indicated that genes *GFED* were rightwardly transcribed as a unit (Hocking and Egan, 1982c). The 2 kb transcript encoding the *B* gene has been shown to initiate transcription to the left of the *Hin*dIII site at 61.3% by Northern analysis (Kalionis *et al.*, 1986b) and the size of this transcript indicated that it probably initiated transcription around 60.9%. However, alleles of gene *G* map to the left of the *Bam*HI site at 58.7% (Finnegan and Egan, 1979), which would require the transcript to be at least 2670 bases in size. Nevertheless it is likely that the *B* gene is represented on a late transcript and since the late promoter for this transcript is likely to be under positive control by the *B* gene product (Finnegan and Egan, 1981), Kalionis *et al.* (1986b) predicted that the *B* gene positively controls its own transcription late in infection (positive autogenous control).

1.5.4 Promoters under *B* gene control.

In the lysogenic state the *B* gene product appears to repress its own transcription, either directly or indirectly (Section 1.5.3(b); Kalionis *et al.*, 1986b). An inverted ten base repeat was found overlapping the -10 region of the *B* promoter and a second copy of this repeat was found overlapping the -35 region (Figure 1.6). Kalionis *et al.* (1986b) indicated that these inverted repeats were likely to be the presumptive binding sites for the protein involved in repressing *B* gene transcription. The simplest model for the autogenous control of *B* gene transcription, would be that the B protein behaves as a classical repressor at its promoter by binding to the inverted repeat sequences to prevent transcription (Kalionis *et al.*, 1986b).

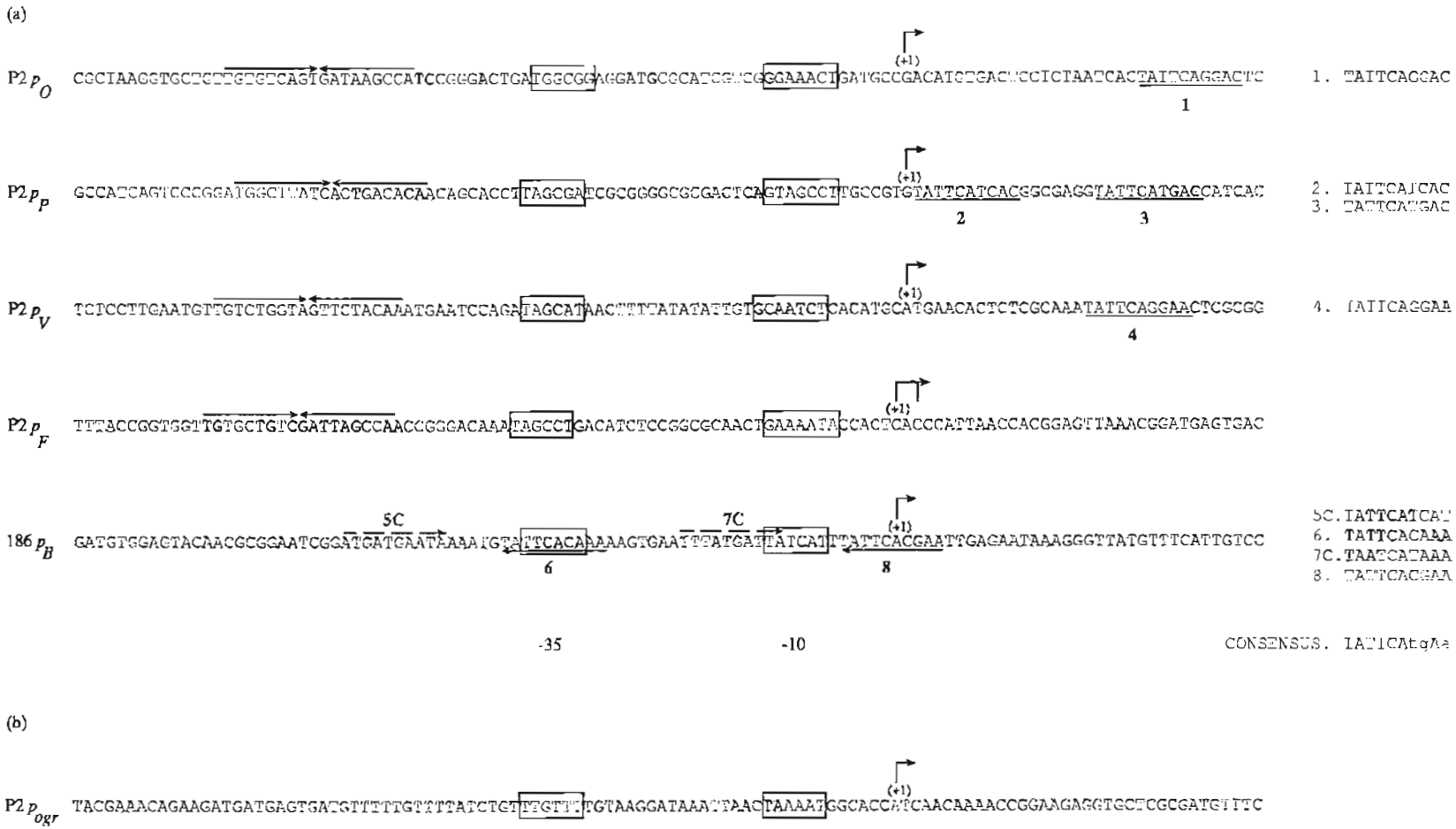
Late gene transcription in 186 is dependent on the protein product of the *B* gene (Finnegan and Egan, 1981). The *B* gene product may act as a positive control factor at the late promoters and therefore the late promoters of 186 may provide evidence for a presumptive *B* gene product recognition sequence. The DNA sequence of the 186 late promoters has not been determined and the characterization of a 186 late promoter is one of the concerns of this work. However, the B protein is known to activate transcription of the P2 late promoters, since viable P2-186 hybrid phages exist that only possess the 186 *B*

Figure 1.6. Comparison of the DNA sequences of the 186 *B* promoter, the P2 late promoters and the P2 *ogr* promoter.

(a) Promoters under *B* gene control. The Figure is essentially the same as presented in Kalionis *et al.* (1986b). Promoter regions of the P2 late genes *O*, *P*, *V*, and *F* (Christie and Calendar, 1983; Christie and Calendar, 1985) and the promoter of the 186 *B* gene (Kalionis *et al.*, 1986b) are presented with the nucleotide sequences aligned at the proposed mRNA start points (+1) of each gene. The -10 and -35 regions of the 186 *B* promoter are boxed. Conserved regions in the P2 late promoters about the -10 and -35 positions are also boxed (Christie and Calendar, 1985). Dale *et al.* (1986) identified an inverted repeat element centered at the -55 position in the P2 late promoters, represented by converging arrows.

Sequences showing homology to the *B box* (Kalionis *et al.*, 1986b) are shown. Regions 1, 2, 3, 4, 6 and 8 are found in the left strand of the sequence and regions 5C and 7C are found in the complementary strand. The *B box* sequences in the *B* promoter form two inverted repeat elements (Kalionis *et al.*, 1986b) shown by converging arrows. The consensus sequence for the *B box* (as determined by Kalionis *et al.*, 1986b) is shown on the right.

(b) DNA sequence of the P2 *ogr* promoter (Birkeland and Lindqvist, 1986; Christie *et al.*, 1986). The proposed mRNA start point (+1) and -10 and -35 regions of the *ogr* promoter (boxed regions) are shown.


(a)
P2 p_O
P2 p_P
P2 p_V
P2 p_F
186 p_B
(+1)
1
2
3
4
5C
6
7C
8
-35
-10
1. TAITCAGCAC
2. IAITCAICAC
3. TAITCATGAC
4. TAITCAGGAA
5C. IATTCATCAT
6. TATTCACAAA
7C. TAATCATAAA
8. TAITCACGAA
CONSENSUS. IATTCAtgAa
(b)
P2 p_{ogr}

gene for the activation of P2 late genes (Section 1.3.3; Hocking and Egan, 1982b). The DNA sequence of four P2 late promoters have been determined (Christie and Calendar, 1983; 1985) and are shown in Figure 1.6). The promoter regions were found to differ considerably from the *E. coli* consensus sequence, consistent with the promoters being positively controlled. In the DNA sequence of the late promoters several regions of homology were evident about the -10 and -35 regions (Christie and Calendar, 1985) and the late promoters also shared elements of an inverted repeat structure centered at the -55 position (Dale *et al.*, 1986). Three of the four late promoters also shared a highly conserved ten nucleotide sequence downstream from the site of initiation of transcription, which corresponded to one arm of the inverted repeat found overlapping the *B* promoter (Figure 1.6). Kalionis *et al.* (1986b) have termed this ten base sequence the *B box*. The implication is that the product which is involved in repressing *B* gene transcription, by binding to the inverted repeats overlapping the *B* promoter, may also bind downstream of the late promoters and could therefore have a role in controlling transcription from the late promoters.

The P2 late promoters are normally under positive control of the Ogr protein (Sunshine and Sauer, 1975). Since the 186 B protein and the Ogr protein have extensive amino acid homology (Kalionis *et al.*, 1986b) and are functionally interchangeable (Hocking and Egan, 1982b), it is possible that the *B box* is also recognized by P2 Ogr. However, the *ogr* promoter does not contain the ten base sequence or any other detectable inverted repeat (Figure 1.6(b)) and therefore may not be autogenously controlled (Kalionis *et al.*, 1986b). In spite of this fact, there is data compatible with autogenous control at the *ogr* promoter (Birkeland and Lindqvist, 1986; Birkeland *et al.*, 1988).

1.5.5 Interaction of the late control proteins with the alpha subunit of *E. coli* RNA polymerase.

Sunshine and Sauer (1975) isolated a mutant of *E. coli* C, *gro*109, which blocked P2 propagation by interfering with late gene transcription, despite the fact that P2 could still replicate in this strain. The *ogr* gene of P2 is defined by *trans*-dominant mutations that overcome the block imposed by the *gro*109 mutation and restore P2 late transcription and protein production. The *gro*109 was found to be structural mutation in the gene

encoding the α subunit of RNA polymerase and was therefore designated *rpoA*109 (Fujicki *et al.*, 1976). This indicated that positive control by the *ogr* gene product involved an interaction with one or both of the α subunits of RNA polymerase.

DNA sequence analysis of independent *ogr* mutations, *ogr*1 and *ogr*52, demonstrated that the mutations were identical missense mutations at amino acid 42 in the *ogr* gene (Birkeland and Lindqvist, 1986; Christie *et al.*, 1986). The *rpoA*109 mutation also interferes with P2-assisted growth of P4 (Sauer *et al.*, 1982) and P4 mutations which supress the *rpoA*109 have missense mutations in the δ gene (Halling *et al.*, 1990), which is involved in transactivation of helper late genes (Souza *et al.*, 1977).

Unlike P2, wild-type 186 can multiply on the *rpoA*109 strain (Sauer, 1979). However, since the 186 *B* gene and the P2 *ogr* gene show extensive amino acid homology (Section 1.5.2; Kalionis *et al.*, 1986b) and are functionally interchangeable (Section 1.3.3; Hocking and Egan, 1982b), a similar interaction with the α subunit is expected for the B protein. No specific role has been assigned for the α subunit of *E. coli* RNA polymerase and it is rarely implicated as having a role in selective transcription. Thus, these control proteins provide a rare opportunity to study control mediated through the α subunit of RNA polymerase.

1.5.6 Late gene transcription by satellite phage P4.

Satellite phage P4 requires all the known morphogenetic genes and lysis genes of a 186 or P2 helper phage for a productive infection (Sauer *et al.*, 1982). P4 infection of a P2 lysogenic cell causes derepression of the P2 prophage (Six and Lindqvist, 1978). Since P2 excision is inefficient, P2 replication occurs *in situ* in the chromosome (Six and Lindqvist, 1978), and the consequent derepression and replication cause the P2 prophage to express its late genes for use by P4. Thus P4 can activate P2 late genes by derepressing P2 prophage and thereby causing expression of P2 late genes under their normal mode of control (Six and Lindqvist, 1978; Geisselsoder *et al.*, 1981). The P4 ε gene is responsible for the derepression of a P2 prophage (Geisselsoder *et al.*, 1981). However, unlike P2, a repressed 186 prophage does not support P4 growth (Sauer *et al.*, 1982).

Derepression of a helper prophage that is unable to replicate will not lead to the expression of the prophage late genes, since late gene transcription in 186 and P2 is

dependent upon phage replication (Finnegan and Egan, 1981; Geisselsoder *et al.*, 1973; Lengyel and Calendar, 1974; Funnell and Inman, 1982). However, under such circumstances, P4 utilises a second mechanism, transactivation, to activate helper late genes (Six,1975). P4 can transactivate the late genes of a replication-deficient P2 or 186 helper (Six, 1975; Souza *et al.*, 1977; Sauer *et al.*, 1982). The use of a replication deficient helper lengthens the latent period of P4 by about twenty minutes, but does not adversely affect the P4 burst size (Bertani and Six, 1988). The δ gene of P4 is responsible for transactivation of both 186 and P2 and can substitute for the 186 *B* gene (Sauer *et al.*, 1982) and the P2 *ogr* gene (Sauer *et al.*, 1982; Halling *et al.*, 1990; Halling and Calendar, 1990), which both show considerable homology to the δ gene (Kalionis *et al.*, 1986b). During transactivation, the sites of initiation of transcription from the P2 late promoters p_O and p_P are the same as during normal P2 development (Christie and Calendar, 1983) and suggests that the late promoters utilised by δ during transactivation of P2 are the same as utilised by Ogr. Transactivation is also dependent on the function of the P4 α gene (Gibbs *et al.*, 1983), which is required for P4 DNA replication (Gibbs *et al.*, 1973), but this is probably an indirect requirement of replication, as transcription of the δ gene depends on α gene function (Harris and Calendar, 1978).

Expression of the P4 late genes (*sid*, δ and *psu*) from the P4 late promoter p_{sid} is also reciprocally transactivated by a P2 helper (Lindqvist, 1974; Barrett *et al.*, 1976; Souza *et al.*, 1977; Harris and Calendar, 1978). The *ogr* gene of P2 is responsible for this reciprocal transactivation of P4 by P2, since expression of the *ogr* gene on a plasmid *in vivo* is sufficient to activate the P4 late promoter p_{sid} (Dale *et al* ., 1986), and Keener *et al.* (1988) have shown that the *ogr* gene product activated transcription from the p_{sid} promoter *in vitro* in the absence of other phage functions.

1.5.7 Control of late transcription in 186.

(i) Role of replication in late transcription

The *B* gene is transcribed in the lysogenic state and its transcription appears to be under autogenous control (Kalionis *et al.*, 1986b). There is no evidence of late gene transcription in the lysogenic state from *in vivo* transcription studies (Finnegan and Egan, 1981) and this raises the question as to why there is no late gene transcription in the

lysogen, even though B protein is present and active in repressing its own transcription. The lack of late transcription in the lysogenic state was consistent with the absence of gene *A* function, due to CI repression, since gene *A* function is also required for late transcription (Finnegan and Egan, 1981)

In the marker rescue experiments of Hocking and Egan (1982a) it was also found that B protein was unable to activate the late promoters of a prophage. In those experiments, it was found that a superinfecting, immunity-insensitive phage mutant in a late gene function (e.g. 186 *vir*2 *Dam*14) could not rescue by complementation the appropriate late function from a prophage, in spite of the fact that the superinfecting phage is insensitive to the 186 repressor and therefore can replicate and presumably provide excess B function required for activation. Kalionis *et al.* (1986b) concluded that the inability of B protein to activate the late promoters was due to the absence of *A* gene function, which presumably provided a template topology required for B to activate the late promoters.

(ii) Requirement of another function under *c*I control

Kalionis *et al.* (1986b) also considered the possibility that another function (apart from genes *B* and *A*) was required for activation of late transcription. The need of a third function was based on the fact that satellite phage P4 could not transactivate a repressed 186 lysogen (Sauer *et al.*, 1982). This inability to transactivate a 186 lysogen was not due to the inability of the P4 δ function to activate the late promoters of 186, since P4 could transactivate a heat-induced, non-permissive 186 *cIts Bam*17 lysogen (Sauer *et al.*, 1982). Nor was it due to the inability of P4 to rescue sufficient late functions from a single copy of 186, since P4 could transactivate a heat-induced, replication-defective (*Aam*) lysogen of 186 *cIts* (Sauer *et al.*, 1982). To explain these data, Kalionis *et al.* (1986b) suggested that the activation of the late promoters by the P4 δ protein required a function under CI repressor control, and that a similar need by 186 B protein would explain the inability of B protein to activate prophage late promoters in the marker rescue experiments of Hocking and Egan (1982a). In such a case, this function would necessarily be *cis*-acting. However, it was also possible that a 186 lysogen produced a function that inhibited P4 development in some manner and that this was the reason for the inability of P4 to transactivate a repressed 186 lysogen.

1.6 POSITIVE CONTROL OF TRANSCRIPTION INITIATION.

1.6.1 Positive control.

Positive control of gene expression is postulated to occur when loci are identified genetically where mutations cause a reduction in the level of expression of other genes. Positive control at the level of transcription has been well documented (Englesberg and Wilcox, 1974; Raibaud and Schwartz, 1984; Busby, 1986; Galloway and Platt, 1985; Platt, 1986; von Hippel *et al.*, 1984). Positive control of transcription may be exerted at the level of initiation (promoter activation) or at the level of termination (antitermination).

This thesis aims to characterize the control of late gene transcription in phage 186. Control of late gene transcription is believed to be by positive regulation at the level of transcription initiation (Section 1.5.4). For this reason, it is appropriate to discuss examples of those systems where positive control is exerted at the level of transcription initiation.

1.6.2 Promoter activation.

Promoter activation is the stimulation of the frequency of initiation from a promoter. Activated promoters require either the participation of some factor that is not always present or not always active. The activator proteins characterized to date can be catagorized into three distinct classes:

(i) New RNA Polymerases

Activation of transcription may be by the production of an entirely new RNA polymerase. The late genes of bacteriophages T3 and T7 are positively regulated by phage-encoded RNA polymerases (Oakley and Coleman, 1977; Dunn and Studier, 1983). The early genes of bacteriophage N4 are transcribed by a phage-encoded, virion encapsulated RNA polymerase and two of the products of N4 early transcription constitute a RNA polymerase required for middle gene expression (Haynes and Rothman-Denes, 1985; Zehring *et al.*, 1983; Zehring and Rothman-Denes, 1983).

(ii) Sigma factors

Positive control of transcription initiation may be through the use of novel σ factors, which positively regulate gene expression by specifically altering promoter

specificity. Examples of positive control involving new σ factors have been observed in *B. subtilis* sporulation (Goldfarb *et al.*, 1983; Johnson *et al.*, 1983; Wong and Doi, 1982), in *E. coli* in the control of the heat shock response by the product of the *rpoH* gene (Grossman *et al.*, 1984; Landick *et al.*, 1984), and in bacteriophage development in the lytic development of the *B. subtilis* phage SPO1 and the activation of late gene transcription in the phage T4 by gene 55 protein (reviewed by Geiduschek and Kassavetis, 1988; Stewart, 1988).

(iii) Accessory factors

The majority of the characterized activator proteins are accessory factors and in this class of activator proteins, participation of the σ^{70} subunit of RNA polymerase is still required for transcription. Examples of accessory factors include the *E. coli* proteins CRP (reviewed by de Crombugghe *et al.*, 1984), AraC (Englesberg and Wilcox, 1974), OmpR (Hall and Silhavy, 1979; 1981a; 1981b), and MalT (Debarbouille *et al.*, 1978; Richet and Raibaud, 1989); the λ CI and CII proteins (Gussin *et al.*, 1983; Hochschild *et al.*, 1983; Wulff and Rosenberg, 1983; Ho *et al.*, 1983; 1986; Ho and Rosenberg, 1988); the T4 middle control protein, Mot (Uzan *et al.*, 1983; Rabussay, 1983; Brody *et al.*, 1983; Pulitzer *et al.*, 1985); the Mu late control protein C (Margolin and Howe, 1986; Margolin *et al.*, 1989); and the P2 Cox protein (Saha *et al.*, 1989). The late control proteins of phages 186, P2 and P4 (B, Ogr and δ respectively) are also likely to be accessory factors, since they appear to show little homology to known σ subunits (Christie *et al.*, 1986; Halling *et al.*, 1990).

1.6.3 Structure of activated promoters and activator binding sites.

In *E. coli*, the promoters recognized by RNA polymerase holoenzyme alone, show DNA sequence homology upstream of the site of transcription initiation. In particular, two well conserved hexanucleotide sequences are located about the positions -10 and -35 relative to the transcription start point (Rosenberg and Court, 1979; Hawley and McClure, 1983). Consistent with the fact that positively controlled promoters are not fully functional or completely inactive in the presence of RNA polymerase alone, the DNA sequences of positively controlled promoters deviate significantly from the *E. coli* consensus promoter (Raibaud and Schwartz, 1984; Busby, 1986).

Promoters requiring an entirely new RNA polymerase or a new σ factor show very little homology to the *E. coli* consensus sequence (reviewed by Chamberlin and Ryan, 1982; Reznikoff *et al.*, 1985; Kustu *et al.*, 1986; Grossman and Losick, 1986). Promoters dependent upon accessory proteins for their function have some similarity to the *E. coli* consensus in the -10 region, but differ markedly in the -35 region, particularly if the accessory protein binds to distinct binding sites on the DNA located close to or overlapping the -35 region of the promoter (Raibaud and Schwartz, 1984; Busby, 1986; Raibaud *et al.*, 1985; Gussin *et al.*, 1983; Ho *et al.*, 1986; Brody *et al.*, 1983). However, they usually possess one residue of the TTG motif in the -35 hexamer consensus sequence (TTGACA) and at least two of the three most highly conserved residues TA---T in the -10 hexamer consensus sequence (TATAAT) (McClure, 1985).

The majority of characterized transcriptional activators which are accessory factors, appear to bind to DNA at various distances upstream of the -35 region (Raibaud and Schwartz, 1984; Busby, 1986). The exceptions to this are the λ CII protein and the related P22 CI protein, which both bind in the region spanning the -35 element (Ho *et al.*, 1983; Ho and Rosenberg, 1988) and the *E. coli* MerR protein, which has been shown to activate transcription of the *merT* promoter by binding between the -35 and -10 regions (O'Halloran *et al.*, 1989). The activator binding sites may show a hyphenated inverted repetitive structure, such as for λ CI repressor and *E. coli* CRP binding sites (Raibaud and Schwartz, 1984), a direct repeat separated by a few bases as for the λ CII and *E. coli* PhoB binding sites (Ho *et al.*, 1986; Surin *et al.*, 1984) or can be unstructured as for the T4 Mot activator protein binding site (Brody *et al.*, 1983). The affinities of transcriptional activators for their DNA binding sites are often low, sometimes several orders of magnitude lower than those of repressors (Ho and Rosenberg, 1988; Maeda *et al.*, 1988; Raibaud *et al.*, 1989).

1.6.4 Mechanism of promoter activation by RNA polymerase accessory factors.

The first step in the initiation of transcription is the binding of RNA polymerase to promoter DNA. RNA polymerase binds non-specifically to DNA and is believed to diffuse along the DNA until it encounters a promoter. A closed complex is then formed upon binding of the RNA polymerase to the promoter. The equilibrium constant for the formation of the closed complex (K_B) ranges from 10^6 M^{-1} to 10^9 M^{-1} (McClure, 1985).

A transcriptionally competent state, known as the open complex, is formed upon isomerization of the closed complex and involves unwinding of the DNA at the site of initiation. The rate constant for the formation of the open complex (k_f) ranges from 10^{-3} s^{-1} to 10^{-1} s^{-1} (McClure, 1985). The product of the equilibrium constant for closed complex formation and the rate constant for open complex formation (K_B x k_f) provides a measure of the strength of a promoter.

Activators may act by increasing the affinity of the polymerase for the promoter or by increasing the rate of the isomerization step (reviewed by Hoopes and McClure, 1987). The CI protein of λ increases the isomerization rate at the p_{RM} promoter by approximately eleven fold (Hawley and McClure, 1982). The cyclic AMP binding protein CRP increases K_B at the *lac* operon promoter, P1, by approximately twenty fold (Malan *et al.*, 1984). The CII protein of λ activates at its target promoters by increasing both the binding and isomerization of RNA polymerase (Shih *et al.*, 1984).

It has been proposed that the activator could conceivably achieve changes in the binding and isomerization constants by either causing a conformational change in the promoter DNA or by favourable protein-protein interaction. Several activator proteins have been shown to stimulate transcription by contacting RNA polymerase bound downstream at the adjacent promoter (Guarente *et al.*, 1982; Hochschild *et al.*, 1983) and appear to contact RNA polymerase through an acidic patch on the protein surface (Bushman and Ptashne, 1988). The binding of λ CII, the *E. coli* MerR and CRP activator proteins to the activator site do cause a conformational change in the DNA, which is also expected to be important in activation (Ho *et al.*, 1983; Gartenberg and Crothers, 1988; O'Halloran *et al.*, 1989). Such alterations in the DNA conformation upon binding of the activator may be a general aspect of transcriptional activation (Travers and Klug, 1987) and may reflect a requirement of the activator protein to introduce a local change into the promoter DNA to bring the -10 and -35 sites into better alignment, as has been proposed for activation by the MerR protein (O'Halloran *et al.*, 1989). However, the mechanism by which the protein-protein interactions and the DNA structural changes increase the binding of RNA polymerase and/or open complex formation at activated promoters have not yet been fully elucidated.

1.6.5 Regulation of activators.

In most cases the activity or expression of activators is strictly controlled (Busby, 1986; Raibaud and Schwartz, 1984). Expression of activators is often controlled by other positive regulators, or by transient expression as a result of temporal development as seen during phage development. In many cases the activity of an activator is modulated by the binding of a ligand, to effect a change in conformation of the activator protein. Binding of the ligand may allow activation to occur, as seen in the activation of the *araBAD* promoter by AraC in the presence of arabinose (Lee *et al.*, 1981) and the activation of transcription of the *dsdA* gene (involved in D-serine deaminase synthesis) by the protein product of the *dsdC* gene in the presence of D-serine (Mcfall, 1987). Binding of the ligand may also abolish activation. Asparagine, the product of the asparagine synthetase gene (*asnA*), turns off the stimulation of *asnA* transcription by the activator AsnC (Kolling and Lother, 1985). Covalent modification may also be used to alter activator activity. The *adaA* gene product activates the genes involved in the adaptive response to alkylating agents in *E. coli* and is irreversibly activated by a post-translational methylation of a cystein residue (Teo *et al.*, 1986). The *adaA* gene product positively controls its own transcription upon conversion of the protein to the activator form by methylation (Teo *et al.*, 1986).

Finally, negative autoregulation is often used to regulate activator expression (reviewed by Raibaud and Schwartz, 1984), enabling the concentration of activator to be maintained at a constant level. Examples of negative autoregulation of expression of activators are seen in the expression of the *E. coli* genes *araC* (Lee *et al.*, 1981), *asnC* (Kolling and Lother, 1985) and *merR* (Ni'Bhriain *et al.*, 1983). The 186 late activator gene *B* also appears to negatively regulate its own transcription (Kalionis *et al.*, 1986b).

1.7 AIMS AND APPROACH.

The purpose of this work was to provide a further understanding of the control of late gene transcription in phage 186. Control of late transcription appears to require an activator function, phage replication and the involvement of the α subunit of RNA polymerase. Thus study of the control of late transcription may provide evidence for a novel form of promoter activation and allow further understanding of the role of the α subunit in the transcription process.

The specific aims were as follows:

(1) To identify the phage functions required directly for late transcription.

(2) To investigate the requirement of phage replication in the activation of late transcription.

(3) To investigate the control of transcription of the late control function *B* and the implications of this control for activation of late transcription.

The first step in the investigation of control of late transcription was the identification and characterization of a 186 late promoter. Evidence from polarity studies and marker rescue studies indicated that the *Bam*HI-*Eco*RI (10.0% to 13.3%) region was likely to encode the promoter for the late transcription unit encoding genes *VUTSRQ* ((Hocking and Egan, 1982c; Finnegan and Egan, 1979). Prior to this work, preliminary DNA sequence determination leftward (with respect to the genetic map) from the *Eco*RI site at 13.3% had allowed the identification of a rightward open reading frame, not completely contained on the *Bam*HI-*Eco*RI (10.0% to 13.3%) restriction fragment, which was identified as the *V* gene upon determination of the base-pair change responsible for the *Vam*38 mutation (Dibbens, 1984). The promoter for the *V* gene was also located by preliminary primer-extension analysis (Dibbens, 1984).

The approach taken in this work was to complete the DNA sequence of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region and to locate potential genes and transcriptional signals by computer assisted analysis. Determination of the DNA sequence enables specific probes to be constructed, allowing the identity and activity of potential transcripts to be confirmed. This should allow the identification and characterization of a 186 late promoter.

Late gene transcription in 186 has been shown to be dependent upon the functions of gene *B* and replication gene *A* (Finnegan and Egan, 1981). A third function, normally under *c*I repressor control, may also be required for late transcription (Kalionis *et al.*, 1986b). Determination of the absolute requirements for late transcription may be studied by the establishment of a late transcription assay system. This may be achieved by cloning a representative late promoter so as to express an assayable, selectable gene function and then determining the phage functions necessary for expression of the reporter function.

In the lysogenic state, the late control gene *B* appears to autogenously repress its own transcription (Kalionis *et al.*, 1986b). There is no evidence of activation of late

transcription in the lysogenic state (Finnegan and Egan, 1981), despite the fact that the late control protein is present and active in repressing its own transcription in this state. To investigate the control of *B* gene transcription and the implications of this control for activation of late transcription, the first step was to determine the phage functions required to exert repression at the *B* promoter. Identification of the phage functions necessary to repress *B* gene transcription can be investigated by characterization of *in vivo* transcripts or by the fusion of the *B* promoter to an assayable function. Investigation of whether the repressive autogenous control is responsible for the absence of activation of late transcription in the lysogenic state can be tested by determining whether activation of late transcription occurs under conditions of repression of *B* gene transcription.

The results are presented in three sections. The first section is concerned with defining the region encoding a late promoter. The second section concerns the characterization of potential transcripts, establishment of a late promoter assay system and determination of the phage requirements for activation of late transcription. The third section investigates the nature of the autogenous repression of *B* gene transcription and its implications for the control of late gene transcription.

CHAPTER 2

MATERIALS AND METHODS

CHAPTER 2. MATERIALS AND METHODS.

2.1 BACTERIAL STRAINS.

The bacterial strains used in this study are described in Table 2.1. Bacterial strains constructed in this work were obtained using the methods described in Section 2.15.

2.2. 186 STRAINS.

Derivatives of the phage 186 used in this study are described below.

186 *cIts* : A heat-inducible mutant with a temperature-sensitive repressor (Baldwin *et al.*, 1966; Woods and Egan, 1974).

186 *cI*10 : A clear plaque mutant, which has a defective *cI* gene (Huddleston, 1970).

186 *vir*1 : A 186 mutant able to grow on a 186 lysogen, isolated as a spontaneous mutant in a stock of the phage 186 *cIam*53 (Woods, 1972).

186 *cIts Bam*17 : A derivative of 186 *cIts* with an amber mutation in the late control gene *B* (Hocking and Egan, 1982d; Kalionis *et al.*, 1986b).

2.3 PLASMID VECTORS AND PLASMID CLONES.

2.3.1 Plasmid vectors.

The plasmid vectors used in this study are described below.

pACYC184: A p15A replicon encoding the genes for chloramphenicol and tetracycline resistance (Chang and Cohen, 1978). Plasmid pACYC184 is shown diagramatically in Figure 2.1(a) and contains a number of unique cloning sites in the tetracycline resistance gene (*Hin*dIII, *Eco*RV, *Bam*HI and *Sal*I).

pEC612 : A plasmid for the expression of native proteins (Brumby, 1986). Plasmid pEC612 is derived from the R1 replicon pOU61 (Larsen *et al.*, 1984). The copy number regulation of pEC612 is as for plasmid pOU61 described below. This plasmid directs efficient translation from the initiating methionine of a cloned gene without the use of the endogenous translation signals or the fusion of foreign amino acids onto the amino terminal end of the protein. The expression relies on the unique *Nde*I cloning site on the

TABLE 2.1

BACTERIAL STRAINS.

stock number	strain	genotype	comments	origin or reference
E251	W3350	F⁻ *galK galT strA*748	*sup*$^+$	This lab. Made by transduction of *strA* allele from CGSC4214 into W3350.
E252	W3350 (186)	E251 (186 *cIts*)	A *sup*$^+$ 186 lysogen.	Finnegan (1979).
E296	W3350 (186)	E251 (186 c*Its* *Pam*16)	A *sup*$^+$ *galK*$^-$ 186 lysogen.	This lab.
E508	C600	*tonA supE*44 *thr leu thi*	*supE*	Appleyard (1954).
E536	W3350	*strA*594 F⁻ *galK galT*	*sup*$^+$	Campbell (1965), Weigle (1966).
E538	S26	h-59 (S26 λiλ cured of the λ lysogen) F⁻ *gal*$^+$ *str*^{S}	*sup*$^+$	From D. Hogness. Garen *et al.* (1965).
E540	S26	h-59 (S26 λiλ cured of the λ lysogen) F⁻ *supE gal*$^+$ *str*^{S}	*supE*	From D. Hogness.
E574	C600 (186)	E508 (186 *cIts*)	A *supE* 186 lysogen. Source of 186 DNA.	This lab. Described by Baldwin *et al.* (1966).
E605	JM101	*lac pro supE*44 *thi* F'*traD*36 *proAB lacI*^{q} lacZ ΔM15	Host for M13 infection.	Messing (1979).
E635	W3350 (186)	E251 (186$^+$)	A *sup*$^+$ 186 lysogen.	This lab.
E813	W3350 (P2)	E536 (P2 *C5 nip*)	A *galK*$^-$ P2 *Cts* lysogen.	This lab.
E857	W3350 (Hy5)	E536 (Hy5 *cIts*)	A *galK*$^-$ Hy5 *cIts* lysogen.	This lab.
E863	C600	*galK lac thr leu supE*44	*supE* strain for analysis of promoters.	From K. McKenney. Shulman *et al.* (1976).

TABLE 2.1 (cont.)

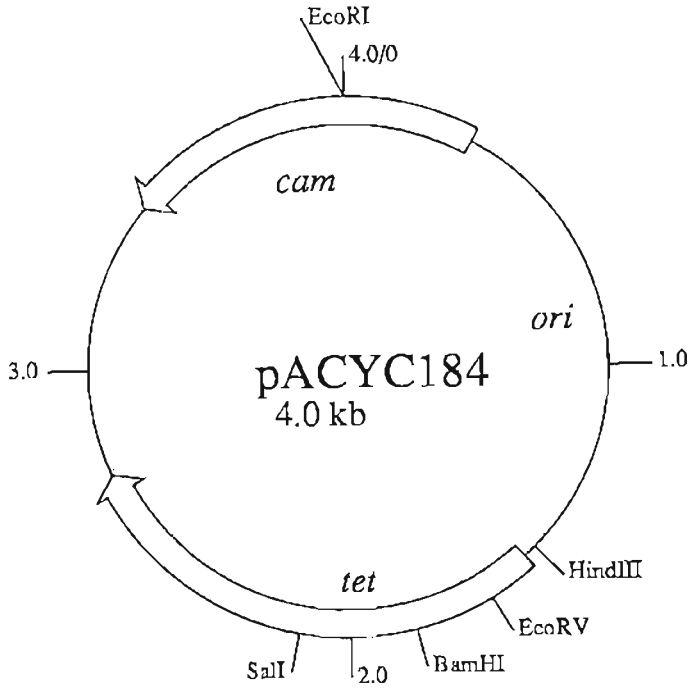
E864	SA1943	*galK lac*$^+$*sup*$^+$	*sup*$^+$ strain for analysis of promoters.	From Smith, Kline and French Labs.
E1017	C600 (186)	E508 (186 *cIts Bam*17)	A *supE* 186 lysogen. Source of 186 *Bam*17 DNA.	This lab. Hocking and Egan (1982d).
E2112	C600 (pACYC184)	E508 pACYC184	Source of pACYC184 DNA.	Chang and Cohen (1978).
E2121	N100 (pKC7)	*recA gal sup*$^+$ carrying plasmid pKC7	Source of pKC7 DNA.	Rao and Rogers (1979).
E2179	C600 (pOU61)	E508 carrying pOU61	Source of pOU61 DNA.	Larsen *et al.* (1984).
E2184	N100 (pKO2)	*recA gal sup*$^+$ carrying plasmid pKO2	Source of pKO2 DNA.	From K. McKenney, de Boer (1984).
E2195	N100 (pKL600)	carrying pKL600	Source of plasmid pKL600.	From K. McKenney.
E2301	P90C (pMC7)	F$^-$ *ara B*1 *lac pro str*^{r} carrying pMC7	Source of pMC7 DNA.	Calos (1978).
E2389	C600 (pEC612)	E508 carrying pEC612	Source of pEC612 DNA.	This lab. Brumby (1986).
E4106	JM105	*thi rpsl endA sbcB*15 *hspR*4 *lac proAB* F' *traD*36 *proAB lacI*^{q} *lacZ* ΔM15	*sup*$^+$ F'*lacI*^{q}	Yanisch-Perron *et al.* (1985).
E4170	W3350 (186)	E251 (186 cI*ts ins*3)	A *sup*$^+$ *galK*$^-$ 186 *int*$^-$ lysogen.	Richardson (1987).
E4308	SA1943 (186)	E864 (186 cI*ts*)	A *sup*$^+$ *galK*$^-$ 186 lysogen.	This work.
E4309	SA1943 (186)	E864 (186 cI*ts Bam*17)	A *sup*$^+$ *galK*$^-$ 186 lysogen.	This work.

Figure 2.1 Circular maps of the cloning vector pACYC184 and the promoter analysis plasmid pKO2.

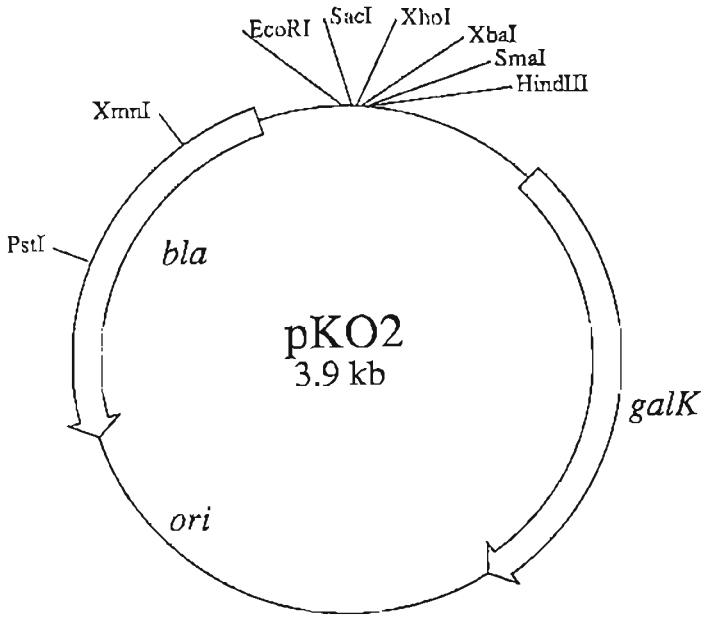
a. Cloning vector pACYC184. The position and direction of transcription of the genes encoding resistance to chloramphenicol (*cam*) and tetracycline (*tet*) are indicated, as are the origin of replication (*ori*) and the positions of relevant restriction sites (Chang and Cohen, 1978).

b. Promoter analysis plasmid pKO2. The position and direction of transcription of the gene encoding resistance to ampicillin (*bla*) and the *galK* gene are indicated. The origin of replication (*ori*) and the positions of relevant restriction sites are indicated (McKenney *et al.*, 1981).

(a)


EcoRI
4.0/0
cam
ori
pACYC184
4.0 kb
3.0
1.0
tet
HindIII
EcoRV
SalI
2.0
BamHI

(b)


EcoRI
SacI
XhoI
XbaI
SmaI
HindIII
XmnI
PstI
bla
pKO2
3.9 kb
galK
ori

vector (5'-CATATG-3'), being preceded by a sequence designed to maximize translational efficiency from the ATG codon of the *Nde*I cloning site. Transcription of cloned gene products is from the strong, hybrid *trc* promoter, the activity of which is regulated by *lacI* repressor (Brosius *et al.*, 1985). Strong transcriptional terminators from the *rrnB* ribosomal RNA operon (Brosius *et al.*, 1981) prevent transcription into the plasmid. Expression of a cloned gene product is accomplished by construction of a *Nde*I site at the methionine codon of the cloned gene product to be expressed (by *in vitro* site-directed mutagenesis). Expression of the gene product is achieved by transfer to high temperature and the addition of the gratuitous inducer IPTG (in a host encoding *lac* repressor). Plasmid pEC612 encodes the gene for ampicillin resistance and is shown diagrammatically in Figure 2.2(b).

pKC7 : A derivative of the plasmid pBR322, in which the *Hin*dIII-*Bam*HI fragment is replaced with the *Hin*dIII-*Bam*HI fragment from the transposon Tn5 (Rao and Rogers, 1979; Auerswald *et al.*, 1981; Mazodier *et al.*, 1985).

Plasmid pKC7 encodes the genes for ampicillin and kanamycin resistance.

pKL600 : A plasmid containing the p_{lac} promoter in an orientation such that it allows the expression of the *galK* gene. This plasmid was obtained from K. McKenney (personal communication) and encodes the gene for ampicillin resistance. Plasmid pKL600 is a derivative of the *galK* promoter analysis plasmid pKO1 (McKenney *et al.*, 1981), which has fewer cloning sites in its polylinker than the derivative pKO2 (de Boer, 1984).

pKO2 : A plasmid for the cloning and quantitation of promoter activity (de Boer, 1984). pKO2 is shown diagrammatically in Figure 2.1(b). This plasmid encodes the gene for ampicillin resistance.

pMC7 : A multicopy $lacI^q$ clone derived from plasmid pMB9 (Calos, 1978). Plasmid pMC7 encodes the gene for tetracycline resistance.

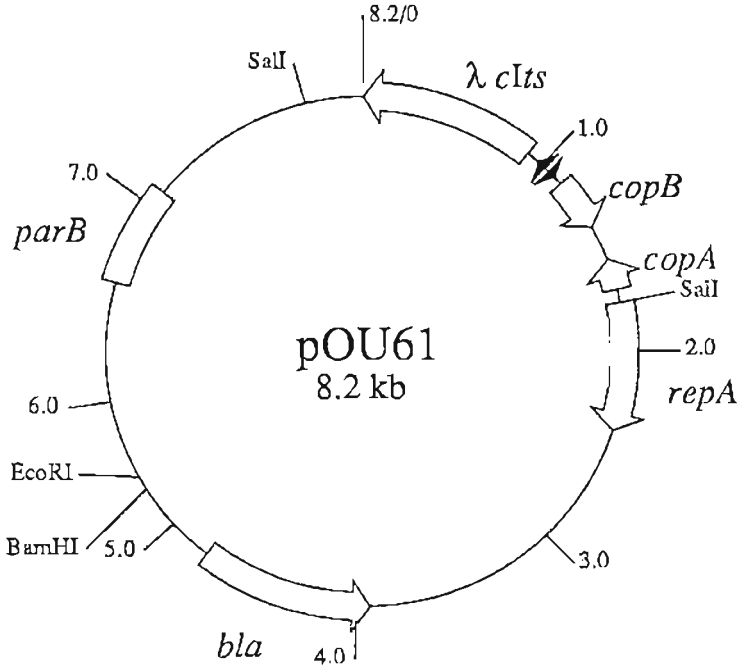
pOU61 : A R1-derived replicon (Larsen *et al.*, 1984) whose copy number is regulated by temperature through the use of the $\lambda\, p_R$ promoter and repression by the λ *cIts* gene product also contained on the vector. The copy number of the plasmid is maintained at the chromosomal equivalent of one per cell at 30°C and plasmid stability maintained by

Figure 2.2 Circular maps of plasmids pOU61 and pEC612.

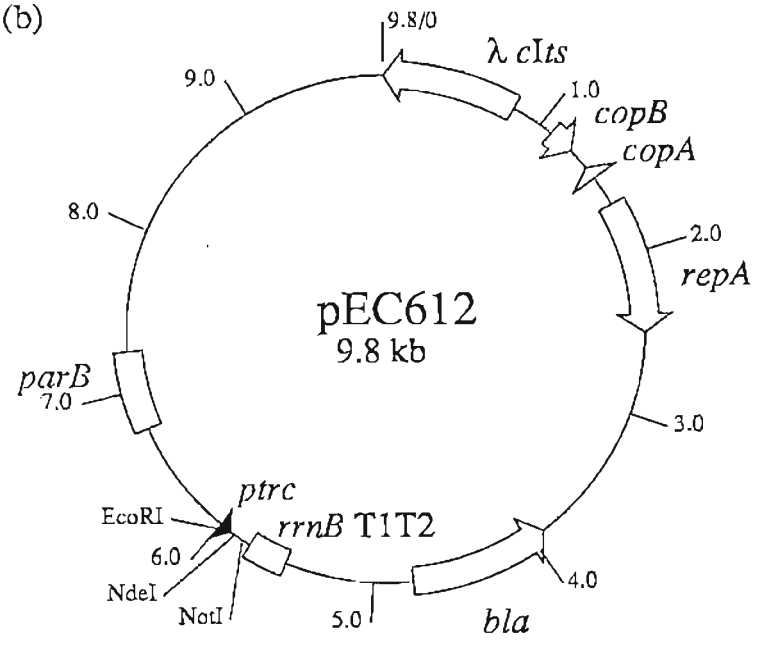
a. Cloning vector pOU61. The function of the genes are described in Larsen *et al.* (1984). The position and direction of transcription of the genes on the plasmid is shown. Initiation of transcription from the origin of replication (just distal to *repA*) is dependent on the expression of the *repA* gene. The *copA* transcipt inhibits translation of *repA*. The *copB* gene product normally represses transcription of *repA*. Plasmid stability is maintained through the partitioning loces *parB* (Rasmussen *et al.*, 1987). Expression of the lambda *cIts* gene is from the p_{RM} promoter and the divergent promoter p_R transcribes *repA*. The promoters are indicated by shaded arrows. The *bla* gene encodes resistance to ampicillin. The positions of relevant restriction sites are also shown.

b. Expression plasmid pEC612 (Brumby, 1986). Functions of the genes is as described in (a). The *trc* promoter (shaded arrow; Brosius *et al.*, 1985) and the *rrnB* terminators T1 and T2 (Brosius *et al.*, 1981) are also shown. Relevant restiction sites are indicated.

(a)


8.2/0
SalI
λ cIts
1.0
copB
copA
SalI
2.0
repA
3.0
4.0
bla
5.0
BamHI
EcoRI
6.0
parB
7.0
pOU61
8.2 kb

(b)


9.8/0
λ cIts
1.0
copB
copA
2.0
repA
3.0
4.0
bla
5.0
rrnB T1T2
ptrc
NotI
NdeI
6.0
EcoRI
parB
7.0
8.0
9.0
pEC612
9.8 kb

the presence of the partitioning locus (Rasmussen *et al.*, 1987). Inactivation of the *cIts* repressor at 42°C results in run-away replication, eventually resulting in cell death (Larsen *et al.*, 1984). Plasmid pOU61 encodes the ampicillin resistance gene and is shown diagrammatically in Figure 2.2(a).

2.3.2 Plasmid clones

Plasmid clones used in this study were constructed in this work and are described below. Plasmid clones constructed in this work were obtained using the methods described in Section 2.28.

pEC428 : Contains the 490 bp *Alu*I-*Alu*I (11.0% to 12.6%) fragment from 186 *cIts* DNA (source of DNA mEC404; Section 2.4.2), cloned into the *Sma*I site of pKO2 in the orientation that the *galK* gene is under control of late promoter p_V (p_V*galK* plasmid).

pEC429 : Contains the 2.2 kb *Sac*I-*Sac*I (63.0% to 70.2%) fragment from 186 *cIts* DNA, carrying the 186 late genes *FED* and the *B* gene under its own control, cloned into the *Sac*I site of plasmid pKO2 in an orientation such that *B* gene transcription is directed towards the *bla* gene of the plasmid.

pEC430 : Contains the 2.2 kb *Sac*I-*Sac*I (63.0% to 70.2%) fragment from 186 *cI*ts DNA, carrying the 186 late genes *FED* and the *B* gene under its own control, cloned into the *Sac*I site of the p_V*galK* plasmid pEC428 in an orientation such that *B* gene transcription is directed towards the *bla* gene of the plasmid.

pEC431 : Contains the 2.2 kb *Sac*I-*Sac*I (63.0% to 70.2%) fragment from 186 *cI*ts *Bam*17 DNA, carrying the 186 late genes *FED* and the *Bam*17 gene under its own control, cloned into the *Sac*I site of the p_V*galK* plasmid pEC428 in an orientation such that *B* gene transcription is directed towards the *bla* gene of the plasmid.

pEC432 : A kanamycin resistant, ampicillin sensitive derivative of plasmid pKO2. Constructed by cloning the 1.3 kb *Eco*RI to *Sma*I restriction fragment of plasmid pKC7 into the large *Eco*RI to *Xmn*I fragment of pKO2.

pEC433 : A kanamycin resistant, ampicillin sensitive derivative of the p_V*galK* plasmid pEC428. Constructed by cloning the 1.3 kb *Eco*RI to *Sma*I restriction fragment of plasmid pKC7 into the large *Eco*RI to *Xmn*I fragment of the p_V*galK* plasmid pEC428.

pEC434 : A plasmid for expression of the *B* gene product. A 357 bp *Nde*I-*Sma*I (66.5% to 67.6%) restriction fragment from M13 clone mEC650 (Section 2.4.2), encoding the B^+ gene coding sequence, cloned into the *Nde*I site and the *Not*I (end-filled; Section 2.28.2) sites of pEC612, to yield the *B* expression plasmid $pp_{trC}B^+$. The *Sma*I site in the M13mp19 polylinker was utilized for cloning.

pEC435 : A plasmid for expression of the *Bam*17 gene product. A 357 bp *Nde*I-*Sma*I (66.5% to 67.6%) restriction fragment, encoding the *Bam*17 coding sequence, from M13 clone mEC650 (Section 2.4.2) cloned into the *Nde*I site and the *Not*I (end-filled; Section 2.28.2) sites of pEC612, to yield the *Bam* expression plasmid $pp_{trC}Bam$. The *Sma*I site in the M13mp19 polylinker was utilized for cloning.

pEC436 : A plasmid carrying a fragment encoding the presumptive *B* promoter-operator region (Kalionis *et al.*, 1986b). A 290 bp *Pst*I-*Nde*I (65.5% to 66.5%) restriction fragment, blunted by treatment with the Klenow fragment of DNA polymerase I (Section 2.28.2), from mEC650 (Section 2.4.2) was cloned into the *Xba*I (end-filled; Section 2.28.2) site of pKO2. The orientation was such that transcription from the *B* promoter was towards the *bla* gene.

pEC437 : A p_V*galK* plasmid carrying a fragment encoding the presumptive *B* promoter-operator region (Kalionis *et al.*, 1986b). A 290 bp *Pst*I-*Nde*I (65.5% to 66.5%) restriction fragment, blunted by treatment with the Klenow fragment of DNA polymerase I (Section 2.28.2), from mEC650 (Section 2.4.2) was cloned into the *Xba*I (end-filled) site of the p_V*galK* pEC428. The orientation was such that transcription from the *B* promoter was towards the *bla* gene.

pEC438 : Contains the 490 bp *Alu*I-*Alu*I (11.0% to 12.6%) fragment from 186 *clts* DNA (source of DNA mEC404; Section 2.4.2), cloned into the *Sma*I site of pKO2 in the orientation that the *galK* gene is under control of late promoter p_{L2}.

pEC439 : Contains the 1.4 kb *Bam*HI-*Eco*RI (10.0% to 13.3%) fragment from wild-type 186 DNA, cloned into the *Sma*I site of pKO2 (after end-filling of the terminii of the fragment; Section 2.28.2) in the orientation that the *galK* gene is under control of late promoter p_{L2}.

pEC440 : A deletion derivative of the p_V*galK* plasmid pEC428, with DNA downstream of the +18 position of the p_V promoter deleted. Constructed by end-filling the

large *Oxa*NI-*Hin*dIII fragment of pEC428 (Section 2.28.2) and subsequent ligation. The *Oxa*NI site (13.1%) is downstream of the p_V promoter and the *Hin*dIII site is in the polylinker of pEC428.

pEC441 : Contains the 541 bp *Eco*RI-*Hin*dIII fragment from M13 clone mEC405 (*Eco*RI and *Hin*dIII are in the polylinker; Section 2.4.2) cloned into the *Eco*RI and *Hin*dIII sites of pKO2. The orientation was such that the *galK* gene is under control of late promoter p_V.

pEC442 : A *B box* deletion derivative of the p_V*galK* plasmid pEC441, with the first *B box* downstream of the p_V promoter deleted (deletion coordinates 1018-1029; Figure 3.4(a)). Constructed by cloning the 541 bp *Eco*RI-*Hin*dIII fragment from M13 clone mEC408 (*Eco*RI and *Hin*dIII are in the polylinker; Section 2.4.2) into the *Eco*RI and *Hin*dIII sites of pKO2. The orientation was such that the *galK* gene is under control of late promoter p_V.

pEC443 : A *B box* deletion derivative of the p_V*galK* plasmid pEC441, with the second *B box* downstream of the p_V promoter deleted (deletion coordinates 1108-1119; Figure 3.4(a)). Constructed by cloning the 541 bp *Eco*RI-*Hin*dIII fragment from M13 clone mEC409 (*Eco*RI and *Hin*dIII are in the polylinker; Section 2.4.2) into the *Eco*RI and *Hin*dIII sites of pKO2. The orientation was such that the *galK* gene is under control of late promoter p_V.

pEC444 : A *B box* deletion derivative of the p_V*galK* plasmid pEC441, with both of the *B boxes* downstream of the p_V promoter deleted (deletion coordinates 1018-1029 and 1108-1119; Figure 3.4(a)). Constructed by cloning the 541 bp *Eco*RI-*Hin*dIII fragment from M13 clone mEC409 (*Eco*RI and *Hin*dIII are in the polylinker; Section 2.4.2) into the *Eco*RI and *Hin*dIII sites of pKO2. The orientation was such that the *galK* gene is under control of late promoter p_V.

pEC445 : A plasmid in which the *galK* gene is under the transcriptional control of the *B* promoter (Kalionis *et al.*, 1986b). A 290 bp *Pst*I-*Nde*I (65.5% to 66.5%) restriction fragment, blunted by treatment with the Klenow fragment of DNA polymerase I (Section 2.28.2), from mEC650 (Section 2.4.2) was cloned into the *Xba*I (end-filled) site of pKO2. This restriction fragment encodes the *B* promoter region (Kalionis *et al.*, 1986b). The orientation was such that transcription from the *B* promoter was towards the *galK* gene.

pEC446 : A kanamycin resistant derivative of plasmid pEC445, in which the *galK* gene is under the transcriptional control of the *B* promoter (Kalionis *et al.*, 1986b). Constructed by cloning the 1.3 kb *Eco*RI to *Sma*I restriction fragment of plasmid pKC7 into the large *Eco*RI to *Xmn*I fragment of pEC445.

pEC447 : Contains the 1.4 kb *Bam*HI-*Bam*HI (63.0% to 67.6%) fragment from M13 clone mEC412 (Section 2.4.2) carrying the 186 late genes *FED* and the *B* gene under its own control, cloned into the *Bam*HI site of the plasmid pEC612 in an orientation such that *B* gene transcription is directed towards the *bla* gene of plasmid pEC612. The *Bam*HI sites of the M13 clone mEC412 were both derived from the polylinker, a result of the construction of the M13 clone used to clone the DNA.

pEC448 : Contains the 1.4 kb *Bam*HI-*Bam*HI (63.0% to 67.6%) fragment from M13 clone mEC413 (Section 2.4.2) carrying the 186 late genes *FED* and the *Bam*17 gene under its own control, cloned into the *Bam*HI site of the plasmid pEC612 in an orientation such that *B* gene transcription is directed towards the *bla* gene of plasmid pEC612. The *Bam*HI sites of the M13 clone mEC412 were both derived from the polylinker, a result of the construction of the M13 clone used to clone the DNA.

pEC449 : Contains the 3.6 kb *Xho*I-*Bgl*II (67.6% to 79.6%) fragment from wild-type DNA cloned into the large *Sal*I to *Bam*HI fragment of pACYC184. The 3.6 kb *Xho*I-*Bgl*II (67.6% to 79.6%) fragment encodes genes *CP69*, *int*, *c*I, *apl*, *c*II, *fil* and *dhr* (Kalionis *et al.*, 1986b; Richardson *et al.*, 1988). The orientation of the plasmid was such that genes *c*I, *int* and *CP69* were under the control of the *tet* promoter of the plasmid.

pEC450 : Contains the 2.5 kb *Ssp*I-*Sau*3A (69.6% to 77.9%) fragment from pEC449 DNA cloned into the large *Eco*RV to *Bam*HI fragment of pACYC184. The 2.5 kb *Ssp*I-*Sau*3A (69.6% to 77.9%) fragment encodes genes *c*I, *apl* and *c*II (Kalionis *et al.*, 1986b). The orientation of the plasmid was such that genes *apl* and *c*II were under the control of the *tet* promoter of the plasmid.

pEC451 : Contains the 1.5 kb *Sal*I-*Sau*3A (73.0% to 77.9%) fragment from pEC449 DNA cloned into the large *Xho*I to *Bam*HI fragment of pACYC184. The 1.5 kb *Sal*I-*Sau*3A (73.0% to 77.9%) fragment encodes genes *apl* and *c*II (Kalionis *et al.*, 1986b). The orientation of the plasmid was such that genes *apl* and *c*II were not under the control of the *tet* promoter of the plasmid.

pEC452 : Contains the 1.3 kb *Hae*III-*Hae*III (71.2% to 75.5%) fragment from pEC449 DNA cloned into the *Eco*RV site of pACYC184. The 1.3 kb *Hae*III-*Hae*III (71.2% to 75.5%) fragment encodes the *c*I gene (Kalionis *et al.*, 1986b). The orientation of the plasmid was such that the *c*I gene was not under the control of the *tet* promoter of the plasmid.

pEC453 : Contains the 1.3 kb *Hae*III-*Hae*III (71.2% to 75.5%) fragment from pEC449 DNA cloned into the *Eco*RV site of pACYC184. The 1.3 kb *Hae*III-*Hae*III (71.2% to 75.5%) fragment encodes the *c*I gene (Kalionis *et al.*, 1986b). The orientation of the plasmid was such that the *c*I gene was under the control of the *tet* promoter of the plasmid.

2.4 M13 VECTORS AND M13 CLONES.

2.4.1 M13 vectors.

The M13 vector used in this study was M13mp18 (Yanisch-Perron *et al.*, 1985), obtained from BRESATEC, South Australia.

2.4.2 M13 clones.

The M13 clones used in this study are described below. M13 clones constructed in this work were obtained using the methods described in Section 2.28. The sequence coordinates of the restriction sites refer to the first base of the site on the appropriate strand, as described in Section 3.2.1.

mEC403 : M13mp8 containing the *r*-strand of the 1.4 kb *Bam*HI-*Eco*RI fragment (10.0% to 13.3%; sequence coordinates 6-1361) from 186 *cIts* DNA cloned into the *Bam*HI and *Eco*RI sites of M13mp8 (Dibbens, 1984).

mEC404 : M13mp9 containing the *l*-strand of the 1.4 kb *Bam*HI-*Eco*RI fragment (10.0% to 13.3%; sequence coordinates 6-1361) from 186 *cIts* DNA cloned into the *Bam*HI and *Eco*RI sites of M13mp9 (Dibbens, 1984).

mEC405 : M13mp18 containing the *r*-strand of the 490 bp *Alu*I-*Alu*I fragment (12.2% to 13.8%; sequence coordinates 648-1137) from mEC403 RF DNA (186 *cIts* DNA) cloned into the *Sma*I site of M13mp18 (this work).

mEC406 : M13mp18 containing the *r*-strand of the 968 bp *Xmn*I-*Eco*RI fragment (11.3% to 13.3%; sequence coordinates 394-1361) from mEC403 RF DNA (186 *cIts* DNA), end-filled (Section 2.28.2) and cloned into the *Sma*I site of M13mp18 (this work).

mEC407 : M13mp18 containing the *r*-strand of the 388 bp *Bam*HI-*Xmn*I fragment (10.0% to 11.3%; sequence coordinates 6-393) from mEC403 RF DNA (186 *cIts* DNA), end-filled (Section 2.28.2) and cloned into the *Sma*I site of M13mp18 (this work).

mEC408 : M13mp18 containing the *r*-strand of the 490 bp *Alu*I-*Alu*I fragment (12.2% to 13.8%; sequence coordinates 648-1137) with the first *B box* downstream of the p_V promoter deleted (deletion coordinates 1018-1029; Figure 3.4(a)). Constructed by oligonucleotide mutagenesis (Section 2.31) of M13 clone mEC405 with oligonucleotide DS 620 (Section 2.5) (this work).

mEC409 : M13mp18 containing the *r*-strand of the 490 bp *Alu*I-*Alu*I fragment (12.2% to 13.8%; sequence coordinates 648-1137) with the second *B box* downstream of the p_V promoter deleted (deletion coordinates 1108-1119; Figure 3.4(a)). Constructed by oligonucleotide mutagenesis (Section 2.31) of M13 clone mEC405 with oligonucleotide DS 619 (Section 2.5) (this work).

mEC410 : M13mp18 containing the *r*-strand of the 490 bp *Alu*I-*Alu*I fragment (12.2% to 13.8%; sequence coordinates 648-1137) with both *B boxes* downstream of the p_V promoter deleted (deletion coordinates 1018-1029 and 1108-1119; Figure 3.4(a)). Constructed by oligonucleotide mutagenesis (Section 2.31) of M13 clone mEC408 with oligonucleotide DS 619 (Section 2.5) (this work).

mEC411 : M13mp18 containing the *r*-strand of the 600 bp *Pst*I-*Eco*RI fragment (12.5% to 13.3%; sequence coordinates 762-1361) from mEC403 RF DNA (186 *cIts* DNA), end-filled (Section 2.28.2) and cloned into the *Sma*I site of M13mp18 (this work).

mEC412 : M13mp18 containing the *r*-strand of the 1.4 kb *Sac*I-*Sac*I (63.0% to 67.6%; the second *Sac*I site was in the M13 polylinker) fragment from M13 clone mEC648 RF DNA, cloned into the *Sac*I site (this work).

mEC413 : M13mp18 containing the *r*-strand of the 1.4 kb *Sac*I-*Sac*I (63.0% to 67.6%; the second *Sac*I site was in the M13 polylinker) fragment from M13 clone mEC649 RF DNA, cloned into the *Sac*I site (this work).

mEC648 : M13mp19 containing the *l*-strand of the 1.9 kb *Hin*dIII-*Xho*I (61.3% to 67.6%) fragment from 186 *cIts* DNA cloned into the *Hin*dIII and *Sal*I sites of M13mp19 (Osborne, 1986).

mEC649 : M13mp19 containing the *l*-strand of the 1.9 kb *Hin*dIII-*Xho*I (61.3% to 67.6%) fragment from 186 *cIts Bam*17 DNA cloned into the *Hin*dIII and *Sal*I sites of M13mp19 (Osborne, 1986).

mEC650 : M13mp19 containing the *l*-strand of the 1.9 kb *Hin*dIII-*Xho*I (61.3% to 67.6%) fragment from 186 *cIts* DNA cloned into the *Hin*dIII and *Sal*I sites of M13mp19 (Osborne, 1986), with a *Nde*I site created at the predicted initiation codon of the *B* gene (Kalionis *et al.*, 1986b) by oligonucleotide mutagenesis (Section 2.31) of mEC648 (Osborne, 1986).

mEC651 : M13mp19 containing the *l*-strand of the 1.9 kb *Hin*dIII-*Xho*I (61.3% to 67.6%) fragment from 186 *cIts Bam*17 DNA cloned into the *Hin*dIII and *Sal*I sites of M13mp19 (Osborne, 1986), with a *Nde*I site created at the predicted initiation codon of the *Bam*17 gene (Kalionis *et al.*, 1986b) by oligonucleotide mutagenesis (Section 2.31) of mEC649 (Osborne, 1986).

2.5 OLIGONUCLEOTIDES.

The oligonucleotides used during the course of this work for specific mutagenesis and DNA sequencing were constructed by BRESATEC and are described below.

Universal sequencing primer (USP) 17-mer:

5'-GTAAAACGACGGCCAGT-3'.

Oligonucleotide DS 619 30-mer:

5'-CTGTTGCCCGATGGCTACGGCGATGTGGTC-3'.

A *l*-strand oligonucleotide (sequence coordinates 1093-1134) used to create a 12 bp deletion (sequence coordinates 1108-1119) of the second *B box* in the *V* gene (Section 4.2.4(b)).

Oligonucleotide DS 620 30-mer:

5'-ATTATCAGCGCCAGTATGGCCGAAACCTAT-3'.

A *l*-strand oligonucleotide (sequence coordinates 1003-1029) used to create a 12 bp deletion (sequence coordinates 1018-1029) in the first *B box* in the *V* gene (Section 4.2.4(b)).

Oligonucleotide DS 621 17-mer:

5'-CGAATCGGCGTCGAGGG-3'.

A *l*-strand oligonucleotide (sequence coordinates 967-983) used as a sequencing primer to determine the sequence of site-directed mutants specified by oligonucleotide DS 620.

2.6 ENZYMES.

Restriction Endonucleases :New England Biolabs, Toyobo or Bethesda Research Labs.

E. coli DNA polymerase I (Klenow fragment) : BRESATEC.

T4 DNA ligase : Boehringer Mannheim and BRESATEC.

Calf intestinal phosphatase : Boehringer Mannheim.

Avian myeloblastosis virus (AMV) reverse transcriptase : Life Science Inc., Florida.

Lysozyme : Sigma Chemical Co.

Proteinase K : Boehringer Mannheim.

RNase A : Sigma Chemical Co. Stock solutions were heated at 80^{o}C for 20 min to inactivate DNases.

E. coli DNase I : Boehringer Mannheim.

T4 Polynucleotide Kinase : Bochringer Mannheim.

All enzymes were stored according to the manufacturers' directions.

2.7 RADIOCHEMICALS.

Radiochemicals d[α-^{32}P]CTP and d[α-^{32}P]ATP of specific activity 3000 Ci/mmol, d[γ-^{32}P]ATP of specific activity 4000 Ci/mmol (radioactive concentrations of 5 mCi/ml) were purchased from BRESATEC. D-[1-^{14}C]-galactose of specific activity of 58 mCi/mmol (radioactive concentration of 200 uCi/ml) was obtained from Amersham.

2.8 CHEMICALS.

All chemicals were of analytical grade or of the highest purity available.

Acridine orange : Sigma Chemical Co.

Acrylamide : Sigma Chemical Co.

Agarose : Sigma Chemical Co.

Amine A : Humpko Sheffield, U.S.A.

Ampicillin (sodium salt) : Sigma Chemical Co. Stock solutions (50 mg/ml in water), millipore filtered and stored at -20°C.

Ammonium persulphate (APS) : May and Baker. Stock solutions at 25% (w/v) in water, prepared freshly for use.

Bacto-tryptone, yeast extract and Bacto-agar : Difco Labs., U.S.A.

Bovine serum albumin (BSA) : Sigma Chemical Co. Acetylated before use to remove nucleases according to the procedure of Gonzalez *et al.* (1977) and kept as a 1.5 mg/ml solution in water at -20°C (gift from M. Pritchard, this laboratory).

5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (BCIG) : Sigma Chemical Co. Stock solutions at 20 mg/ml in dimethyl formamide, were kept at -20°C.

Brij 58 (polyoxyethylene 20 cetyl ether) : Sigma Chemical Co.

Calf thymus DNA : Sigma Chemical Co. Sonicated, heat denatured and stored as a 10 mg/ml solution in 10 mM Tris-HCl, pH 8.0, 0.1 mM EDTA at -20°C.

Casein amino acids (vitamin free) : Difco labs, U.S.A..

Chloramphenicol : Sigma Chemical Co. Stock solutions (30 mg/ml in 95% ethanol), stored at -20°C.

CsCl : Bethesda Research Labs.

Deoxyribonucleoside triphosphates (dNTP) : Sigma Chemical Co. Stock solutions at 20 mM (prepared in 5 mM Tris-HCl, pH 8.0, 0.1 mM EDTA) were kept at -20°C.

Dithiothreitol : Sigma Chemical Co. Stored in the dark as a 1 M solution in water, millipore filtered and kept at -20°C.

E. coli tRNA : Sigma Chemical Co. Extracted three times with phenol/TE before use and kept as a 10 mg/ml solution in TE at -20°C.

Ethanol (95%) : Redistilled before use and stored at -20°C. RNase-free ethanol was obtained by sterilising the distillation apparatus and collection bottles in a 110°C oven overnight before use.

Ethidium bromide : Sigma Chemical Co. Stored as a 10 mg/ml solution in water at 4°C in the dark.

Ethylenediaminetetraacetic acid (EDTA) : Disodium salt. Sigma Chemical Co.

Ficoll 400 : Phamacia Fine Chem.

Formamide : B.D.H. Labs. Aust. De-ionised (Section 2.29.3d) and stored in the dark at -20°C.

Glyoxal : Technical grade glyoxal was obtained from B.D.H. Labs. Aust. De-ionised (Section 2.29.3d) and stored in the dark at -80°C.

8-hydroxy quinoline : Sigma Chemical Co.

Isopropanol : May and Baker Ltd.

Isopropyl-β-D-thiogalactopyranoside (IPTG) : Sigma Chemical Co. Stock solutions were at 24 mg/ml in water and stored at -20°C.

Kanamycin (sulphate) : Sigma Chemical Co. Stock solutions (50 mg/ml in water), millipore filtered and stored at-20°C.

Low-gelling-temperature agarose (LGT) : Bethesda Research Labs.

MacConkey agar: Difco Labs., U.S.A.

Mixed bed resin (508-X8(D)) : Bio-Rad Labs.

β -Mercaptoethanol : Sigma Chemical Co.

N, N'-methylene-bis-acrylamide (bis) : Sigma Chemical Co.

N, N, N', N'-tetramethylethylenediamine (TEMED) : Eastern Kodak Co.

Nonidet P40 : B.D.H. Labs., Aust.

Phenol : AR grade, B.D.H. Labs Aust. Redistilled and stored in the dark at -20°C.

Polyethylene glycol (PEG) 8000 : for phage preparations and general use was from Sigma Chemical Co. M13 phage preparations for sequencing were prepared using PEG 8000 from B.D.H. Labs. Aust.

Polyvinyl pyrrolidone : May and Baker Ltd.

Ribonucleoside triphosphates (NTP) : Sigma Chemical Co. Stock solutions at 20 mM (prepared in 5 mM Tris-HCl, pH 8.0, 0.1 mM EDTA) were kept at -20°C.

Sodium azide : Ajax Chemicals Ltd.

Sodium dodecyl sulphate (SDS): Sigma Chemical Co.

Tetracycline : Upjohn Pty. Ltd. Stock solutions (25 mg/ml in 75% ethanol), stored at -20°C.

Tetramethylammonium chloride : Aldrich Chemical Co.

Toluene Scintillation Fluid contained 0.35% (w/v) PPO (2,5-diphenyl oxazole) and 0.035% (w/v) POPOP (1,4-bis [2-(5-phenyl-oxazolylphenyl)] benzene in toluene. Both PPO and POPOP were obtained from Sigma Chemical Co. Toluene was obtained from May and Baker Ltd.

Thiamine : Sigma Chemical Co.

Thymine : Sigma Chemical Co.

Trizma base and Tris 7-9 : Sigma Chemical Co.

Urea : Sigma Chemical Co.

Zinc Chloride : May and Baker Ltd.

2.9 MEDIA.

2.9.1 Liquid media.

All media and buffers were prepared in glass distilled water and were sterilised by autoclaving for 25 min at 120°C and 120 kPa.

L (Luria) broth : 1% Bacto-tryptone, 0.5% yeast extract, 1% NaCl, pH 7.0.

T broth: 1% Bacto-tryptone, 0.5% NaCl, pH 7.0.

2xYT broth : 1.6% Bacto-typtone, 1% yeast extract, 0.5% NaCl, pH 7.0.

M13 minimal medium : 1.05% K_2HPO_4, 0.45% KH_2PO_4, 0.1% $(NH_4)_2SO_4$, 0.05% $Na_3citrate.2H_2O$ and water to 1000 ml. This solution was autoclaved, cooled to 45°C and the following added from separately prepared sterile solutions; 10 ml of 20% glucose, 0.8 ml of 1 M $MgSO_4$, 0.5 ml of 1% thiamine-HCl.

M63 : 0.1 M KH_2PO_4, 0.015 M $(NH_4)_2SO_4$, 0.002 mM $FeSO_4.7H_2O$ in water adjusted to pH 7.0 with KCl.

Additions to growth media were 1.0% vitamin-free casamino acids, 0.2% glucose, 5 ug/ml thiamine-HCl, 2-5 ug/ml thymine, and 50 ug/ml of individual amino acids.

Antibiotics were added to rich media at the following concentrations; ampicillin at 50 ug/ml, chloramphenicol at 30 ug/ml, kanamycin at 50 ug/ml and tetracycline at 25 ug/ml. For minimal media half the above listed concentrations were used.

2.9.2 Solid media.

L plates : 1.5% Bacto-agar was added to L broth.

TB plates : 1.5% Bacto-agar was added to T broth.

M13 minimal plates : 1.5% Bacto-agar was added to M13 minimal media.

MacConkey-galactose plates : 4% MacConkey agar base, autoclaved, then 1% galactose was added and the plates poured.

Antibiotic plates : L, MacConkey and TB plates were supplemented with antibiotics at the following concentrations; ampicillin at 50 ug/ml, chloramphenicol at 30 ug/ml, kanamycin at 50 ug/ml and tetracycline at 25 ug/ml. M13 minimal plates were supplemented with antibiotic at half the above listed concentrations. Plates were poured from 30 ml of the appropriate medium, dried overnight at 37^{o}C and stored at 4^{o}C.

Soft agar overlay : 1% Bacto-tryptone, 0.7% Bacto-agar, 0.5% NaCl, pH 7.0.

YT soft agar overlay : 0.8% Bacto-tryptone, 0.5% yeast extract, 0.5% NaCl, 0.7% Bacto-agar, pH 7.0.

2.10 BUFFERS.

20x SSC : 3.0 M NaCl, 0.3 M Na_3citrate, pH 7.4.

10x TBE : 0.89 M Tris-HCl, 0.89 M boric acid, 2.7 mM EDTA, pH 8.3.

10x TAE : 0.4 M Tris-acetate, 0.2 M Na acetate, 10 mM EDTA, pH 8.2.

10x TE : 0.1 M Tris-HCl, pH 8.0, 1 mM EDTA.

10x TM : 0.1 M Tris-HCl, pH 8.0, 0.1 M $MgCl_2$.

TM used for preparation and storage of phage (Section 2.13, 2.16) was 10 mM Tris-HCl, pH 7.1, 10 mM $MgSO_4$.

100x Denhardts solution : 2% Ficoll 400, 2% BSA, 2% Polyvinyl pyrrolidone.

10x Glycerol loading buffer : 50% glycerol, 0.40% Bromophenol Blue, 0.20% Xylene Cyanol, 10 mM EDTA.

Formamide loading buffer : 95% Formamide, 0.1% Bromophenol Blue, 0.1% Xylene Cyanol, 10 mM EDTA.

2.11 DNA MOLECULAR WEIGHT MARKERS.

*Hpa*II digest of pUC19 DNA at 500 ng/ul : Obtained from BRESATEC. Fragment sizes in bp: 501, 489, 404, 331, 242, 190, 147, 111, 110, 67, 34 (x2) and 26.

*Hin*dIII digest of phage λ DNA at 200 ng/ul : Obtained from BRESATEC. Fragment sizes in bp: 23130, 9416, 6557, 4361, 2322, 2027, 564 and 125.

*Hin*dIII/*Eco*RI digest of phage λ DNA combined with *Hin*dIII digest of λ DNA, at 200 ng/ul : Obtained from BRESATEC. Fragment sizes in bp: 23130, 21226, 9416, 6557, 5148, 4973, 4361, 4277, 3530, 2322, 2027, 1904, 1584, 1330, 983, 831, 564 and 125.

*Eco*RI digest of phage SPP-1 DNA at 500 ng/ul : Obtained from BRESATEC. Fragment sizes in bp: 7840, 6960, 5860, 4690, 3370, 2680, 1890, 1800, 1450, 1330, 1090,880, 660, 480 and 380.

Sizes were as given by the manufacturer.

2.12 MISCELLANEOUS MATERIALS.

Fuji Rx medical X-ray film : Fuji Photo Film Co.

Positive/negative Land Pack film : Polaroid.

Nitrocellulose : Schleicher and Schuell BA85 (0.45 u).

Ultrafiltration membrane filters : Millipore (0.45 u).

Glass microfilters 2.5 cm : Whatman Ltd.

DE81 2.3 cm filters : Whatman Ltd.

Dialysis membrane (18/32) : Union Carbide.

Microtitre plates (96 V shaped wells) : Flow Lab. Inc.

Centrifuge rotors: Beckman.

2.13 STORAGE OF BACTERIA AND BACTERIOPHAGE.

Bacterial stocks for short term storage were maintained on the appropriate plates at 4°C. Long term storage of bacterial cultures was at -80°C after addition of glycerol to a final concentration of 40%.

Low titre stocks of M13 recombinant phage were maintained in 2xYT broth at -20°C. Low titre 186 phage stocks were passed through 0.45 u Millipore membrane filters and stored at 4°C. High titre 186 phage stocks prepared by CsCl block gradient centrifugation (Section 2.23.1), were dialysed three times against one litre of TM and stored at 4°C.

2.14 GROWTH OF BACTERIAL STRAINS.

Stationary phase bacterial cultures were prepared by inoculating broth with a single colony of bacteria from a plate stock or a loopful of bacteria directly from the glycerol stock, and incubating overnight in capped flasks at the appropriate temperature (usually 30°C or 37°C) in a New Brunswick gyrotary water bath.

Log phase cultures and indicator bacteria were prepared by diluting a fresh stationary culture one hundred fold into sterile broth and incubating with aeration in a gyrotary water bath at the appropriate temperature, until the required cell density was reached (usually $3x10^8$ cfu/ml, which occurs at A_{600}= 0.8 in L broth). Cell density was measured by observing the A_{600} using a Gilford 300 T-1 spectrophotometer. Indicator bacteria were chilled and kept on ice until required.

2.15 CONSTRUCTION OF BACTERIAL STRAINS.

2.15.1 Construction of 186 lysogens.

The phage were plated for single plaques on a lawn of the bacterial strain to be lysogenized and the plates were incubated at 30°C overnight. The centre of a turbid plaque was streaked for single colonies on the appropriate plate and the plate incubated at 30°C. Single colonies were tested, along with the appropriate controls of a lysogen and a non-lysogen, for immunity to 186 by cross-streaking against 186 *c*I10 and 186 *vir*1 (Section 2.2). A colony, which was capable of growth over 186 *c*I10, but not 186 *vir*1 was considered to be lysogenic. This colony was purified two times by streaking for single colonies on the appropriate plates at 30°C and several of these colonies tested for immunity to 186 as described above.

To make 186 amber lysogens of non-permissive (*sup*$^+$) strains, 10^7-10^9 phage were spotted onto the approprate *sup*$^+$ lawn and the plate incubated at 30^{o}C. The centre of the spot was then streaked for single colonies on the appropriate plates and these were tested for lysogeny and purified as described above.

2.15.2 Transformation with plasmids.

2.15.2(a) Preparation of competent cells.

Competent cells were prepared from 20 - 50 ml cultures grown in L broth with aeration to A_{600}= 0.3 - 0.6. After chilling on ice for 10 min, the cells were collected by centrifugation (6,000 rpm, 10 min, 4^{o}C, JA20 rotor), resuspended in 10 ml ice cold 100 mM $CaCl_2$ and left on ice for at least twenty minutes. Cells were again collected by centrifugation as above and resuspended in 1 - 2 ml of ice cold 100 mM $CaCl_2$ and left on ice for at least 2 hours. Competent cells were always prepared and used on the same day.

2.15.2(b) Transformation.

Plasmid DNA or an aliquot of the reaction mix from a ligation reaction (Section 2.28.3(a)) was added to 0.1 ml competent cells in chilled, sterile glass tubes and kept on ice for 10 min. The cells were heat shocked by incubation at 37^{o}C for 5 min and left on ice for a further 10 min. L broth (1 ml) was added to each tube and the tubes were incubated in a gyratory water bath at the appropriate temperature for 1 - 2 hours, to allow the expression of the antibiotic resistance gene(s) present on the plasmid. An aliquot (0.01 - 0.1 ml) from each transformation was plated on the appropriate plates supplemented with antibiotic(s) and the plates incubated at the appropriate temperature.

Plasmids transformed into bacterial strains using the procedure described above were purified two times by streaking for single colonies on the appropriate plates supplemented with antibiotic(s).

2.15.3 Curing cells of plasmids.

Cells were cured of ColE1-related plasmids by several passages in M13 minimal media without the antibiotic conferred by the plasmid to be cured. Cells were then spread

onto plates with and without the selective antibiotic, to assess the extent of plasmid loss. Single colonies were tested by spotting suspensions of the cells (in TM) onto plates with and without antibiotic. Single colonies which were not resistant to antibiotic, were purified by streaking for single colonies two times and were assumed to be cured of the plasmid.

2.16 PHAGE AND BACTERIAL ASSAYS.

2.16.1 Phage assays.

186 phage lysates were assayed for phage by mixing 0.1 ml of the phage diluted in TM buffer, 0.2 ml of log phase indicator bacteria (Section 2.14) and 3 ml melted (0.7%) soft agar overlay and pouring the mixture onto TB plates. The agar was allowed to solidify and the plates were inverted and incubated overnight at the appropriate temperature (usually 30°C or 37°C). Plaques were counted and scored as plaque forming units per ml (pfu/ml).

2.16.2 Bacterial assays.

Bacterial viable counts were measured by diluting a culture of cells in the appropriate media and spreading the appropriate dilution onto the appropriate plates supplemented with antibiotics when necessary. The plates were incubated at the appropriate temperature and the colonies were counted and scored as colony forming units per ml (cfu/ml).

2.17 186 BURST ANALYSIS.

186 phage burst analysis was performed by UV irradiation of cells lysogenic for 186 (UV induction). The phage burst size was determined by calculating the difference between the number of infectious centres before the phage burst and the highest number of pfu/ml produced after the phage burst. Cells lysogenic for wild-type 186 were grown at 30°C in M13 minimal media to A_{600}= 0.3 ($3x10^8$ cfu/ml). For UV irradiation, cultures were centrifuged (6,000 rpm, 10 min, 4°C, JA20 rotor) and bacteria resuspended in an equivalent volume of 10 mM $MgSO_4$. Portions were transferred to glass Petri dishes to a

depth of 1 mm and UV irradiated at 15 J/m^2 (254 nm) with a 15 W General Electric germicidal lamp. Bacteria were then diluted into M13 minimal media and antibiotics and incubated with aeration at 30°C. Samples were taken after 4 hours, treated with chloroform and assayed for free phage.

2.18 GALACTOKINASE ASSAYS.

Galactokinase assays were carried out to quantitate the level of *galK* expression from clones in the McKenney promoter analysis vectors (pKO2 or derivatives constructed in this work, Section 2.3.2). Assays were usually performed in cells with the genotype $galK^-E^+T^+$ (E863 or E864, Table 2.1). For assays performed in cells with a $galK^+$ genotype, the level of chromosmal galactokinase activity was deducted from the total galactokinase activity observed from the plasmid clone. The level of *galK* expression from clones in the McKenney promoter-analysis vectors correlates with the strength of a promoter (McKenney *et al.*, 1981; Rosenberg *et al.*, 1983). Crude estimations of the level of *galK* expression from a particular clone could be obtained from the colour of colonies containing the clone on MacConkey galactose plates. On these plates strains deficient in galactokinase activity ($galK^-$) give white colonies and strains with galacktokinase activity ($galK^+$) give red colonies (McKenney *et al.*, 1981).

The galactokinase assay procedure was adapted from the method described by Wilson and Hogness (1966) and Adhya and Miller (1979) with modifications as described by Dodd *et al.* (1990).

2.18.1 Preparation of lysates.

Cultures to be assayed for galactokinase activity were grown overnight at the appropriate temperature in M63 medium containing 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 1% glucose, 0.2% casein amino acids (vitamin free), 10 ug/ml thiamine and antibiotics. The overnight culture was diluted one hundred fold into the same medium and grown at the appropriate temperature to an A_{650}= 0.2 - 0.6. Galactokinase expression from late promoters, which were activated upon heat induction of a 186 *cIts* lysogen (Section 4.2.4), were determined by growing cultures at 30°C, transferring to 39°C (heat-induction) and taking samples at the appropriate times (as indicated in the text). Determination of the

galactokinase activity from a promoter activated by the expression of the *B* gene product from a plasmid under temperature and/or IPTG control (pEC434, Section 5.2.4), was by transferral of the culture from 30°C to 42°C and the addition of IPTG (1 mM final concentration) and sampling at the appropriate time (as indicated in the text).

To assay, 1 ml aliquots were taken from the culture and placed into eppendorf tubes and the cells lysed by the addition of 40 ul of lysis buffer (100 mM dithiothreitol, 100 mM EDTA, 50 mM Tris-HCl, pH 8.0) and 3 drops of toluene, followed by vortexing for 30 seconds and incubation at room temperature until the toluene had formed a single drop at the surface of the tube. The aqueous phase was removed into another tube and placed on ice. If a promoter giving more than 50 galactokinase units were to be assayed, the cells were diluted five to ten fold in growth medium before the lysis procedure and aliquots were taken from this diluted lysate to be assayed for galactokinase activity.

2.18.2 Galactokinase assay.

To assay for galactokinase activity, 20 ul of the lysate was added to 80 ul of the reaction solution [4 mM rATP, 5 mM $MgCl_2$, 125 mM Tris-HCl, pH 7.9, 1.25 mM dithiothreitol, 4 mM NaF and 200 nM ^{14}C-galactose (8 - 12 uCi)]. The mixture was incubated at 32°C and aliquots (5x 15 ul), were taken at intervals from 0-30 min, samples spotted onto dry DE81 2.3 cm filters, allowed to bind for 2 min at room temperature and then placed into individual pots containing 15 ml of water to stop the reaction. After all the samples had been taken, filters were batch-washed twice in 1 litre of water, except for filters in which total amount of radioactivity added to each filter (average of the unwashed filters) was to be determined. Filters were dried for 15 min at 100°C, scintillation fluid added and radioactivity present on each filter was measured in a Beckman LS 7500 liquid scintillation spectrometer.

The bound cpm (washed filters) were plotted against time and the slope of the line (Δcpm/ min) was calculated. Galactokinase enzyme units (nmol galactose phosphorylated/ min/ ml of cells/ A_{650}) were calculated using the following formula : Units = Δ cpm/min x 1/total cpm x $1/A_{650}$ x 10,400.

2.18.3 Copy number estimation.

The copy numbers of plasmid clones in the McKenney vector were determined using the technique of Projan *et al.* (1983), except for the substitution of lysozyme for lysostaphin to lyse the cells and of RNase A instead of pancreatic ribonuclease to remove RNA. A sample of the culture (1 ml) to be assayed for galactokinase activity, was used to determine the copy number of the plasmid clone directing *galK* expression. Whole cell lysates prepared as described by Projan *et al.* (1983) were fractionated by agarose gel electrophoresis , stained with ethidium bromide and photographed under short wave UV light (Section 2.26.1). The copy number of the clone of interest relative to the copy number of plasmid pKO2 (at 37^0C) was determined by comparison of plasmid band intensities (corrected for cell densities as determined from the A_{650} of the individual cultures) from the photographic negative using a Zeinch scanning laser densitometer. Any significant differences observed in the copy number of the plasmid clones, were noted and the *galK* units were adjusted accordingly.

2.19 PREPARATION OF PHAGE STOCKS.

2.19.1 Low titre phage stocks.

Low titre stocks (10^9 - 10^{10} pfu/ml) of 186 strains were prepared by heat induction or liquid infection, as described by Hocking and Egan (1982c).

2.19.2 High titre phage stocks by heat induction.

A 186 *c*Its lysogenic culture was grown overnight at 30°C and then diluted one hundred fold into 2 aliquots of 500 ml L broth and incubated at 30°C with aeration to A_{600}= 0.8. The culture was transferred to a 39°C bath and the A_{600} followed at regular intervals. Once lysis was evident upon a decrease in the A_{600} of the culture, chloroform (2 ml) was added and the culture was left at 4°C for 10 min to lyse any remaining cells. Bacterial debris was removed by centrifugation (9,000 rpm, 4°C, 20 min, JA10 rotor) and the supernatant was decanted. NaCl and PEG 8000 were added to a final concentration of 0.5 M and 10% (w/v), respectively, and precipitation was allowed to proceed overnight at 4°C. The precipitate was collected by centrifugation (9,000 rpm, 4°C, 20 min, JA10

rotor), resuspended in 8 ml of TM and then purified by CsCl block gradient centrifugation (Section 2.23.1).

2.19.3 High titre phage stocks by liquid infection.

A fresh overnight culture of E508 was diluted one hundred fold into two aliquots of 500 ml of L broth and incubated at 37°C with aeration to A_{600}= 0.5. The culture was infected (moa= 0.1) with phage from a low titre stock (Section 2.19.1). and incubation was continued at 37°C until lysis was evident from a drop in the A_{600} of the culture, or for 4 hours after infection if lysis did not occur. Chloroform was then added, the culture was treated as described in Section 2.19.2 and the phage were purified by CsCl block gradient centrifugation (Section 2.23.1).

2.20 PHAGE DNA PREPARATION.

Phage DNA was prepared either by phenol extraction of CsCl purified high titre phage stocks (Section 2.24.2), or by using the method described below.

A 50 ml culture of E508 was infected at A_{600}= 0.5 with phage at a moa of 0.1 and incubated at 37°C until lysis occurred as determined from the A_{600} of the culture, or for 4 hours if lysis did not occur. Alternatively, a 50 ml culture of a 186 *cIts* lysogen was incubated with aeration to A_{600}= 0.8 and then transferred to a 39°C bath and incubated until lysis was complete as described above. Chloroform was then added and the culture was left at 4°C for 10 min. Cell debris was removed by centrifugation (10,000 rpm, 10 min, 4°C, JA20 rotor) and then DNase I (50 ug) and RNase A (100 ug) were added and the mixture incubated for 1 hour at 4°C. The lysate was again centrifuged (10,000 rpm, 10 min, 4°C, JA20 rotor) to remove any remaining cell debris. The phage particles were then pelleted by centrifugation (20,000 rpm, 3 hours, 4°C, JA20 rotor) and the pellet was resuspended overnight in 400 ul of TE. Proteinase K buffer (40 ul; 100 mM Tris-HCl, pH 8.0, 100 mM NaCl, 100 mM EDTA, 5.0% SDS) was added, followed by the addition of 5 mg proteinase K, and the digestion was allowed to proceed for 1 - 2 hours at 37°C. The phage DNA was then extracted twice with phenol (Section 2.24.3) and the DNA ethanol precipitated using Na acetate (final concentration of 0.3 M) and 2.5 volumes of ethanol,

washed with 70% ethanol (v/v in TE) and 95% ethanol, dried *in vacuo* and resuspended in 100 ul TE.

2.21 PLASMID DNA PREPARATION.

2.21.1 Analytical preparation.

The following procedure based on the method of Birnboim and Doly (1979) gave sufficient plasmid DNA for several restriction analyses from 1.5 ml of a stationary phase, plasmid-containing culture.

Bacteria grown in the appropriate medium containing the required antibiotic(s) were pelleted by centrifugation (10,000 g, 1 min, room temperature, eppendorf centifuge) and the pellet resuspended in 100 ul of freshly prepared lysis buffer (25 mM Tris-HCl, pH 8.0, 10 mM EDTA, 50 mM glucose, 5 mg/ml lysozyme). After 15 - 30 min incubation on ice, 200 ul of freshly prepared 0.2 M NaOH, 1% SDS added, mixed gently and the tube left on ice for 5 min. 3 M Na acetate, pH 4.6 (150 ul) was then added and the tube was inverted several times and placed on ice for 15 - 30 min. The mixture was centrifuged twice (10,000 g, 10 min, 4^{o}C, eppendorf centrifuge) to remove chromosomal DNA and the plasmid DNA in the supernatant was precipitated by addition of 2.5 volumes of ethanol (Section 2.24.3). The pellet was washed in 70% ethanol (v/v in TE) and 95% ethanol, dried *in vacuo* and redissolved in 20 ul TE. RNase A (final concentration of 50 ug/ul) was added, the solution incubated for 30 min at 37^{o}C and restriction analysis performed (Section 2.25).

2.21.2 Large scale preparation.

Preparative amounts of plasmid DNA were prepared by centrifugation in a CsCl gradient. A plasmid containing cell culture was grown overnight in the appropriate medium containing the required antibiotic. The overnight was used to inoculate two 500 ml cultures and these grown with aeration for 12 - 16 hours. Cells were sedimented by centrifugation (5,000 rpm, 10 min, 4^{o}C, JA10 rotor), resuspended in 7.5 ml of 25% sucrose, 50 mM Tris-HCl (pH 8.0), and then 2.0 ml of a freshly prepared 10 mg/ml solution of lysozyme in 25% sucrose, 50 mM tris-HCl (pH 8.0)was added. The cells were

placed on ice for 5 min and 3 ml of 0.25 M EDTA, pH 8.0, added. After gentle mixing, the tube was again placed on ice for 5 min. Detergent solution (12 ml of 1% Brij 58 (w/v) and 0.4% Na deoxycholate (w/v) in 50 mM Tris-HCl, pH 8.0, 25 mM EDTA) were added, mixed gently, and the mixture left on ice for a further 10 min. After centrifugation (18,000 rpm, 30 min, 4°C, JA20 rotor), the supernatant was removed and the plasmid DNA was purified by CsCl equilibrium gradient centrifugation (Section 2.23.2).

2.22 M13 REPLICATIVE FORM (RF) PREPARATION.

M13 RF DNA was prepared by modification of the preparative plasmid method (Section 2.21.2). An overnight culture of JM101, grown in M13 minimal medium, was subcultured into 800 ml of 2xYT broth and grown at 37°C with aeration to A_{600}= 0.8. The overnight culture of JM101 was also subcultured into 5 ml of 2xYT broth, grown with aeration to A_{600}= 0.6, the required M13 single stranded DNA phage stock (200 ul; Section 2.28.4(d)) added and the M13 infected culture grown at 37°C for one hour. The M13 infected culture was then added to the 800 ml culture of JM101 at A_{600}= 0.8 and grown for six to eight hours at 37°C. Cells were collected by centrifugation (5,000 rpm, 10 min, 4°C, JA10 rotor) and the M13 RF DNA isolated by the plasmid preparation procedure described previously (Section 2.21.2), followed by CsCl equilibrium density gradient centrifugation (Section 2.23.2) to further purify the RF DNA.

2.23 CsCl DENSITY GRADIENT CENTRIFUGATION.

2.23.1 CsCl block density gradient for preparation of high titre phage stocks.

CsCl block density gradient centrifugation was used to prepare high titre 186 phage stocks for the preparation of DNA by phenol extraction (Section 2.24.2).

Two CsCl solutions of density 1.6 gm/ml and 1.35 gm/ml were prepared in sterile TM and were used to form a block gradient by adding 4 ml of the 1.35 gm/ml solution and underlaying it with 1.0 ml of the 1.6 gm/ml solution in a 10 ml polycarbonate Oakridge tube. The high titre phage suspension in TM (Section 2.19.2, 2.19.3), was carefully layered on top of the gradient and the tube centrifuged (45,000 rpm, 90 min, 8°C, Beckman Ti50 rotor). The opaque phage band was collected by piercing the side of the

tube with a syringe and dialysed three times against one litre of TM and stored at 4°C. A titre of 10^{12} - 10^{13} pfu/ml was usually obtained by this method.

Dialysis tubing (Section 2.12) was prepared by boiling 10 cm strips in 2% sodium bicarbonate, 1 mM EDTA for 10 min and storing in 20% ethanol.

2.23.2 CsCl equilibrium density gradient for plasmid purification.

CsCl equilibrium density gradient centrifugation was used to prepare plasmid and M13 RF DNA. DNA was purified by adding 0.95 gm of CsCl and 200 ul of a 10 mg/ml solution of ethidium bromide solution, per ml of plasmid DNA solution (Section 2.21.2) and the solution loaded into a 10 ml polycarbonate tube and centrifuged to equilibrium (45,000 rpm, 44 hours, 20°C, Beckman Ti50 rotor). The bands were visualised under subdued fluorescent light and the lower of the two bands, containing the plasmid or M13 RF DNA, was collected by piercing the tube from the bottom. The upper band contained chromosomal DNA and nicked plasmid DNA. Ethidium bromide was removed by three extractions with isopropanol equilibrated with 5 M NaCl, 10 mM Tris-HCl, pH 8.0, 1 mM EDTA. The DNA solution was then diluted two fold with water and ethanol-precipitated with 6 volumes of ethanol (Section 2.24.3). The pellet was washed with 70% ethanol (v/v) and 95% ethanol, dried *in vacuo*, resuspended in TE and stored at 4°C. The concentration and purity of the DNA was determined by obtaining the spectra of absorbance over the range 230-340 nm on a Varian Superscan 3 ultraviolet/visible spectrophotometer (A_{260}= 1.0 represents a concentration of 50 ug/ul). A_{260}/A_{280} and A_{260}/A_{230} ratios were greater than 1.8 for all DNA used, indicating low protein contamination.

2.24 PHENOL EXTRACTION OF DNA.

2.24.1 Phenol equilibration and storage.

To 50 ml of redistilled phenol was added 2.5 mg of 8-hydroxy-quinoline. Buffer equilibration was carried out by the addition of 50 ml of 1 M Tris-HCl (pH 8.0) and the mixture heated until the phenol and aqueous layers had become one. The phases were allowed to separate and the phenol phase was taken and mixed with 50 ml of TE and allowed to stand until the phases separated. The aqueous phase was again removed and the

phenol phase was mixed twice more with TE. Phenol equilibrated with TE in this manner was stored under TE and kept frozen in 10 ml aliquots at -20°C until required.

2.24.2 Phenol extraction of bacteriophage DNA.

A high titre phage stock (10^{12} pfu/ml; Section 2.23.1) was diluted to 0.9 ml in TE, and then 0.1 ml of 100 mM Tris-HCl, pH 8.0, 100 mM NaCl, 100 mM EDTA, 5.0% SDS was added, followed by 5 mg of proteinase K. After incubation at 37°C for 60 min, the solution was again diluted (to 5 ml) with TE, and an equal volume of TE-equilibrated phenol (Section 2.24.1) added. The mixture was gently shaken for 5 min and the phases were separated by centrifugation (7,000 rpm, 5 min, 20°C, JA20 rotor). The aqueous phase was collected and re-extracted at least twice with an equal volume of TE equilibrated phenol. The phenol phases were washed with an equal volume of TE, and Na acetate (pH 4.6) was added to the pooled aqueous phases to a final concentration of 0.3 M, followed by 2.5 volumes of ethanol. DNA was precipitated at -20°C for 30 min and was collected by centrifugation (18,000 rpm, 20 min, 4°C, JA20 rotor). The pellet was washed in 70% ethanol (v/v) and 95% ethanol, dried *in vacuo* and dissolved in TE and stored at 4°C. The concentration and purity of the DNA was determined, as described in Section 2.23.2.

2.24.3 Phenol extraction and ethanol precipitation of DNA solutions.

DNA solutions were vortexed thoroughly with one volume of TE-equilibrated phenol, kept at room temperature for 5 min and centrifuged (10,000 g, 3 min, room temperature, eppendorf centifuge or 7,000 rpm, 5 min, 20°C, JA20 rotor for 50 ml Oakridge tubes). The aqueous phase was removed, and the process repeated until no material was visible at the interface of the aqueous and phenol phases. The phenol phases were washed with an equal volume of TE. The aqueous phases were pooled and Na acetate (pH 4.6) was added to 0.3 M followed by 2.5 volumes of ethanol. DNA was ethanol precipitated at -20°C for 15 min and the DNA collected by centrifugation (10,000 g, 3 min, 4°C, eppendorf centifuge or 16,000 rpm, 20 min, 4°C, JA20 rotor). The pellet was rinsed in 70% ethanol (v/v) and 95% ethanol, dried *in vacuo* for 10 min, dissolved in TE and stored at 4°C.

Ethanol precipitation of solutions that did not require phenol extraction was by the addition of sodium acetate (pH 8.0) to 0.3 M final concentration and 2.5 volumes of ethanol. DNA was precipitated at -20°C for 15 min and the DNA collected by centrifugation in 50 ml Oakridge tubes (16,000 rpm, 20 min, 4°C, JA20 rotor) or in eppendorf tubes (10,000 g, 15 min, 4°C, eppendorf centrifuge) and the pellet rinsed in 70% (w/v) ethanol and 95% ethanol, dried *in vacuo* and dissolved in TE and stored at 4°C.

2.25 RESTRICTION ANALYSIS OF DNA.

Analytical digestion of DNA with restriction endonucleases was carried out for two to twelve hours at the recommended temperature with a two to five-fold excess of enzyme (2 to 5 units per ug of DNA) in a volume of 10 - 20 ul. Digestion buffers were those specified by the manufacturers' catalogues and were stored at -15°C as 10x stocks. For double digestions with enzymes requiring different salt conditions, the enzyme with the lowest NaCl concentration was used first and the concentration of NaCl was then raised to the appropriate amount and the second enzyme was added. Preparative digests of 20 - 50 ug of DNA were in 50 - 100 ul reaction volumes and were incubated overnight at 37°C. Reactions were terminated by heating to 70°C for 10 min and DNA was tested for complete digestion by agarose gel electrophoresis on minigels (Section 2.26.1).

2.26 GEL ELECTROPHORESIS.

2.26.1 Agarose gel electrophoresis of DNA.

Agarose gel electrophoresis (1% to 2% w/v agarose in TAE, stored at 65°C) was carried out using horizontal minigels prepared by pouring 9.5 ml of gel solution onto a 7.5 cm x 5.0 cm glass microscope slide or 18 ml of gel solution onto a 7.5 cm x 10.0 cm glass microscope slide, with the appropriate comb set in place. Glycerol loading buffer (10x; Section 2.10) was added to the samples to a final concentration of 5% glycerol, 0.04% bromophenol blue, 0.02% xylene cyanol, 1 mM EDTA and electrophoresis was carried out in TAE buffer at 75 - 150 mA.

DNA was visualised by staining gels with ethidium bromide (4 mg/ml (w/v) in TAE). Gels were photographed under short wavelength UV light. Approximate

concentrations of DNA solutions were determined by comparing the intensity of the ethidium bromide stained bands with that of known concentrations of DNA molecular weight markers (Section 2.11).

Preparative gels were poured on glass slides, as described above, or in a perspex tray (14 x 11 x 0.3 cm) and run horizontally, submerged in TAE. Low gelling temperature (LGT) agarose gel solution (1.0% w/v in TAE) was cooled to 37°C before pouring and electrophoresis was carried out at 75 - 100 mA at 4°C. DNA was visualized as described above.

2.26.2 Polyacrylamide gel electrophoresis.

2.26.2(a) Non-denaturing gels.

A 30% gel stock (acrylamide:bis, 30:0.8) was prepared by dissolving 146.1 gm acrylamide and 3.9 gm bis-acrylamide in 500 ml of glass distilled water. The solution was de-ionised (Section 2.26.2d) and then de-gassed for 30 min using a vacuum pump. Generally 50 ml of acrylamide gel solution (5% in 1x TBE) was prepared and 375 ul of freshly prepared 25% (w/v) APS and 100 ul TEMED added. Gels were poured at room temperature, with polymerization of the gel solution occurring within 5 - 10 min. Polymerization was allowed to proceed for 60 min on the bench to ensure polymerization was complete. Gel dimensions were 20 x 40 x 0.05 cm or 17 x 26 x 0.05 cm. Pre-electrophoresis was at 15 mA for 30 min. Formamide loading buffer (Section 2.10) was added to the samples (one volume of formamide loading buffer to one volume of sample) and electrophoresis at 25 mA. End-labelled pUC19 *Hpa*II DNA fragments (Section 2.11) were used as molecular weight markers.

2.26.2(b) Denaturing (sequencing) gels.

Stock gel solution (6% polyacrylamide; acrylamide:bis 19:1; 8 M urea in TBE) was prepared by dissolving 57 gm acrylamide, 3 gm bis-acrylamide and 480.5 gm urea in 400 ml glass-distilled water at room temperature. The solution was made to 900 ml with glass-distilled water and was then de-ionised (Section 2.26.2d). One hundred ml of 10x TBE was added and the solution was de-gassed, as described in Section 2.26.2(a).

Polymerisation was carried out by adding 400 ul of freshly prepared 25% (w/v) APS and 100 ul of TEMED, to 80 ml of gel stock solution. Gel dimensions were either 20 x 40 x 0.025 cm or 40 x 40 x 0.025 cm. Polymerisation was allowed to occur for 60 min at room temperature.

Pre-electrophoresis was in TBE for 30 min at 800 V with the comb in place to prevent well distortion. The comb was removed immediately prior to loading and the wells were flushed with TBE to remove urea and any unpolymerised material. Electrophoresis was at 1000 V - 1200 V. Band distortion due to localised heating near the centre of the gel was eliminated by placing a metal plate in contact with the outside gel plate.

2.26.2(c) Denaturing (sequencing) gels to resolve band compressions.

Sequencing gels containing 25% or 40% formamide (v/v) were prepared by including the appropriate amount of de-ionized formamide (Section 2.26.2d) in the gel stock solution (Section 2.26.2b). Polymerisation required the addition of 600 ul freshly prepared 25% APS (w/v in water) and 140 ul of TEMED to 80 ml of gel stock solution. Pre-electrophoresis was for 2 to 3 hours at 500 V. Electrophoresis conditions were as described in Section 2.26.2(b).

2.26.2(d) De-ionization of solutions.

Solutions were de-ionized by adding 10% (w/v) of mixed bed resin (Section 2.8) and gently stirring the solution at room temperature for 30 min. Mixed bed resin was removed by filtration.

2.26.3 Autoradiography.

Fuji Rx medical X-ray film was used for autoradiography. Gels were wrapped in plastic (Vitafilm) and exposed at room temperature for up to 24 hours, or for longer periods at -80^{o}C with Tungstate intensifying screens. All sequencing gels were exposed overnight at room temperature. Specific conditions for autoradiography of gels are given in the Figure legends.

2.27 ISOLATION OF DNA FRAGMENTS FROM GELS.

2.27.1 Extraction of DNA from agarose gel slices.

DNA to be extracted from agarose for the purpose of cloning was detected by staining the gel with acridine orange (0.001% w/v in TAE) for 10 min and then de-staining with TAE for at least 15 min. Bands were visualized under fluorescent light. If the amount of DNA in the band to be isolated was less than 1 ug, the gel was stained with ethidium bromide for 5 min (4 mg/ml (w/v) in TAE) rather than acridine orange, and then de-stained for 5 min. The gel was kept in the dark after the addition of ethidium bromide and the bands were visualized by brief exposure to long wavelength UV light.

Agarose containing the desired DNA fragment was excised from the gel with a sterile scalpel blade and the DNA was removed from the agarose slice by either one of the two methods described below.

2.27.1(a) Extraction of DNA from low gelling temperature (LGT) agarose gel slices.

Extraction of DNA from LGT agarose was performed when the DNA fragments to be isolated were greater than 10 kb in size. The LGT agarose gel slice containing the desired DNA fragment was melted at 65°C for 30 min in an eppendorf tube and diluted with an equal volume of TE buffer. If the amount of DNA was less than 1 ug, *E. coli* carrier tRNA (Section 2.8) was added (20 ug). The solution was vortexed and kept at 65°C for a further 15 min. The liquid LGT was then transferred to a 37°C heating block for 15 min and half a volume of ice cold phenol (equilibrated with TE) added and the tubes were immediately vortex-mixed and placed on ice for 5 min. The aqueous phase was recovered after centrifugation (10,000 g, 3 min, room temperature, eppendorf centrifuge) and the phenol layer washed with TE. The pooled aqueous phase was ethanol precipitated after the addition of Na acetate (pH 4.6) to 0.3 M (Section 2.24.3).

2.27.1(b) Extraction of DNA from agarose using the Geneclean procedure.

If the DNA fragments to be isolated were less than 10 kb in size, extraction from normal agarose using the Geneclean procedure (Bio 101, La Jolla, California) was performed according to the manufacturers protocol. DNA fragments of size greater than

10 kb were not isolated using this procedure, since various extents of shearing of DNA due to the extraction process were observed.

2.27.2 Recovery of DNA from polyacrylamide gel slices.

The radioactive fragments of interest were located after autoradiography (Section 2.26.3) by comparison of their sizes to radioactive DNA markers (Sections 2.11. and 2.28.2). For kinasing reactions (Section 2.30.3), oligonucleotides of known size were kinased (Section 2.30.3) and used as radioactive DNA size markers. The autoradiograph was aligned with radioactive ink marks, which had been placed on the edges of the gel prior to autoradiography, and was used as a template to locate the bands of interest. The bands were excised from the gel using a sterile scalpel blade and the DNA was eluted from gel slices by adding 400 ul of gel elution buffer (500 mM ammonium acetate, 1 mM EDTA, 0.1% SDS, pH 7.6) to each individual slice in an eppendorf tube and incubating the tubes overnight at 37°C with constant agitation (Maxam and Gilbert, 1980). The eluate was collected and 2.5 volumes of ethanol added, the DNA precipitated (Section 2.24.3), dissolved in TE and stored at 4°C. The eluate was collected, Na acetate (pH 8.0) added to 0.3 M and the DNA precipitated after the addition of 2.5 volumes of ethanol (Section 2.24.3). For kinased oligonucleotides, 400 ul of TE was added to the gel slice and incubated overnight at 37°C.

2.28 PLASMID AND M13 CLONING.

2.28.1 Preparation of vector DNA for cloning.

Plasmid and M13 cloning vectors were prepared by digesting the vector (Section 2.21.1) or RF DNA (Section 2.22) with the appropriate restriction enzyme(s) (Section 2.25). The volume was adjusted to 50 ul (or by ethanol precipitation (Section 2.24.3) if the restriction volume was greater than 50 ul) and made to 50 mM Tris-HCl, pH 8.0, 0.1 mM $ZnCl_2$, 0.1 mM EDTA. Calf intestinal phosphatase (0.5 - 2.0 U) was added and the mixture incubated for two hours at 37°C. The phosphatase was inactivated by heating to 70°C for 20 min and the mixture extracted with phenol and DNA was ethanol precipitated (Section 2.24.3). If vector DNA were to be purified from any undigested DNA by

fractionation on an agarose gel, followed by recovery of the vector DNA from a gel slice (Section 2.27.1(b)), the phosphatasing reaction mix was not phenol extracted prior to loading onto the gel. Vector DNA was finally dissolved in TE at a concentration of 20 - 50 ng/ul and stored at $4^{o}C$.

2.28.2 End-labelling and end-filling using the large fragment of DNA Polymerase I (Klenow).

DNA to be cloned, which required the ends to be blunt, was end-filled using the Klenow fragment of *E. coli* DNA polymerase I in a reaction containing 50 mM NaCl, 6 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$, 0.05 mM dNTPs (added from a stock containing 0.25 mM of each of the four dNTPs in 5 mM Tris-HCl, pH 8.0, 1 mM EDTA) and 1 unit of Klenow fragment. The reaction was incubated at $37^{o}C$ for 15 min, terminated by heating at $70^{o}C$ for 10 min and ethanol precipitated (Section 2.24.3).

DNA restriction fragments to be used as radioactive size markers, or for analysis of restriction patterns on polyacrylamide gels, were end-labelled with ^{32}P in a reaction mix containing 50 mM NaCl, 6 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$, 1 unit of Klenow fragment and 2 - 4 uM d[α-^{32}P]CTP or d[α-^{32}P]ATP. After incubation at $37^{o}C$ for 15 min the reaction was terminated by heating at $70^{o}C$ for 10 min followed by the addition of formamide loading buffer (Section 2.10) and the samples were loaded onto the gel or stored at $4^{o}C$ until required.

One tenth of any solution containing DNA restriction fragments to be fractionated on polyacrylamide gels (Section 2.26.2(a)) and then cloned (Sections 2.28.3 and 2.28.4), was end-labelled to allow detection of the fragments by autoradiography.

2.28.3 Plasmid cloning.

2.28.3(a) Ligation.

Restricted, phosphatased plasmid vector (20 - 50 ng; Section 2.28.1) was ligated with the DNA fragment to be cloned in a 3:1 molar ratio of insert:vector. The ligation was carried out in a volume of 20 ul containing 50 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 10 mM dithiothreitol, 1 mM rATP, and T4 DNA ligase (0.2 units for staggered end

ligations or 1.0 unit for blunt end ligations). The mixture was incubated for 3 to 16 hours at 15°C and then transformed into bacterial cells (Section 2.15.2).

2.28.3(b) Transformation.

Competent cells were prepared as described in Section 2.15.2(a). An aliquot of the ligation reaction (1 - 10 ul) was transformed into the appropriate strain as described in Section 2.15.2(b) and plated onto selective plates. The following controls were included for each bacterial strain to be transformed with ligated DNA: (1) restricted, phosphatased vector (to test for uncut vector DNA), (2) restricted, phosphatased and religated vector (to assess the phosphatasing reaction), (3) uncut vector, to test the transformation efficiency of the competent cells (a transformation efficiency of 10^5 - 10^6 transformants/ug was achieved for most strains used), (4) untransformed cells (to test for spontaneous antibiotic resistance or contamination in the competent cells).

2.28.3(c) Identification of plasmid recombinants.

Often plasmid recombinants were identified by one or a combination of the following phenotypes : sensitivity to antibiotics (if the fragment had been cloned into an antibiotic resistance gene), the colour of colonies on MacConkey galactose plates (for fragments containing a promoter cloned into *galK* fusion vectors), the ability to complement a 186 phage (deficient in the cloned gene function) in infection, or immunity to 186 (if the fragment cloned expressed the *c*I gene). Plasmid DNA was prepared (Section 2.21.1) and the identity and orientation of the insert DNA confirmed by diagnostic restriction analysis (Section 2.25) and agarose gel electrophoresis (Section 2.26.1). For plasmid recombinants which did not show a characteristic phenotype, plasmid DNA was prepared (Section 2.21.1), and the identity and orientation of the insert DNA determined by diagnostic restriction analysis (Section 2.25) and agarose gel electrophoresis (Section 2.26.1).

If a large number of putative clones were to be screened, colony hybridization was performed. This procedure was carried out essentially as described by Maniatis *et al.* (1982) and was based on the procedure of Grunstein and Hogness (1975). Colonies to be tested were spotted onto the appropriate plate with the required antibiotic. Controls

containing the parent vector and a clone containing a fragment overlapping the region of interest were spotted onto the same plate. After incubation at the appropriate temperature, colonies were transferred to a nitrocellulose filter by placing a dry filter on top of the agar plate. The filter and underlying agar were marked with a needle (to allow later alignment). The filter was allowed to become completely wetted, removed and the agar plate reincubated at the appropriate temperature to allow the colony spots to regenerate. The nitrocellulose filter was placed colony side up, for 5 min, on Whatman 3mm filter papers, which were pre-soaked in each of the following solutions : 5% SDS; denaturing solution (0.5 M NaOH, 1.5 M NaCl); neutralizing solution (1.5 M NaCl, 0.5 M Tris-HCl, pH 8.0); and 2x SSPE (0.36 M NaCl, 20 mM NaH_2PO_4, pH 7.4, 2 mM EDTA). The nitrocellulose filter was then dried colony side up on Whatman 3mm filter paper for 30 min at room temperature and baked for 2 hours at 80°C in a vacuum oven. The filter was floated on the surface of a solution of 6x SSC in a beaker for 1 min and then was submerged for 5 min to thoroughly wet the filter. The filter was then washed in a solution of 50 mM Tris-HCl, pH 8.0, 1 M NaCl, 1 mM EDTA, 0.1% SDS at 42°C for 2 hours to remove any agar fragments or loose bacterial debris. The filter was placed in a plastic bag and pre-hybridization and hybridization carried out, using an appropriate radioactive probe, which was prepared by oligo-labelling (Section 2.30.1). The filter was prehybridized for 2 - 12 hours at 42°C in 50% formamide, 5x SSC, 5x Denhardts, 50 mM Na phosphate (pH 6.5) and 0.1 mg/ml sonicated and heat-denatured calf thymus DNA (denatured as described in Section 2.32.5). Hybridization was at 42°C overnight by the addition of heat-denatured, oligo-labelled probe (denatured as described in Section 2.32.5) to the prehybridization solution. After hybridization, the filter was washed 4 times for 5 min in 2x SSC, 0.1% SDS at room temperature followed by 2 washes of 1 hour each in 1x SSC, 0.1% SDS at 65°C. The filter was then placed on Whatman 3mm paper, as ymetrical marks were made around it with radioactive ink and the filter wrapped in plastic (Vitafilm). The filter was autoradiographed, as described in Section 2.26.3. The autoradiograph was aligned with the agar plate and colonies, which hybridized with the probe, were tooth-picked from the plate and purified by streaking for single colonies. Plasmid DNA was prepared (Section 2.21.1) and tested for insert size and orientation by restriction analysis (Section 2.25) and agarose gel electrophoresis (Section 2.26.1).

2.28.4 M13 cloning.

2.28.4(a) Ligation.

M13 vector DNA, which had been digested with the appropriate restriction enzyme(s), treated with calf intestinal phosphatase and purified (Section 2.28.1), was kept at a constant 20 ng per ligation. The DNA fragment to be cloned was mixed with M13 vector in a insert to vector molar ratio of 3:1. The required ratio was achieved by adding 10 ng/kb of insert DNA. Ligation was carried out in a 20 ul reaction containing 50 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 10 mM dithiothreitol, 1 mM rATP, and T4 DNA ligase (0.2 units for staggered end ligations or 1.0 units for blunt end ligations). The reaction was incubated for 2 - 12 hours at 15°C and the ligation mix used to transfect competent JM101 cells.

2.28.4(b) Transfection.

For transfection with M13 clones, competent JM101 cells were prepared from a 50 ml culture of JM101 in 2xYT broth (inoculated from a fresh overnight in M13 minimal medium) grown at 37°C with aeration to A_{600}= 0.8. The culture was left on ice for 10 min, and the cells collected by centrifugation (6,000 rpm, 10 min, 4°C, JA20 rotor). The cells were resuspended in 1 - 2 ml ice-cold 100 mM $CaCl_2$ and left on ice for at least two hours. Competent cells were always prepared and used on the same day. An aliquot (1 - 10 ul) of the ligation mix (Section 2.28.4(a)) was added to 0.2 ml of competent cells in chilled, sterile glass tubes and the mixture kept on ice for 40 min. The cells were then heat shocked by incubation at 45°C for 2 min and 3 ml of molten YT soft agar containing 20 ul of IPTG (24 mg/ml), 20 ul of BCIG (20 mg/ml) and 0.2 ml of log phase (A_{600}= 1.0) JM101 were added. The mixture was poured onto an M13 minimal plate and the plates were incubated overnight at 37°C. M13 recombinants (clones) usually appeared as white plaques whereas parental M13 phage plaques were blue.

The following controls were included with each transfection: (1) restricted, phosphatased vector (to test for contamination by uncut vector), (2) restricted, phosphatased and religated vector (to test that the vector DNA alone was not giving rise to

white plaques), (3) uncut vector (to test the transfection efficiency of the competent cells), (4) untransfected cells (to test for M13 contamination).

2.28.4(c) Identification of M13 recombinants.

To identify M13 recombinants, white plaques were tooth-picked from the appropriate plates (Section 2.28.4(b)) and single stranded DNA phage stocks were prepared (Section 2.28.4(d)). The phage were analysed for inserted DNA by lysing the phage and sizing the single stranded DNA by agarose gel electrophoresis (Section 2.28.4(f)). Single stranded DNA was prepared (Section 2.28.4(e) from M13 phage containing inserts identified in this way, and if necessary the DNA was subjected to a complementarity test (Section 2.28.4(g)) to determine the orientation of the insert. The DNA sequences of the clones were then determined (Section 2.29).

2.28.4(d) Preparation of M13 single stranded DNA phage stocks.

A fresh stationary phase culture of JM101 grown in M13 minimal medium, was diluted one hundred fold into 2xYT broth and two ml aliquots of the diluted culture dispensed into 10 ml screw-capped polycarbonate tubes. M13 plaques were toothpicked into the cultures and incubated at 37^{o}C with constant agitation for 5 - 7 hours. The cultures were then centrifuged in a bench centrifuge (6,000 rpm, 10 min, room temperature) and the supernatants, containing the M13 single stranded DNA phage, carefully transferred into an eppendorf tube and centrifuged (10,000 g, 10 min, room temperature, eppendorf centrifuge). One ml of supernatant was transferred into another eppendorf tube and 270 ul of PEG solution (20% PEG w/v, 2.5 M NaCl) was added and the solutions mixed. Phage particles were allowed to precipitate for 15 min at room temperature and the phage pellets collected by centrifugation (10,000 g, 5 min, room temperature, eppendorf centrifuge) and the supernatants withdrawn by aspiration. Tubes were briefly recentrifuged again and any traces of the supernatant removed. The phage pellets were resuspended in 200 ul of TE. Phage stocks prepared in this manner were stored at -20^{o}C.

2.28.4(e) Preparation of M13 single stranded DNA.

To prepare M13 single stranded DNA, the phage stocks (Section 2.28.4(d)) were phenol extracted with one half volume TE-saturated phenol (Section 2.24.3) at room temperature and the phases separated by centrifugation (10,000 g, 3 min, room temperature, eppendorf centrifuge). A portion of the aqueous phase (150 ul) was carefully withdrawn, avoiding the interface of the two phases, the DNA ethanol precipitated after the addition of 15 ul of 3 M Na acetate (pH 4.6) and 400 ul of ethanol (Section 2.24.3), washed with 1 ml of 95% ethanol, dried *in vacuo* and dissolved in 25 ul of TE.

2.28.4(f) Sizing of M13 single stranded DNA by agarose gel electrophoresis.

Potential recombinant phage were tested for the presence of the cloned DNA by treating 10 ul of single stranded DNA phage stock (Section 2.28.4(d)), with 2 ul of SDS lysis buffer (0.1% bromophenol blue, 3 mM EDTA, pH 8.0, 300 mM NaCl, 28% glycerol, 2% SDS), heating the mixture at 65^{o}C for one hour and incubation on ice for 5 min. DNA liberated from the phage in this manner was electrophoresed on minigels (Section 2.26.1). Recombinant M13 DNA had a lower mobility on these gels than a control with no inserted DNA.

2.28.4(g) Complementarity test for M13 single stranded DNA clones.

To determine the orientation of cloned DNA, complementation tests were performed. M13 single stranded DNA template DNA (2 ul; Section 2.28.4(e)) to be tested was added to 2 ul of reference DNA (a M13 single stranded DNA clone of known orientation and spanning the region, from which the test clone was derived) and 8 ul of 0.25 M NaCl. The mixture was incubated at 60^{o}C for one hour and then placed on ice for 5 min. Glycerol loading buffer was added to the samples and DNA was electrophoresed on a 1% agarose minigel (Section 2.26.1) at 50 - 100 mA. DNA was visualised by ethidium bromide staining (4 mg/ml EtBr in TAE). If the test DNA was not complementary to the reference clone a single band was observed, whereas clones containing DNA sequences complementary to the reference clone were able to hybridize in this region, increasing their molecular weight and retarding their mobility on the gel. The following controls were performed : (1) reference DNA plus M13 single stranded DNA

without an insert was treated, as described above, and was shown not to produce any hybrid bands. (2) As a positive control the reference DNA was hybridized, as described above, to a M13 single stranded DNA clone of known opposite orientation.

2.29 DNA SEQUENCING.

The dideoxynucleotide chain termination sequencing technique (Sanger et al., 1977; 1980; Schreier and Cortese, 1979) was used to determine DNA sequence. DNA sequencing was performed using BRESATEC Dideoxy Sequencing kits and the recommended protocol with the following modifications:

2.29.1 Annealing.

The M13 universal primer (2.5 ng, 1 ul) was annealed to M13 DNA (6 ul; Section 2.28.4(c)) in a reaction volume of 10 ul containing 1x TM (final concentration) by incubation in a 65^{o}C oven for one hour and cooling to room temperature for 30 min.

2.29.2 The extension reaction.

The sequencing reactions were carried out in microtitre plates. d[α-^{32}P]CTP (0.6 uCi per clone to be sequenced) was dried *in vacuo* and resuspended in the appropriate amount of each of the four dNTP/ddNTP reaction mixes. Annealed DNA (2 ul) from each clone was added to 4 separate wells and 2 ul of the appropriate label-dNTP/ddNTP solution added to each well and mixed by spinning for 5 seconds in a microtitre centrifuge (Hettisch Universal 2S, 2000 rpm). Nucleotide concentrations were as given by the manufacturer. The microtitre plate was incubated at 37^{o}C for 10 min to allow heating to 37^{o}C. Klenow enzyme solution (2 ul; 10 units of Klenow enzyme diluted to 0.1 units/ul in TM just before use) was dispensed to the side of each well and the sequencing reactions were commenced by a 5 second centrifugation to mix the enzyme solution with the reaction mix and were incubated at 37^{o}C for 15 min. Reactions were chased by the addition of 2 ul of dNTP-Klenow enzyme solution (0.25 mM of each dNTP and 0.05 units/ul of Klenow enzyme solution in TE, prepared immediately before use) to each tube and incubating for a further 15 min at 37^{o}C. The reaction was terminated by the addition of 4 ul of formamide loading buffer (95% de-ionised formamide, 0.1% bromophenol blue,

0.1% xylene cyanol, 10 mM EDTA, 0.01 M NaOH) followed by a brief spin. Immediately prior to loading the sequencing reactions, the samples in the microtitre tray were heated to 100°C for 5 min and samples (0.5 ul) were then loaded onto 6% denaturing polyacrylamide gel and electrophoresed (Section 2.26.2(b)).

After electrophoresis, one of the glass plates was removed and the gel fixed by washing it with two litres of 10% acetic acid, 20% ethanol (v/v in water). The addition of ethanol prevented the gel from swelling and wrinkling during the fixing process. The gel was dried in a 110°C oven for 30 - 45 min and then autoradiographed (Section 2.26.3).

2.30 PREPARATION OF RADIOACTIVE DNA PROBES.

2.30.1 Oligo-labelling of DNA fragments.

Oligo-labelling was performed using Oligo-labelling kits purchased from BRESATEC and following the recommended protocol. DNA to be used as a probe was restricted (Section 2.25) with the appropriate restriction enzymes and the required DNA fragment isolated using agarose gel electrophoresis (Section 2.26.1) and the Geneclean procedure (Section 2.27.1(b)). The DNA to be used as a probe was generally of a size of 1 kb or less. Denaturation of the DNA fragment and randomly primed synthesis of complementary DNA in a reaction containing d[α-^{32}P]CTP was as described in the protocol. Unincorporated label was removed by extraction of the DNA using the Geneclean procedure (Section 2.27.1(b)) and the DNA resuspended in 100 - 200 ul TE.

2.30.2 Preparation of radioactive DNA probes by primer extension on M13 single stranded DNA clones.

The preparation of ^{32}P-DNA probes from M13 single stranded DNA clones was adapted from the procedure of Bruening *et al.* (1982). M13 single stranded DNA clones with inserts of the same polarity as the RNA to be detected, were used to prepare ^{32}P-DNA probes where only the strand complementary to the RNA was made radioactive.

M13 universal primer (2.5 ng; Section 2.5) was annealed to 8 ul of M13 clone DNA in 10 mM Tris-HCl, pH 8.0, 10 mM $MgCl_2$ at 60°C for one hour and the mixture was allowed to cool to room temperature for 30 min. The extension reaction was

performed for 15 min at 37°C, in a volume of 20 ul containing 10 mM Tris-HCl, pH 8.0, 10 mM $MgCl_2$, 5 U of Klenow fragment, 50 uCi (2.8 uM) each of d[α-^{32}P]ATP, d[α-^{32}P]CTP and unlabelled dATP and dGTP (each at 36 uM). The reaction was chased for 5 min at 37°C with 2 ul of unlabelled dNTP solution (0.25 mM dNTPs) and the Klenow enzyme subsequently inactivated by heating at 70°C for 10 min. The extended product was digested with the appropriate restriction enzyme(s), as described in the text and Figure legends, for 4 - 10 hours at 37°C (Section 2.25) and the resulting fragments were fractionated by electrophoresis on a 5% non-denaturing polyacrylamide gel (Section 2.26.2(a)). The radioactive DNA fragment to be used as a probe was located, isolated and extracted from the gel (Section 2.27.2), concentrated by ethanol precipitation (Section 2.24.3) and resupended in 100 - 200 ul of TE.

2.30.3 Preparation of radioactive DNA probes by kinasing oligonucleotides.

Radioactive DNA probes for use in detecting mutants created by oligonucleotide site-directed mutagenesis (Section 2.31), were made by labelling the 5'-end of the oligonucleotide using polynucleotide kinase. Oligonucleotide (50 ng) was kinased in a 10 ul reaction containing 1x TM (10 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$), 1 mM dithiothrietol, 50 uCi of r[γ-^{32}P]ATP (final concentration 2.5 uM) and 2 U of polynucleotide kinase. Incubation was at 37°C for 1 hour, and formamide loading buffer (5 ul) added and the solution boiled for 5 min. The sample was loaded onto a 20% polyacrylamide gel (Section 2.26.2(a)) and electrophoresis was carried out, as described in Section 2.26.2(a). The gel was autoradiographed (Section 2.26.3) and the radioactive oligonucleotide was isolated from the gel (Section 2.27.2). The oligonucleotide was eluted from the gel slice by incubation at 65°C for 2 - 10 hours in TE.

2.31 OLIGONUCLEOTIDE SITE-DIRECTED MUTAGENESIS.

The method used for oligonucleotide site-directed mutagenesis of M13 cloned DNA was derived from the procedures of Zoller and Smith (1982; 1984). The oligonucleotides used in this work are described in Section 2.5. Before use in the mutagenesis reaction, the oligonucleotides were tested for their ability to hybridize

specifically to the region of interest by using the oligonucleotide as a primer in a sequencing reaction (Section 2.29).

2.31.1 Kinasing the oligonucleotide.

Since the synthetic oligonucleotides did not contain a 5'-phosphate group, it was necessary to add a phosphate onto the 5'-end using polynucleotide kinase, in order to allow ligation of the 5'-end of the oligonucleotide to the 3'-end of the extended strand in the extension-ligation reaction (Section 2.31.2). The oligonucleotide (50 ng) was kinased as described in Section 2.30.2, except for the substitution of 1 mM rATP (final concentration) for r[γ-^{32}P]ATP in the kinasing reaction. Incubation was at 37^{o}C for 1 hour. The reaction was stopped by the addition of 10 mM EDTA (pH 8.0) and heat inactivation of the enzyme at 70^{o}C for 10 min.

2.31.2 Extension-ligation reaction.

M13 single stranded DNA (200 ng, 2 ul) to be mutagenised was mixed with 10 ng of the kinased oligonucleotide containing the appropriate mutation (Section 2.31.1), and 2 ng of cold kinased M13 universal sequencing primer, in a volume of 15 ul containing 5 ul of TM (100 mM Tris-HCl, pH 7.5, 100 mM $MgCl_2$) and 5 ul 200 mM NaCl. The mixture was heated to 65^{o}C for 5 min, cooled slowly to room temperature and placed at 4^{o}C for 15 min.

After the annealing step, 5 ul of dNTP solution (0.5 mM of each dNTP), 5 ul of 10 mM rATP, 17 ul H_2O, 2 U of Klenow fragment and 1 U of T4 DNA ligase were added to the annealed DNA and the extension/ligation reaction was allowed to proceed at room temperature for 4 hours.

2.31.3 Transfection.

JM101 competent cells were prepared (Section 2.28.4(b)) and 2 - 20 ul of the extended-ligated DNA was transfected into the cells as described in Section 2.28.4(b), but without the addition of IPTG or BCIG to the agar before plating.

2.31.4 Identification of mutated M13 clones.

Plaques obtained after the transfection of the extended-ligated DNA, and non-mutated control plaques, were spotted onto a master plate and a replica plate, seeded with a lawn of JM101 and the plates incubated overnight at 37°C. The phage were then tested for the presence of the desired mutation, directed by the oligonucleotide, by plaque hybridization using the relevant oligonucleotide as a probe, as described below.

The solvent TMACl (tetramethylammonium chloride) was used for the stringent washing of filters since TMACl eliminates the preferential melting of AT versus GC base pairs (Melchior and von Hippel, 1973; Orosz and Wetmur, 1977), and thus, the temperature for stringent washing (in 3M TMACl) is based solely on the length of the probe (Ullrich *et al.*, 1984; Wood *et al.*, 1985).

2.31.4(a) Transfer of plaques to nitrocellulose.

A nitrocellulose filter was placed on the replica plate containing the plaques to be tested and left for 5 - 15 min. The filter and the agar plate were marked with asymetric dots using a needle (to aid in the alignment of the filter and the plate for the identification of mutants). The filter was then removed and allowed to dry, DNA-side up, at room temperature for 30 min, after which it was baked at 80°C under vacuum for 2 hours. The filter was then washed in 10 mM Tris-HCl (pH 8.0) by placing the filter on the surface of the solution for 1 min and then submerging it for 5 min. This procedure helped to reduce non-specific (background) hybridization.

2.31.4(b) Hybridization.

Pre-hybridization was performed in petri dishes at 37°C overnight in 20 ml of the following solution : 6x NET (0.9 M NaCl, 90 mM Tris-HCl, pH 7.6, 6 mM EDTA); 5x Denhardts solution; 0.5% (v/v) Nonidet P40; and 100 ug/ml of sonicated and denatured calf thymus DNA. The calf thymus DNA was boiled for 5 min and then snap-chilled before addition to the rest of the solution. After pre-hybridization, the solution was removed and 20 ml of the same solution containing the appropriate ^{32}P labelled oligonucleotide (Section 2.30.3) was added. Hybridization was carried out at 37°C for 12 - 24 hours.

2.31.4(c) Washing.

After hybridization, the filter was washed twice, non-stringently, in 100 ml of 6x SSC for 10 min at room temperature. The filters were then rinsed in TMACl wash solution (3M TMACl, 2 mM EDTA, 0.05 M Tris-HCl, pH 8.0, 1% SDS) at room temperature and placed on a piece of Whatman 3MM paper, on which assymetrical marks had been placed using radioactive ink (to allow the orientation of the autoradiograph with the filter). After autoradiography overnight, all plaques including the unmutated controls showed hybridization to the probe. To identify mutant phage, the filters were washed in TMACl wash solution for 1 hour at the temperature specified by Wood *et al.* (1985) to allow discrimination between the desired mutant and the wildtype. The filter was autoradiographed as above. After the stringent wash, 2% - 10% of the plaques hybridized with the probe. Plaques which hybridized with the radioactive oligonucleotide under these stringent wash conditions, were then tooth-picked from the master plate and purified by streaking for single plaques.

2.31.4(d) Confirmation of the mutation by DNA sequencing.

To confirm that the phage identified by plaque hybridization contained the correct mutation, M13 single-stranded DNA was prepared (Section 2.28.4(e)), and the DNA sequence was determined (Section 2.29) using either the M13 universal sequencing primer or another oligonucleotide, which would anneal to the M13 clone at a position upstream from the mutation. In particular, the DNA sequence complementary to the entire oligonucleotide was confirmed to be as expected. If the primer to be used in the sequencing reaction had a mismatch with the DNA sequence, to which it was to be annealed, the hybridization was carried out at 65^{o}C, cooled slowly to room temperature and was then placed at 4^{o}C for 15 min, to obtain better annealing.

2.32 RNA ANALYSIS.

All procedures for analysis of RNA required care to be taken in avoiding ribonucleases. Gloves were worn at all times and all glassware was sterilised in a 110^{0}C oven overnight. All other equipment was autoclaved or immersed in 1 M KOH for 15 min and rinsed thoroughly with sterile glass-distilled water.

2.32.1 RNA preparation.

This method was adapted from a procedure by Court *et al.* (1980) and is essentially as described by Kalionis *et al.*, (1986b). Cultures were grown in L broth or minimal media supplemented with antibiotics with aeration at 30^{o}C to A_{600}= 0.8. For heat-induced cultures of 186 *cIts* lysogens, the cultures were incubated with aeration at 30^{o}C to A_{600}=0.8 and heat induced by transfer to a 39^{o}C water bath. Aliquots of 10 ml were taken at the times indicated in the text and Figure legends, placed into 50 ml polypropylene tubes and transferred immediately to ice. Sodium azide (NaN_3) was added to a final concentration of 0.02 M to stop cell metabolism, and the aliquots were kept chilled on ice until all time samples were collected. Cells were collected by centrifugation (9,000 rpm, 10 min, 4^{o}C, JA20 rotor) and resuspended in 2 ml of freshly prepared lysis buffer (10 mM Tris-HCl, pH 7.5, 1 mM EDTA, 10 mM NaN_3, 4 mg/ml lysozyme). Lysis was accomplished by freezing the samples in a dry ice/ethanol bath and then placing the tubes immediately into a room temperature water bath and allowing the lysate to thaw for 10 min. The freeze/thaw cycle was carried out two further times and SDS was then added to a final concentration of 0.2% to ensure complete lysis and to inhibit the action of RNases. The tubes were then incubated at 45^{o}C for 3 min. Samples were extracted twice with an equal volume of phenol equilibrated with RNA buffer (20 mM Na acetate, pH 5.2, 20 mM KCl, 10 mM $MgCl_2$) and the phenol phases were washed once with RNA buffer. The aqueous phases were pooled and nucleic acids precipitated by addition of one tenth volume of 3 M Na acetate (pH 4.6) and 2.5 volumes RNase free ethanol. Tubes were placed at -20^{o}C for 30 min and nucleic acids were collected by centrifugation (18,000 rpm, 20 min, 4^{0}C, JA20 rotor). The pellets were rinsed with RNase free ethanol, dried *in vacuo* and finally redissolved in 4.5 ml of 0.1 mM EDTA (pH 8.0).

2.32.2 Removal of DNA from RNA preparations.

A simple method to remove contaminating DNA (and residual protein) was based on the procedure of Glisin *et al.* (1974) and relies on the high buoyant density of RNA, which allows it to pellet in CsCl solutions whereas both DNA and protein have lower buoyant densities and remain in solution.

The RNA sample (in 4.5 ml of 0.1 mM EDTA, pH 8.0) (Section 2.32.1) was mixed with 4.5 ml of 7.5 M CsCl and the resulting solution carefully overlayed onto a 2.5 ml pad of 5.2 M CsCl in a 10 ml polyallomer tube. After centrifugation (30,000 rpm, 20°C, 16 hours, SW41 rotor), the supernatant (10 ml) was carefully removed by aspiration and the tubes were cut below the level of the CsCl pad with a sterile scalpel blade. The remaining supernatant was removed and the gelatinous pellet was dissolved in 0.1 mM EDTA (pH 8.0). The RNA was ethanol precipitated twice (with RNase free ethanol) and resuspended in 0.1 mM EDTA (pH 8.0).

RNA concentrations were determined using a Varian superscan spectrophotometer by measuring the absorbance over the range 190 - 340 nm (A_{260}= 1.0 represents a concentration of 40 ug/ml). RNA was stored at -80°C until required. Yields of DNA-free RNA after centrifugation through the CsCl pad were generally 0.5 - 2.5 mg/10 ml aliquot of cell culture.

2.32.3 Glyoxylation of RNA samples.

RNA samples (Section 2.32.2) were denatured with 1 M de-ionised glyoxal in 10 - 20 ul of 10 mM Na phosphate, pH 6.5, 0.1 mM EDTA at 50°C for 30 - 45 min. Dimethylsulphoxide (DMSO) was omitted from the glyoxylation procedure of McMaster and Carmichael (1977), because of excessive breakdown of RNA even when redistilled DMSO was used (B. Kalionis, personal communication). This did not affect the denaturation process, as judged by the absence of any change in the mobility of molecular weight markers (B. Kalionis, personal communication).

2.32.4 Slot blot analysis of RNA.

2.32.4(a) Slot blot procedure.

RNA slot blots were performed using the procedure of Thomas (1983). RNA was denatured using glyoxal (Section 2.32.3) and then diluted to a volume of 50 ul in 6x SSC (in 10 mM Na phosphate, pH 6.5, 0.1 mM EDTA). RNA was loaded onto nitrocellulose (which had been soaked in 20x SSC for 30 min) by suction using a Schleicher and Schuell minifold II apparatus. Each loading position was pre-washed using 20x SSC and after the RNA solution loaded, was washed through twice with 20x SSC. The filter was then dried for 30 min at room temperature, RNA-side up, and baked at 80^{o}C under vacuum for 2 hours. To remove glyoxal, 10 mM Tris-HCl (pH 8.0) at 100^{o}C was added to the filters and they were agitated slowly until the buffer reached room temperature. To reduce non-specific hybridization, the filters were floated, RNA-side up, in 6x SSC for 1 min and then submerged for 5 min

2.32.4(b) Prehybridization, hybridization and washing.

Pre-hybridization and hybridization conditions of ^{32}P-DNA probes to nitrocellulose-bound RNA, were as described by Thomas (1980). Pre-hybridization was at 42^{o}C for 12 - 20 hours in plastic bags in 60% formamide, 6x SSC, 6x Denhardts solution, 0.06 M Na phosphate, pH 6.5 and 0.1 mg/ml sonicated and denatured calf thymus DNA (boiled for 5 min then snap-chilled before addition to the pre-chilled pre-hybridization mix). Hybridization was at 42^{o}C for 18 - 24 hours in a solution of 60% formamide, 6x SSC, 1x Denhardts solution, 0.025 M Na phosphate and 0.1 mg/ml sonicated and denatured calf thymus DNA, which contained the radioactive probe ($1x10^{6}$ - $3x10^{7}$ cpm). Probes used for RNA analysis were prepared by primer extension on M13 single stranded clones (Section 2.30.1) with the same polarity as the RNA. Probes were heat-denatured at 100^{o}C for 5 min, snap-chilled and diluted into pre-chilled hybridization buffer. After hybridization, the hybridization buffer was removed and the filters were washed four times for 5 min at room temperature in 250 ml 2x SSC, 0.1% SDS and then twice at 60^{o}C in 250 ml 0.1x SSC, 0.1% SDS. Filters were covered with plastic film (Vitafilm) and autoradiographed at -80^{o}C with an intensifying screen.

2.32.4(c) Quantitation of hybridization to RNA.

After autoradiography of nitrocellulose filters (Section 2.32.4(b)), the intensity of the slots was quantitated using a Zeinch scanning laser densitometer. The level of hybridization was represented by the total peak area (above background) obtained by laser densitometric scanning of autoradiographs of RNA slot blot hybridized with a specific probe. The specific activity of each probe was normalized by hybridization to a sample of RNA of known concentration to allow comparison of the amount of RNA hybridizing to one probe relative to another probe.

2.33 DETERMINATION OF 5'-ENDS OF RNA TRANSCRIPTS BY PRIMER EXTENSION ANALYSIS.

This procedure was based on the method described by McKnight *et al.* (1981) with modifications described by Dodd *et al.* (1990). Radioactive DNA restriction fragments to be used as primers were prepared, as described in Section 2.30.2.

The radioactive DNA primer and 10 ug of *in vivo* RNA (Section 2.32.2) were precipitated with ethanol (Section 2.24.3), redissolved in 10 ul of 200 mM NaCl, 10 mM Tris-HCl (pH 8.3), and then heated at 100°C for 3 min. RNA and primer were annealed at 60°C for 1.5 hours, the tubes allowed to cool to room temperature and 24 ul of reaction mix was added to give a final concentration of 10 mM Tris-HCl (pH 8.3), 10 mM $MgCl_2$, 10 mM dithiothreitol, 500 uM of each of the four dNTPs, 60 mM NaCl and 15 units of AMV reverse transcriptase was added to the reaction mix and the tubes were incubated at 42°C for one hour. Nucleic acids were precipitated with RNase free ethanol (as described in Section 2.24.3). The pellet was washed with 70% (v/v) and 95% ethanol and dried *in vacuo*. Nucleic acids were redissolved in 2 ul of 0.1 mM EDTA, an equal volume of formamide loading buffer added to the samples and the samples were heated to 100°C for 3 min. Electrophoresis was carried out on a 6% denaturing polyacrylamide gel (Section 2.26.2). The gel was fixed, as described in Section 2.26.2, and autoradiographed (Section 2.26.3).

2.34 COMPUTER ASSISTED ANALYSIS.

The database management system of Staden (1980) was used for the storage, management and general analysis of DNA sequences.

2.34.1 Prediction of protein coding regions (genes).

The method of assessing DNA sequences for their protein coding potential based on codon usage (Gribskov *et al.*, 1984), was used for analysis of DNA sequence data. To analyse the protein coding potential of any region in phage 186, the program CODONPREF was used. CODONPREF is a sequence analysis package (version 6.0) from the Genetics Computer Group, University of Wisconsin (Devereux *et al.*, 1984). CODONPREF uses the technique of Gribskov *et al.* (1984), which is similar to that of Staden and McLachlan (1982), and the program is based on the concept of synonymous codons and compares the codon usage of any sequence with that from a group of highly expressed genes in *E. coli* (Grantham *et al.*, 1981). Codon usage of all possible reading frames was compared (in window lengths of 25 codons) to the codon usage of the standard. Any frame, which had a similar codon usage pattern to that of the standard, was predicted to be a protein coding frame. Calculation of the score was described in detail by Gribskov *et al.* (1984).

2.34.2 Searching for signals.

The program SCAN (Kalionis et al., 1986a) was used to predict the location of potential promoter sequences for transcription, and ribosome-binding sites. This program uses a weight matrix to evaluate each section of the DNA sequence by the same principal used in the program of Staden (1984).

Promoters were predicted using a weight matrix composed of the frequency of each base at each position of the 112 *E. coli* promoters compiled by Hawley and McClure (1983), with variable spacings (15 - 21 bp) between the -10 and -35 positions. The following threshold scores were used for various spacings between the -10 and -35 regions : 66.0 (17 bp); 66.8 (16 and 18 bp); 67.8 (15 and 19 bp); 68.3 (20 bp); and 68.8 (21 bp) (Kalionis et al., 1986a). [A spacing of 17 between the -35 and -10 regions has been shown to be the optimal for promoter activity (Aoyama *et al.*, 1983).] Using a method similar to

the method used in this study Mulligan *et al.* (1984) showed that the degree of homology of a promoter to known promoters was related to the strength of that promoter *in vitro*.

Potential ribosome-binding sites were predicted using a weight matrix based on the rules of Stormo *et al.* (1982). These rules are based on the degree of homology to the Shine-Dalgarno sequence (Shine and Dalgarno, 1974).

2.34.3 Searching for secondary structure.

Potential stem-loop structures were searched for using the program COMSTR (A.V. Sivaprasad, personal communication). COMSTR displayed the structures in a 2-D form and calculated an approximate ΔG value for the stability of the structure using the rules of Tinoco *et al.* (1973) as modified by Steger *et al.* (1984). Dot matrix analysis was also used for the detection of inverted repeat structures, since COMSTR does not detect secondary structures, which have assymetrical bulges in the stem (A.V. Sivaprasad, personal communication).

The SCAN program was written by I. Dodd (this laboratory) and COMSTR was written by A.V. Sivaprasad (this laboratory). All computer analysis was performed on a Vax 11-785 computer.

RESULTS: SECTION I

CHAPTER 3

DNA SEQUENCE OF A LATE PROMOTER REGION

CHAPTER 3. DNA SEQUENCE OF A LATE PROMOTER REGION.

3.1 INTRODUCTION.

To identify the phage functions involved in the control of late gene transcription in 186, the approach was to clone a late promoter so as to direct transcription of a reporter function and then determine the functions required to achieve transcription from the late promoter. In order to locate a 186 late promoter, the approach taken was to determine the DNA sequence of a region expected to encode a 186 late promoter, and to use computer assisted analysis to determine the coding capacity and to locate potential transcriptional and translational control signals in the DNA sequence.

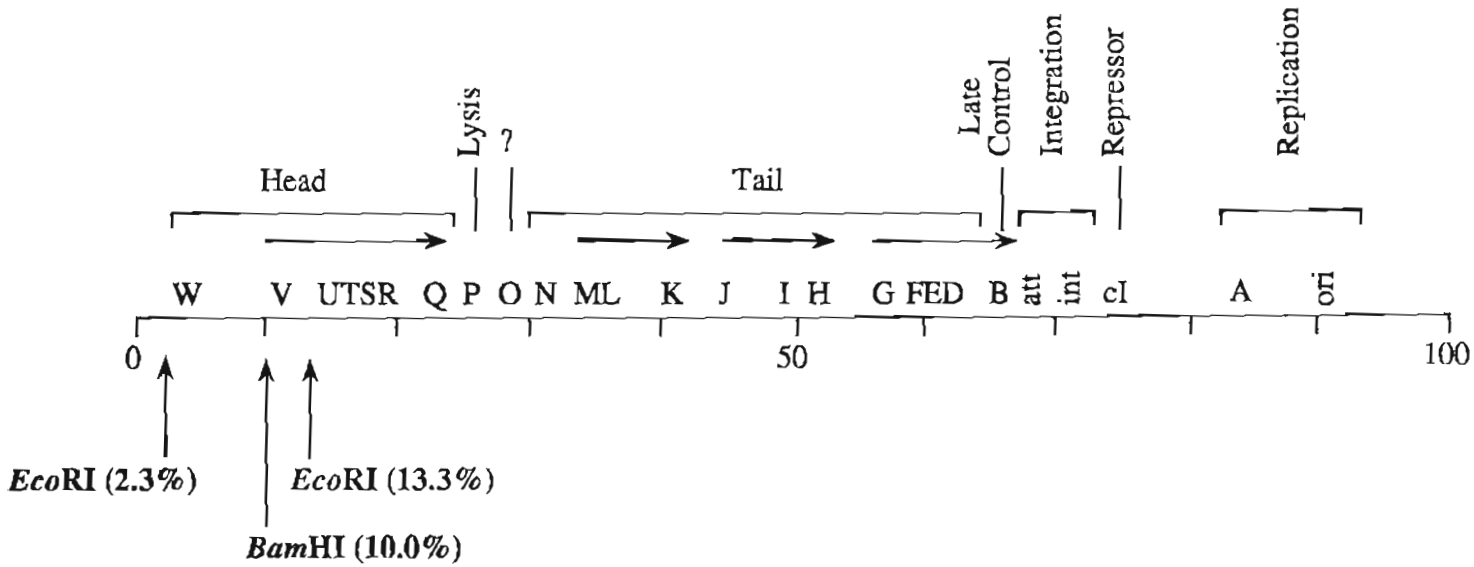
Polarity studies had indicated that the head genes *VUTSRQ* were rightwardly transcribed as a unit (Hocking and Egan, 1982c). The single amber mutation in gene *V* (*Vam*38) has been mapped to the *Bam*HI-*Eco*RI (10.0% to 13.3%) region by marker rescue, as shown in Figure 3.1 (Finnegan and Egan, 1979). Amber mutations in the adjacent head gene *W* mapped to the *Eco*RI-*Bam*HI (2.3% to 10.0%) region (Figure 3.1; Finnegan and Egan, 1979), but did not exhibit polar effects on genes *VUTSRQ* (Hocking and Egan, 1982c). This indicated that genes *W* and *V* were probably represented on different transcripts and that the promoter for the rightward *VUTSRQ* transcription unit was located between genes *W* and *V*. Since the *Vam*38 mutation mapped to the *Bam*HI-*Eco*RI (10.0% to 13.3%) region, the DNA sequence of this restriction fragment was determined as a first step to locate the promoter for the *VUTSRQ* transcript.

The DNA sequence of one strand of DNA to the left of the *Eco*RI site at 13.3% has been previously determined (Dibbens, 1984). Analysis of this preliminary sequence allowed the prediction of the 5'-end of a potential gene, which was identified as the *V* gene upon the determination of the base pair change responsible for the *Vam*38 mutation (Dibbens, 1984). The aim of this chapter was to complete the DNA sequence of the entire *Bam*HI-*Eco*RI (10.0% to 13.3%) region and to use computer assisted analysis to locate potential coding regions and potential control signals in the DNA involved in the expression of these genes.

Figure 3.1 The physical and genetic map of phage 186 showing the location of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region.

(a) Physical and genetic map of phage 186 as determined by Hocking and Egan (1982a). Functions of genes, as determined by Hocking and Egan (1982c,d), are shown above the Figure. The genetic mapping (Hocking and Egan, 1982a) and the physical mapping (Finnegan and Egan, 1979) have been described previously (Section 1.3.1). The location of the *Eco*RI sites at 2.3% and 13.3%, and the *Bam*HI site at 10.0% are as determined by Saint and Egan (1979).

Amber alleles in gene *W* have been mapped to the *Eco*RI-*Bam*HI (2.3% to 10.0%) region, while the single amber allele in gene *V* has been mapped to the *Bam*HI-*Eco*RI (10.0% to 13.3%) region (Finnegan and Egan, 1979).


Lysis
?
Late Control
Integration
Repressor
Replication
Head
Tail
W V UTSR Q P O N ML K J I H G FED B att int cI A ori
0
50
100
EcoRI (2.3%)
EcoRI (13.3%)
BamHI (10.0%)

3.2 RESULTS AND DISCUSSION.

3.2.1 Sequencing strategy.

The *Bam*HI-*Eco*RI (10.0% to 13.3%) region (Figure 3.1) from 186 *cIts* DNA has been previously cloned into the M13 vectors M13mp8 and M13mp9 using the *Bam*HI and *Eco*RI sites, to yield the clones mEC403 and mEC404 respectively (Dibbens, 1984). Clone mEC403 contains the *r*-strand and has been used to determine the DNA sequence 455 bp rightward (with respect to the genetic map) from the *Bam*HI site at 10.0%, while clone mEC404 contains the *l*-strand and has been used to determine the DNA sequence 533 bp leftward from the *Eco*RI site at 13.3% (Dibbens, 1984).

The *Bam*HI-*Eco*RI (10.0% to 13.3%) restiction fragment has been sized by agarose gel electrophoresis (Section 2.26.1) at 1.4 kb (data not shown). The strategy chosen to complete the sequencing of this region was to clone smaller DNA fragments spanning the *Bam*HI-*Eco*RI (10.0% to 13.3%) region into the vector M13mp18, followed by the determination of the DNA sequence of these clones using the modified Sanger dideoxy chain-termination method (Section 2.29). The region of the 186 genome sequenced is presented in Figure 3.1 and the strategy used to determine the DNA sequence of the fragment is shown in Figure 3.3.

Subdigestion of the 1.4 kb *Bam*HI-*Eco*RI (10.0% to 13.3%) restriction fragment from RF DNA (Section 2.22) of the M13 clone mEC403 (Section 2.4.2) with *Alu*I (Section 2.25) and end-labelling with ^{32}P (Section 2.28.2), generated five bands on polyacrylamide gel electrophoresis (Section 2.26.2(a)), sized at 490 bp, 280 bp, 220 bp, 200 bp and a probable doublet band of size 80 bp (Figure 3.2). These DNA fragments were isolated from a non-denaturing polyacrylamide gel (Section 2.26.2(a)), after end-filling of the termini (Section 2.28.2), and cloned into M13mp18 *Sma*I vector (Sections 2.28.1 and 2.28.4) and recombinants in both orientations identified (Section 2.28.4). The DNA sequence generated from these clones (Figure 3.3) allowed the determination of the complete DNA sequence, using the sequence data obtained from mEC403 and mEC404 (Dibbens, 1984), except for an *Alu*I fragment of 79 bp (sequence coordinates 567-647, Figure 3.3), the orientation of which was unknown. The preliminary DNA sequence indicated the presence of a unique *Pst*I site on the *Bam*HI-*Eco*RI (10.0% to 13.3%)

Figure 3.2 Restriction analysis of the *Bam*HI-*Eco*RI (10.0% to 13.3%) fragment from 186 *cIts*.

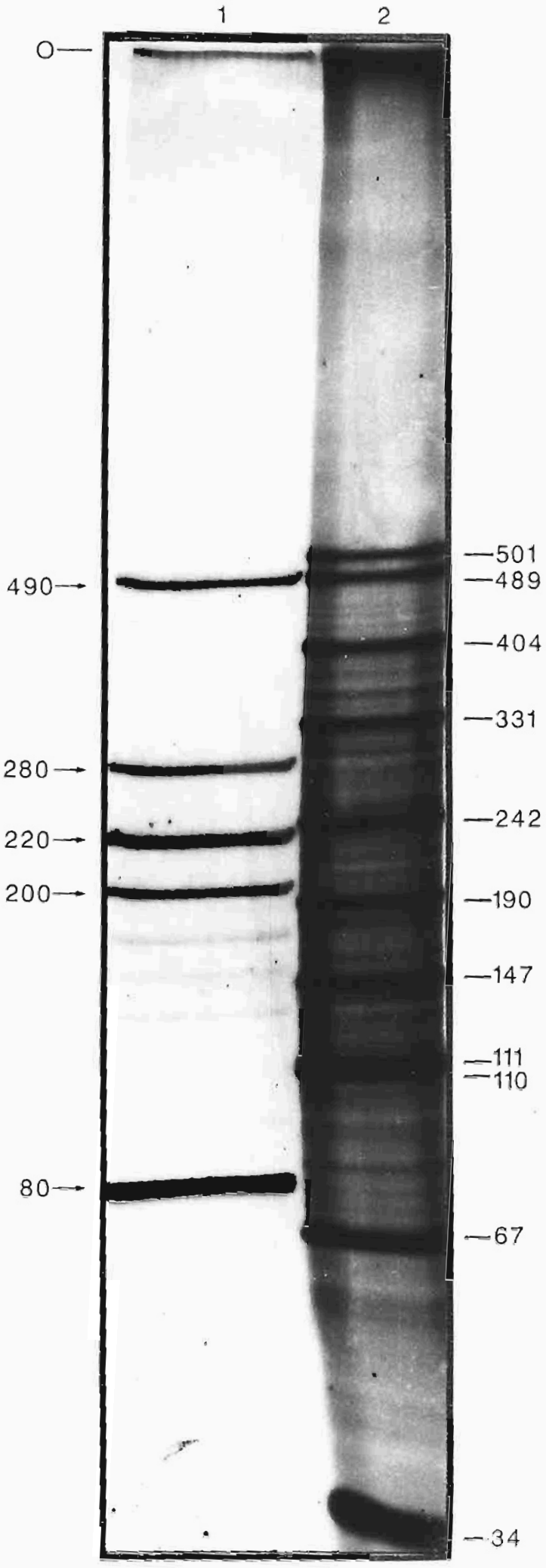
The 1.4 kb *Bam*HI-*Eco*RI (10.0% to 13.3%) restriction fragment from mEC403 RF DNA (Section 2.4.2) was digested with *Alu*I (Section 2.25), the DNA end-labelled (Section 2.28.2) and fractionated on a 5% non-denaturing polyacrylamide gel (Section 2.26.2(a)).

Plasmid pUC19 *Hpa*II DNA fragments (Section 2.11) were end-labelled (Section 2.28.2) and fractionated on the gel to serve as size markers. Electrophoresis and autoradiography were as described in Sections 2.26.2(a) and 2.26.3. The gel was autoradiographed for 6 hours at room temperature.

Gel Tracks

1. The 1.4 kb *Bam*HI-*Eco*RI (10.0% to 13.3%) restriction fragment from mEC403 digested with *Alu*I. The predicted sizes of the resultant restriction fragments are indicated on the left of the Figure.

2. pUC19 *Hpa*II molecular weight markers. The sizes of the fragments are indicated on the right of the Figure.


1
2
O
490
280
220
200
80
501
489
404
331
242
190
147
111
110
67
34

fragment at 12.2% (Figure 3.3) and this restriction site was used to determine the orientation of the DNA sequence generated from the 79 bp *Alu*I restriction fragment. To obtain the DNA sequence leftward from the *Pst*I (12.2%) site, the *Bam*HI-*Eco*RI (10.0% to 13.3%) fragment from mEC403 DNA was subdigested with *Pst*I (Section 2.25) and the DNA fragments resulting from the *Pst*I digestion force cloned (Section 2.28.4) into the vector M13mp18 restricted with *Pst*I and *Bam*HI (Section 2.28.1). Determination of the DNA sequence of a *Bam*HI-*Pst*I (10.0% to 12.2%) clone allowed the orientation of the 79 bp *Alu*I (11.9% to 12.7%) fragment to be determined (Figure 3.3).

Completion of the DNA sequence of both strands of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region required sequencing over the restriction sites used to generate the DNA sequence, and for this purpose the restriction sites for *Xmn*I (11.3%) and *Oxa*NI (13.1%) were utilized (Figure 3.3). Subdigestion of the *Bam*HI-*Eco*RI (10.0% to 13.3%) fragment with either of these enzymes (Section 2.25) generated fragments sized by agarose gel electrophoresis (Section 2.26.1) at 0.9 kb and 0.4 kb (data not shown), as predicted from the DNA sequence. The *Bam*HI-*Eco*RI (10.0% to 13.3%) restriction fragment from mEC403 RF DNA was digested with *Xmn*I (Section 2.25), the termini end-filled (Section 2.28.2) and the DNA fragments (isolated from a 1.5% agarose gel; Section 2.27.1(b)) cloned into M13mp18 vector restricted with *Sma*I (Section 2.28.1; Section 2.28.4). This yielded M13 clones containing the *Bam*HI-*Xmn*I (10.0% to 11.3%) restriction fragment in both orientations and clones containing the *Xmn*I-*Eco*RI (11.3% to 13.3%) restriction fragment also in both orientations. The same procedure was also performed after digestion of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region with *Oxa*NI. Determination of the DNA sequence rightward and leftward from both the *Oxa*NI and *Xmn*I sites allowed the completion of the DNA sequence (Figure 3.3).

The DNA sequence of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region is presented in Figure 3.4(a), together with the results of the computer assisted analysis showing potential genes (Section 3.2.2(a)) and potential transcription control signals (Section 3.2.5). The complete sequence is 1362 bp in length and sequence coordinates are from the *Bam*HI site mapped at 10.0% (Saint and Egan, 1979). Chromosomal coordinates of restriction sites are calculated from the *Bam*HI site at 10.0% (Saint and Egan, 1979), with the assumption that 300 bp is approximately equal to 1% of the chromosome. The *Eco*RI site has been

Figure 3.3 Sequencing strategy of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region from 186 *cIts*.

(a) The physical and genetic map of phage 186 indicating the *Bam*HI-*Eco*RI (10.0% to 13.3%) fragment sequenced in this work. Functions of genes, as determined by Hocking and Egan (1982c,d), are shown above the Figure. The genetic mapping (Hocking and Egan, 1982a) and the physical mapping (Finnegan and Egan, 1979) have been described previously (Section 1.3.1). The location of the *Bam*HI-*Eco*RI (10.0% to 13.3%) restriction fragment is as determined by Saint and Egan (1979).

(b) The *Bam*HI-*Eco*RI (10.0% to 13.3%) region expanded to show the sequencing strategy employed. The relevant restriction sites used to determine the nucleotide sequence of the region are shown. The sequence coordinates of the restriction sites refer to the first base of the site on the *l*-strand and the chromosomal coordinates are calculated from the *Bam*HI site at 10.0% (Saint and Egan, 1979), with the assumption that 300 bp = 1% of the chromosome. The *Eco*RI site remains designated at 13.3% (Saint and Egan, 1979) until DNA is sequenced to the *cos* site at 0% and final values calculated.

The arrows above and below the map represent gel readings used to generate the DNA sequence of the *Bam*HI-*Eco*RI (10.0% to 13.3%) fragment. The length of the arrows indicates the extent of sequence generated from each of the restriction sites used. The rightward arrows above the map represent gel readings from M13 clones used to generate the DNA sequence of the *l*-strand and the leftward arrows below the map represent gel readings used to generate the DNA sequence of the *r*-strand.

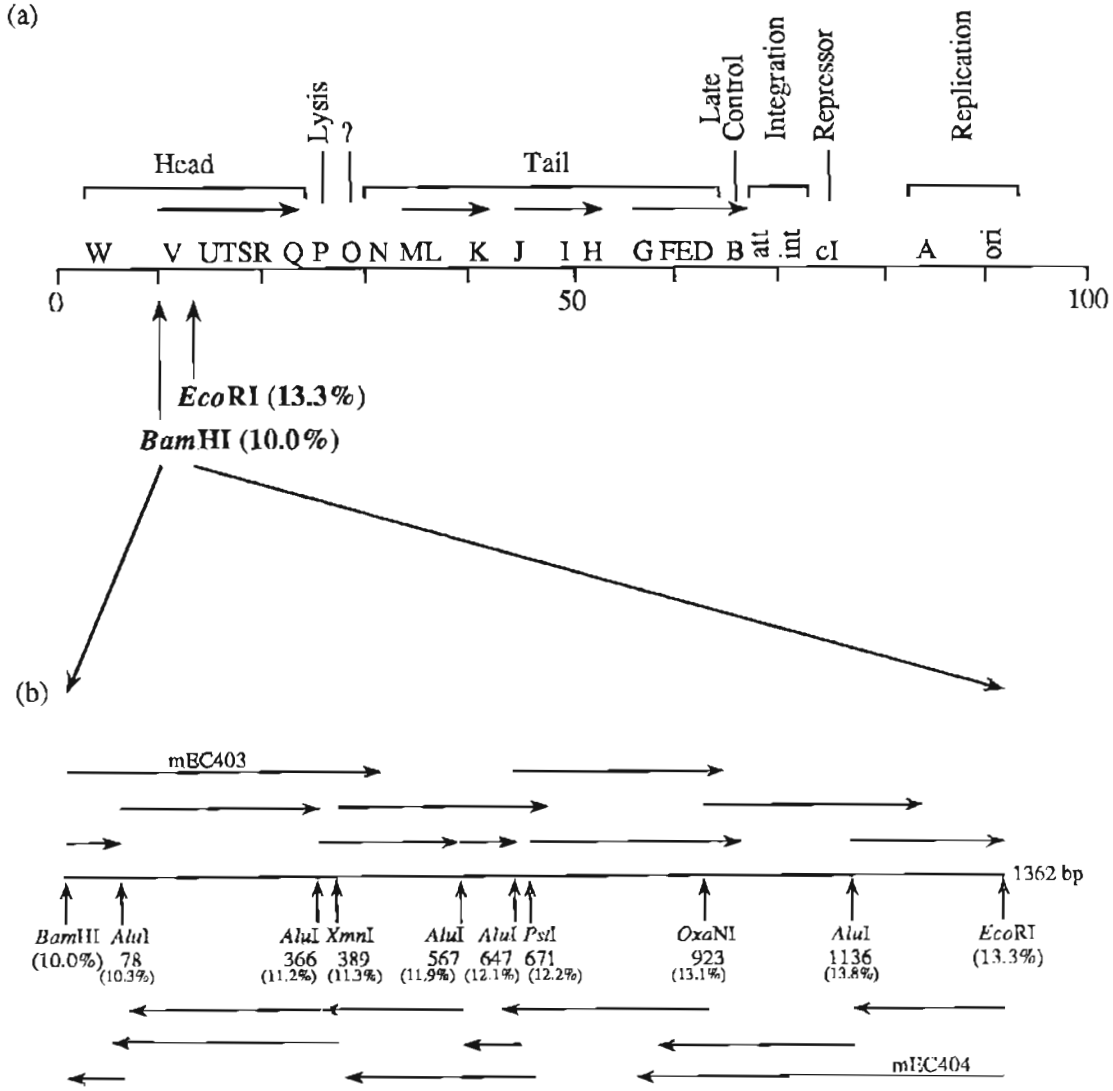

(a)
Lysis
?
Late Control
Integration
Repressor
Replication
Head
Tail
W V UTSR Q P O N ML K J I H G FED B att int cI A ori
0
50
100
EcoRI (13.3%)
BamHI (10.0%)
(b)
mEC403
1362 bp
BamHI (10.0%)
AluI 78 (10.3%)
AluI 366 (11.2%)
XmnI 389 (11.3%)
AluI 567 (11.9%)
AluI 647 (12.1%)
PstI 671 (12.2%)
OxaNI 923 (13.1%)
AluI 1136 (13.8%)
EcoRI (13.3%)
mEC404

Figure 3.4 DNA sequence of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region from 186 *cIts*.

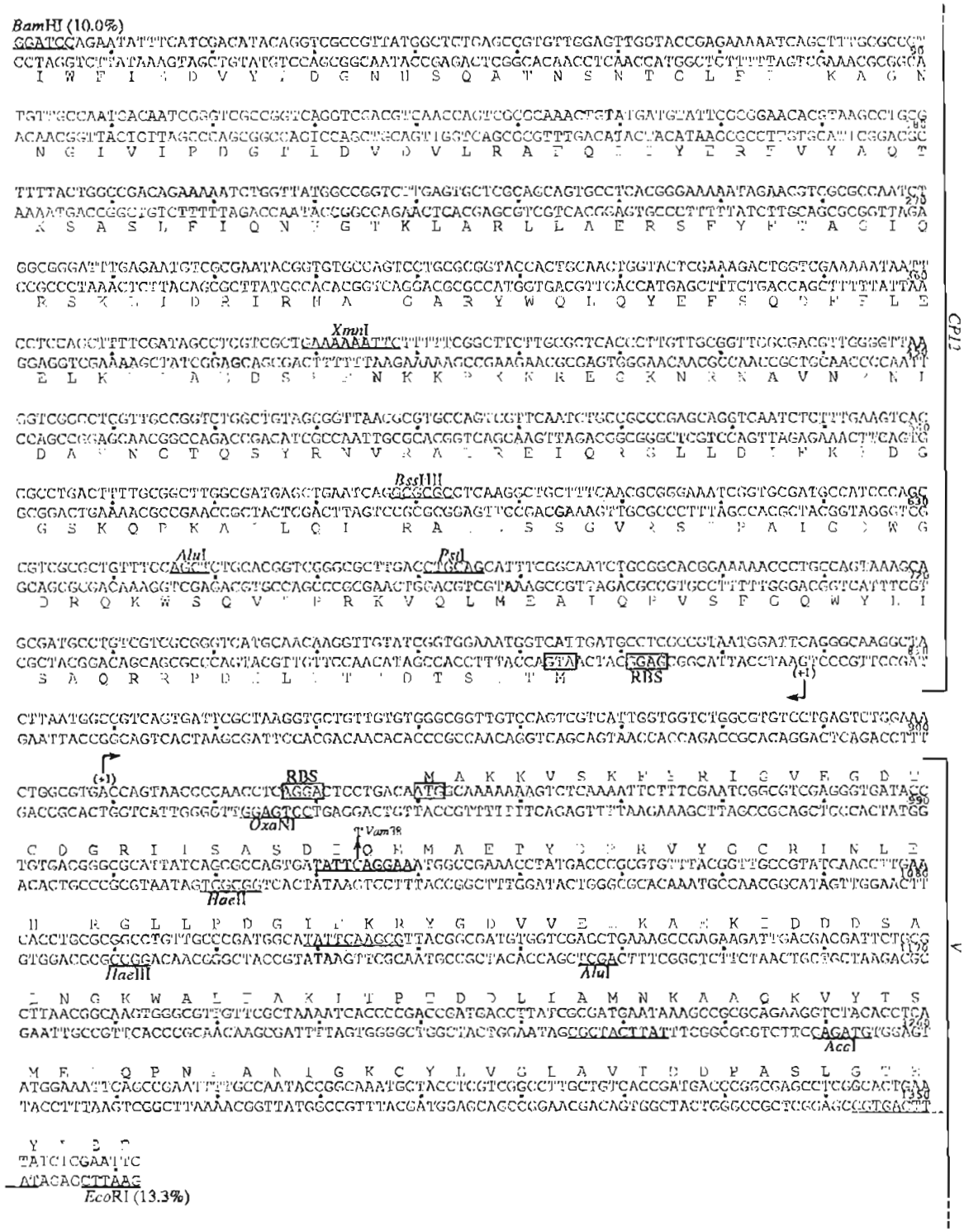
(a) The nucleotide sequence of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region from 186 *cIts* and the amino acid sequence (in single letter code) of predicted proteins encoded on the fragment are shown. The upper strand represents the *l*-strand (Kalionis and Egan, 1981) and numbering is from the first base of the *Bam*HI site at 10.0% (Saint and Egan, 1979). Relevant restriction sites are indicated.

Gene *CP12* is translated from right to left. Gene *V* is translated from left to right. Predicted ribosome-binding sites (RBS) are boxed as are the predicted initation codons (Section 3.2.2(a)). The base pair change responsible for the *Vam*38 mutation is at sequnce coordinate 1024 (Dibbens, 1984) and is shown. Predicted transcription initiation sites (Section 3.2.5(b)) are shown (+1). Underlined regions show homology to the *B box* sequence (Section 3.2.7; Kalionis *et al.*, 1986b).

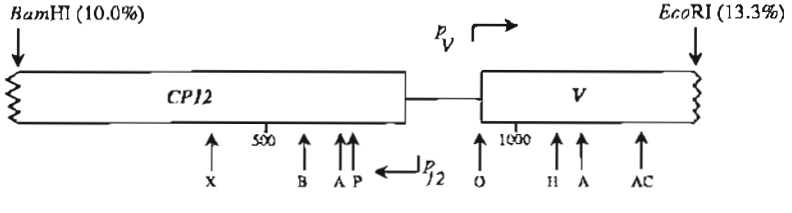
The sequence coordinates of the restriction sites described in the text refer to the first base of the site on the *l*-strand and the chromosomal coordinates are calculated from the *Bam*HI site at 10.0% (Saint and Egan, 1979), with the assumption that 300 bp = 1% of the chromosome. The *Eco*RI site remains designated at 13.3% (Saint and Egan, 1979) until DNA is sequenced to the *cos* site at 0% and final values calculated.

(b) Diagrammatic representation of the DNA sequence of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region, showing genes *V* and *CP12* (which are only partially represented on the fragment) and the late promoters p_V and p_{12}. Relevant restriction sites are indicated, A-*Alu*I, AC-*Acc*I, B-*Bss*HII, H-*Hae*II, O-*Oxa*NI, P-*Pst*I and X-*Xmn*I.

(a)


BamHI (10.0%)
XmnI
BssHII
AluI
PstI
RBS
(+1)
CPI2
V
OxaRI
HaeII
HaeIII
AluI
AccI
EcoRI (13.3%)

(b)


BamHI (10.0%)
EcoRI (13.3%)
P_V
CPI2
V
590
1090
X
B
A P
P_{I2}
O
H
A
AC

mapped at 13.3% (Saint and Egan, 1979) and will remain assigned to this position until further DNA sequencing to the *cos* site at 0% is completed. Unless otherwise specified, all sequence positions and coordinates refer to those in Figure 3.4(a).

3.2.2 Analysis of the DNA sequence.

3.2.2(a) Protein coding potential.

A very powerful method for the prediction of protein coding regions from the DNA sequence can be made on the basis of open reading-frame length, codon usage (Shepherd, 1981, Fickett, 1982; Staden, 1984b; Gribskov *et al.*, 1984; McLachlan *et al.*, 1984) and the presence of ribosome-binding sites (Shine and Delgarno, 1974; Stormo *et al.*, 1982). The coding potential of open reading-frames of at least 25 codons was analysed using the computer program CODONPREF (Section 2.34.1), which determines the coding potential of a region, based on the method of Gribskov *et al.* (1984), by using the codon frequency standards of highly expressed *E. coli* genes (Grantham *et al.*, 1981). Potential translation initiation sites were located using the program SCAN (Section 2.34.2), which uses a weight matrix employing the ribosome-binding site rules of Stormo *et al.* (1982).

The analysis of the DNA sequence of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region allowed the detection of two extensive divergent open reading-frames, both not completely contained within the *Bam*HI-*Eco*RI (10.0% to 13.3%) region. The rightward open reading-frame extended from sequence coordinates 940-1362 (>141aa; Figure 3.4(a)) and the leftward open reading-frame extended from sequence coordinates 775-1 (>258aa; Figure 3.4(a)). Two ribosome binding sites were detected in the DNA sequence of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region using the SCAN program (Table 3.1), and each of these was associated with the initiating methionine of each of the open reading-frames. The ribosome-binding site at sequence coordinate 927, associated with the rightward open reading-frame, satisfied Stormo rule 6 (Stormo *et al.*, 1982), as did the ribosome-binding site at sequence coordinate 784, associated with the leftward open reading-frame (Table 3.1).

Both open reading-frames were detected by the CODONPREF program (Section 2.34.1) using the codon usage of highly expressed *E. coli* genes as a standard (Grantham *et*

TABLE 3.1

Ribosome-binding sites.

Predicted protein	Ribosome-binding site[a]	Stormo rule[b]	Sequence position[c]
CP12	GAGG-5-ATG	6	775
V	AGGA-9-ATG	6	940

Notes to Table 3.1

a. The numbers between the proposed ribosome-binding site and the initiation codon refer to the number of intervening bases.

b. Stormo *et al.*(1982).

c. Sequence position corresponding to the A residue of the ATG initiation codon.

al., 1981). The codon usage plots for each strand of the DNA sequence of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region employing the *E. coli* codon usage frequencies are shown in Figures 3.5 and 3.6. On the basis of codon usage and the presence of potential ribosome-binding sites, it was concluded that both open reading-frames represented genes. The rightward open reading-frame was designated *CP13* (for Computer Protein initiating near 13%) and the leftward open reading-frame was designated *CP12* (for Computer Protein initiating near 12%).

The base-pair change responsible for the *Vam*38 mutation, a C to T transition at sequence coordinate 1024 (Figure 3.4(a)), has been previously determined (Dibbens, 1984). This mutation created an amber stop codon in the predicted reading-frame of *CP13* and demonstrated that the rightward open reading-frame was that of the *V* gene. The known amber mutants in the *W* gene map to the *Eco*RI-*Bam*HI (2.3% to 10.0%) region (Finnegan and Egan, 1979) and determination of the DNA sequence of this region and the nucleotide changes responsible for the amber mutations in gene *W* will demonstrate if *CP12* is gene *W* or a previously undetected gene. A diagrammatic representation of the nucleotide sequence of the 1.4 kb *Bam*HI-*Eco*RI (10.0% to 13.3%) region is presented in Figure 3.4(b), showing the predicted gene content and transcriptional control signals (Section 3.2.5(b); Section 3.2.6).

3.2.2(b) Properties of the predicted proteins.

Many of the physical properties of a protein can be determined by analysis of the amino acid sequence (as determined by translation of the DNA sequence) and may give an indication as to the functional role of the protein. The properties of the portions of the proteins predicted by translation of the DNA sequence of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region, together with the analysis of the codon usage (Section 3.2.3) are presented in Table 3.2.

(i) DNA-binding motifs

Proteins which exert their effect by interacting with the negatively charged backbone of DNA are often highly basic (e.g. λ Cro 16.7%, λ CII 17.5%). The portions of the translation products of genes *CP12* and *V* present on the *Bam*HI-*Eco*RI (10.0% to 13.3%) region showed a slightly higher percentage of basic amino acids (lys. and arg.;

Figure 3.5 Codon usage plots for the *l*-strand of the DNA sequence of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region from 186 *cIts*.

The codon usage plots for the three reading frames on the *l*-strand of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region are shown using the program CODONPREF (Section 2.34.1).

The DNA sequence of the *l*-strand of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region was analysed with window lengths of 25 codons (Section 2.34.1) using the codon usage standards of highly expressed *E. coli* genes (Grantham *et al.*, 1981). The X-axis gives the sequence position for the centre of each 25 codon window analysed and also shows for each reading frame all methionine codons (large bars above the line), all termination codons (large bars overlapping the line) and the open reading frames associated with these codons. The *V* open reading frame is shown. Below each plot are shown the occurrence of rare codons, indicated by small bars. The score is shown on the Y-axis and is calculated from the frequency of occurrence of each codon (f), the total number of occurrences of its synonymous family (F), and the calculated occurrences of the codon (r) and its synonymous family (R) in a random sequence with the same base composition as that being analysed. The codon preference statistic (p) is given by [(f/r)/(r/R)] and the actual plotted statistic is calculated using the logarithmic derivative (Gribskov *et al.*, 1984). The score indicates the degree to which the codon usage of a section of the DNA sequence is related to that of the chosen standard (in this case the codon usage of *E. coli* genes).

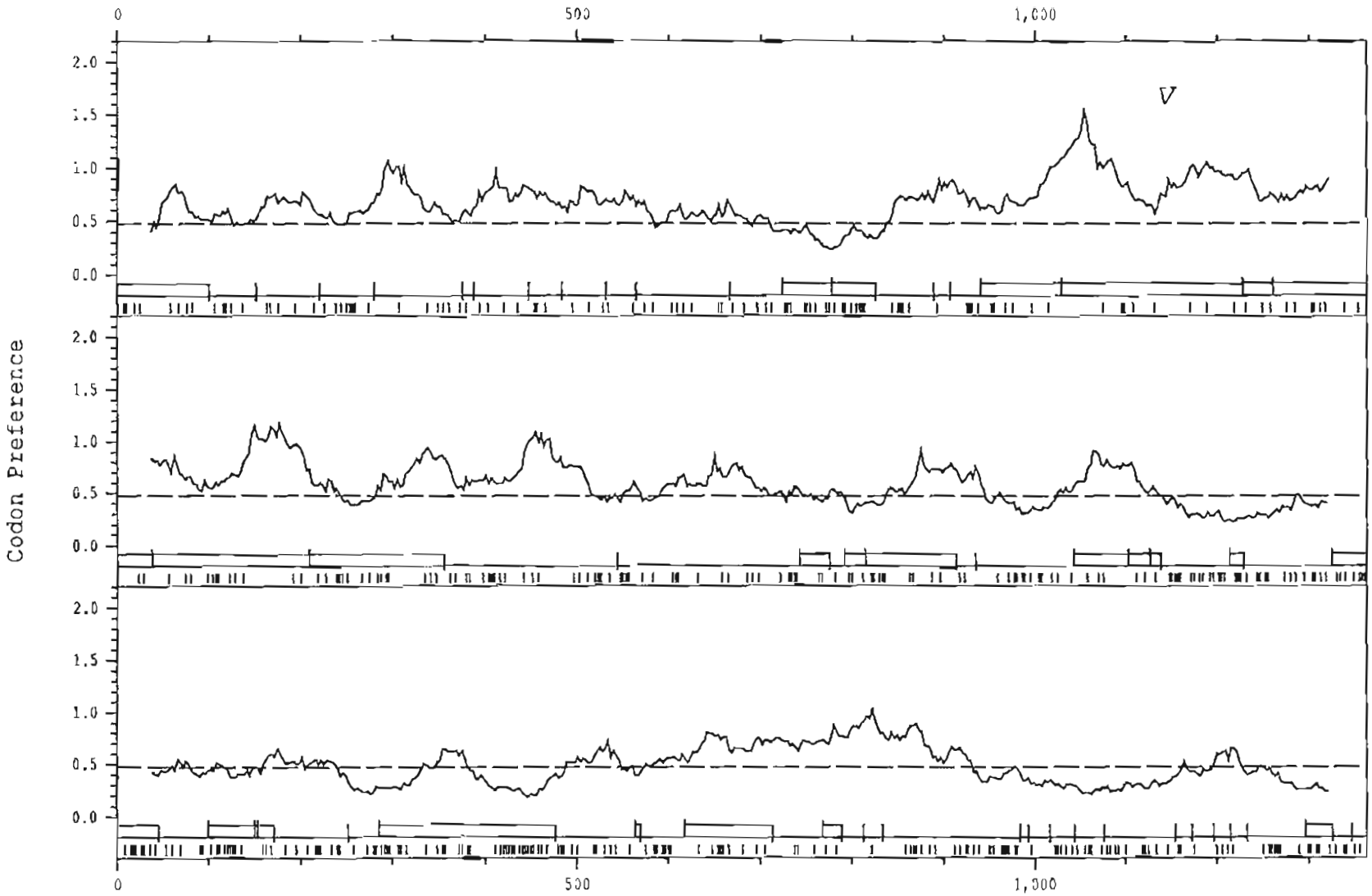

Codon Preference
0
500
1,000
2.0
1.5
1.0
0.5
0.0
V

Figure 3.6 Codon usage plots for the *r*-strand of the DNA sequence of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region from 186 *cIts*.

The codon usage plots for the three reading frames on the *r*-strand of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region are shown using the program CODONPREF (Section 2.34.1).

The DNA sequence of the *r*-strand of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region was analysed with window lengths of 25 codons (Section 2.34.1) using the codon usage standards of highly expressed *E. coli* genes (Grantham *et al.*, 1981). The X-axis gives the sequence position for the centre of each 25 codon window analysed and also shows for each reading frame all methionine codons (large bars above the line), all termination codons (large bars overlapping the line) and the open reading frames associated with these codons. The *CP12* open reading frame is shown. Below each plot are shown the occurrence of rare codons, indicated by small bars. The score is shown on the Y-axis and is calculated from the frequency of occurrence of each codon (f), the total number of occurrences of its synonymous family (F), and the calculated occurrences of the codon (r) and its synonymous family (R) in a random sequence with the same base composition as that being analysed. The codon preference statistic (p) is given by [(f/r)/(r/R)] and the actual plotted statistic is calculated using the logarithmic derivative (Gribskov *et al.*, 1984). The score indicates the degree to which the codon usage of a section of the DNA sequence is related to that of the chosen standard (in this case the codon usage of *E. coli* genes).

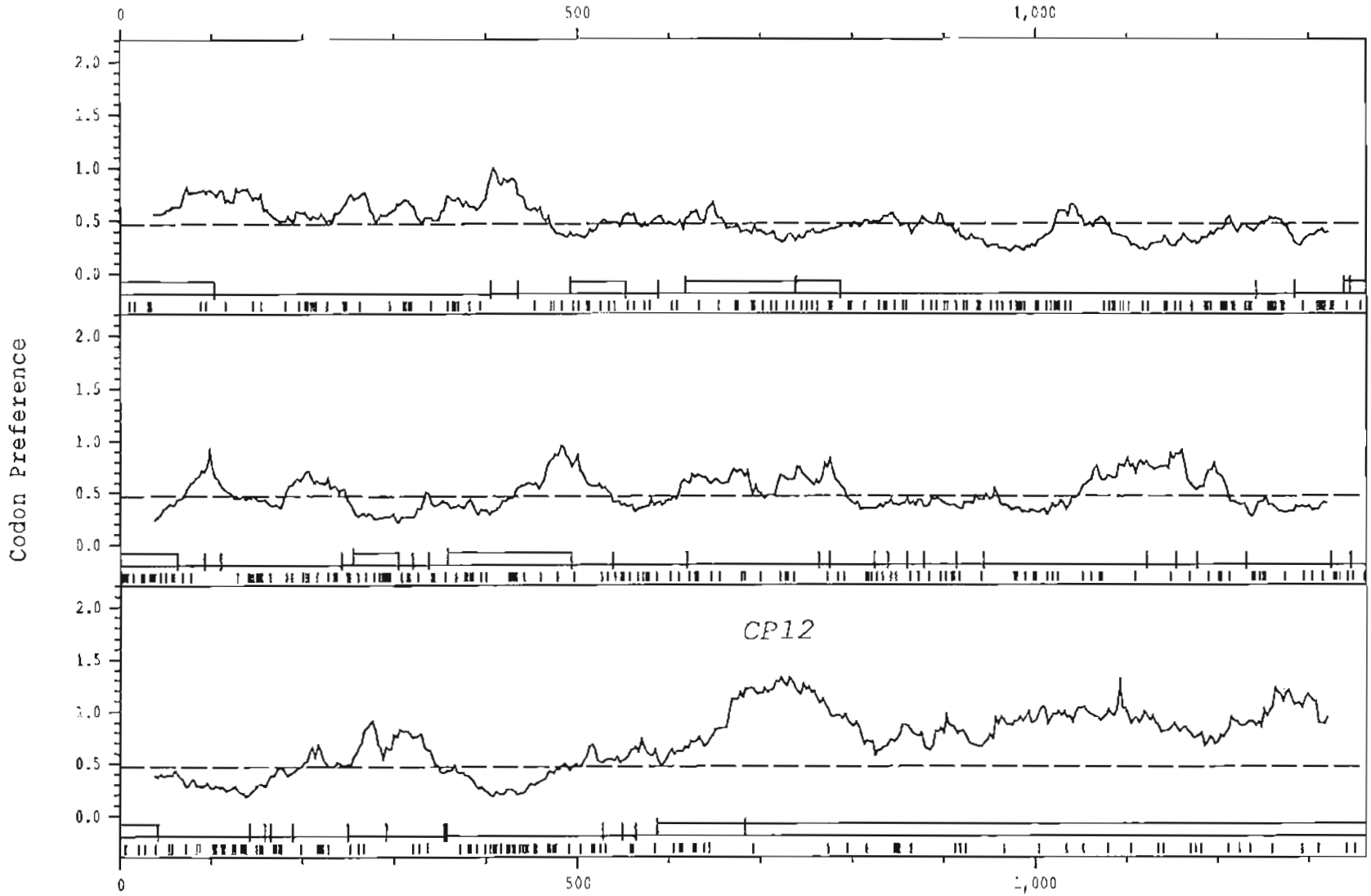

Codon Preference
0
500
1,000
2.0
1.5
1.0
0.5
0.0
CP12

TABLE 3.2

Properties of proteins predicted from the DNA sequence.

		CP12	V			CP12	V
Ala	GCA	8	1	Lys	AAA	5	7
	GCC	5	5		AAG	10	4
	GCG	3	5			15	11
	GCT	3	2				
		19	13	Met	ATG	2	4
						2	4
Arg	AGA*	0	0				
	AGG*	0	0	Phe	TTC	7	4
	CGA*	5	1		TTC	7	2
	CGC	12	2			14	6
	CGG*	1	0				
	CGT	2	3	Pro	CCA	0	0
		20	6		CCC	1	1
					CCG	7	4
Asn	AAC	12	2		CCT	0	0
	AAT	1	3			8	5
		13	5				
Asp	GAC	12	6	Ser	AGC	6	2
	GAT	3	7		AGT	1	1
		15	13		TCA	1	2
					TCC	6	0
Cys	TGC	0	2		TCG	2	0
	TGT	0	1		TCT	1	1
		0	3			17	6
Gln	CAA	1	0	Ter	TAA	0	0
	CAG	17	3	(stop)	TAG	0	0
		18	3		TGA	0	0
						0	0
Glu	GAA	7	6				
	GAG	8	3	Thr	ACA	1	0
		15	9		ACC	8	7
					ACG	3	0
Gly	GGA*	2	0		ACT	0	1
	GGC	10	8			12	8
	GGG*	1	1				
	GGT	4	2	Trp	TGG	5	1
		17	11			5	1
His	CAC	1	1	Tyr	TAC	6	4
	CAT	3	0		TAT	2	2
		4	1			8	6

TABLE 3.2 (cont.)

		CP12	V			CP12	V
Ile	ATA*	2	1	Val	GTA	0	0
	ATC	5	5		GTC	5	6
	ATT	13	4		GTG	3	1
		20	10		GTT	3	1
						11	8
Leu	CTA*	0	0				
	CTC	5	3	TOTAL		258	141
	CTG	14	3	Amino Acids			
	CTT	3	4				
	TTA	1	0				
	TTG	2	2				
		25	12				

	CP12	V
BASIC[a]	35	17
BASIC(%)	13.6	12.1
ACIDIC[a]	30	22
ACIDIC(%)	11.6	15.6
HYDROPHOBIC[a]	94	50
HYDROPHOBIC(%)	36.4	35.5
POLARITY INDEX,%[b]	50.0	44.0
MODULATING CODONS[c](*)	11	3
MODULATING CODONS*(%)	4.3	2.1

Notes to Table 3.2

a. Basic (Lys+Arg), acidic (Glu+Asp) and hydrophobic (Ala+Val+Leu+Ile+Phe+Trp).

b. Proteins below a polarity index of 40% are considered likely candidates for membrane-associated proteins (Capaldi and Vanderkooi, 1972). Polar amino acids (Asp+Asn+Gln+Lys+Ser+Arg+Thr+His).

c. Proteins that are strongly expressed have a low percentage (0.6%) and weakly expressed proteins have a higher percentage (>5.2%) of modulating codons (ATA, AGA, AGG, CGA, CGG, GGA, CTA) as described by Grosjean and Fiers (1982).

13.6% and 12.1% respectively, Table 3.2) than compared with the average distribution of basic amino acids for most proteins (11%; Doolittle, 1981). A clustering of basic amino acid residues (7/11) was noted 116 amino acids from the amino terminal end of *CP12* (sequence positions 430-398, Figure 3.4(a)), but overall both proteins appeared to have an even distribution of basic residues.

In addition, several short peptide motifs that interact with DNA have been identified. One is the helix-turn-helix structural motif seen in the DNA-binding regions of λ Cro-like DNA-binding proteins (Pabo and Sauer, 1984; Scheritz *et al.*, 1985; Anderson *et al.*, 1985) and which can be predicted using the weight matrix system of Dodd and Egan (1987). A further structural motif identified in DNA-binding is the zinc-dependent DNA-binding domain (Miller *et al.*, 1985; Evans and Hollenberg, 1988). It was of interest to examine the translation products of genes *CP12* and *V* for any amino acid sequences showing homology to these DNA-binding motifs.

Using the weight matrix of Dodd and Egan (1990), which uses a larger reference set of presumed helix-turn-helix sequences than the original method (Dodd and Egan, 1987), no amino acid sequence which scored significantly was found in the portion of the translation product of gene *V* (I. Dodd, personal communication). Nor did gene *V* appear to contain any amino acid sequence showing homology to the zinc finger DNA-binding motifs of the C_2H_2 class (transcription factor TFIIIA is the prototype for this class; F/Y-X-C-X_{2-4}-C-X_3-F-X_5-L-X_2-H-X_{3-4}-H, X any amino acid; Evans and Hollenberg, 1988) or of the C_x class (C_4: C-X_2-C-X_{13}-C-X_2-C; C_5: C-X_5-C-X_9-C-X_2-C-X_2-C-X_4-C; C_6: C-X_2-C-X_6-C-X_6-C-X_2-C-X_6-C; Evans and Hollenberg, 1988).

However, the weight matrix of Dodd and Egan (1990) detected in the predicted amino acid sequence of *CP12* a potential helix-turn-helix DNA-binding motif at position 25 of the presumptive protein (I. Dodd, personal communication), which had a SD (standard deviation) score on the new weight matrix of 3.6 and indicated an estimated 60% probability that this sequence represented a true HTH motif. [For comparison, the bacteriophage transcriptional control proteins λ Cro, 434 Cro and P2 Cox score 4.0, 3.7 and 3.1 respectively, using the weight matrix of Dodd and Egan (1990).] This indicated that the product of *CP12* may have DNA-binding capacity and the role of this potential DNA-binding ability is discussed further in Section 3.2.4. No amino acid sequence in

CP12 showing any homology to the zinc finger DNA-binding motifs (Evans and Hollenberg, 1988) was found.

(ii) Polarity index.

A polarity index of less than 40% may indicate that a protein interacts with the cell membrane (Capaldi and Vanderkooi, 1972). Both the translation products of *CP12* and *V* showed a polarity index of greater than 40% (44.0% and 50.0% respectively, Table 3.2) in the portion of the genes contained on the *Bam*HI-*Eco*RI (10.0% to 13.3%) region, and it was therefore unlikely that they represented cell membrane-associated proteins.

3.2.3 Codon usage.

The percentage of modulating (rare) codons in a gene, which correspond to minor tRNA species in *E. coli*, have been shown to correlate with the degree of expression of genes (Ikemara, 1981a,b; Grosjean and Fiers, 1982). It has been postulated that genes with a high percentage of rare codons may be translated more slowly (Grosjean and Fiers, 1982; Gouy and Gautier, 1982; Konigsberg and Godson, 1983) and there is evidence to support this *in vivo* and *in vitro* (Pedersen, 1984; Robinson *et al.*, 1984; Bonekamp *et al.*, 1984). Thus strongly expressed genes appear to have a low content of modulating codons (0.6%) and weakly expressed genes a higher content (5.2%). The portions of the translation products of both *CP12* and *V* encoded on the *Bam*HI-*Eco*RI (10.0% to 13.3.%) showed an intermediate level of modulating codons (4.3% and 2.1% respectively; Table 3.2) and this is expected to be reflected in the expression of these genes *in vivo*.

3.2.4 Homology with late genes *O* and *P* of phage P2.

At the DNA sequence level, 186 and P2 share extensive homology in the late regions of their genomes (Younghusband and Inman, 1974). The related phage P2 displays divergent operons for genes *PQ* and *ONMLKRS* at a parallel position on its chromosome to that of genes *W* and *V* on the 186 chromosome. The first genes represented on each of these operons in P2 (*P* and *O*) have been located in the nucleotide sequence (Christie and Calendar, 1983) and it was therefore of interest to inspect the nucleotide and amino acid sequence of this region in P2 for any homology with the genes predicted on the *Bam*HI-*Eco*RI (10.0% to 13.3%) region (Section 3.2.2(a)).

In the portions of the coding regions available for comparison, the *V* gene of 186 and the *O* gene of P2 were found to share extensive homology at both nucleotide and amino acid levels. As shown in Figure 3.7(a), 186 gene *V* and P2 gene *O* were found to show 87% homology at the nucleotide level and translation of the DNA demonstrated 96% homology at the amino acid level. In the portion of the P2 *P* gene sequenced (Christie and Calendar, 1983), 186 *CP12* and the P2 *P* gene were found to share 51 of the first 65 nucleotides (Figure 3.7(b)). At the amino acid level this corresponded to 19 out of the first 21 amino acids being identical (Figure 3.7(b)). Thus the extensive amino acid homology betwwen 186 genes *V* and *CP12* and the corresponding genes in P2 indicated that the genes products were likely to have similar roles in phage development.

In vitro reconstitution experiments indicated that a 186 *Vam*38 lysate was an efficient tail donor (Hocking and Egan, 1982c), suggesting that the *V* gene product was involved in head formation. However under the electron microscope not only were defective heads evident, but there were also very few tails of normal length (Hocking and Egan, 1982c). Gene *V* was tentatively assigned as having a role in phage head formation, but the reason for the abnormal tails remains unknown. The P2 *O* gene product is involved in phage head formation, but itself is not found in the phage head and may be involved in the cleavage of the major head protein N (Lengyel *et al.*, 1973). Considering the extensive amino acid homology between the products of the 186 *V* and the P2 *O* genes, a similar role for the 186 *V* gene product in morphogenesis is envisaged.

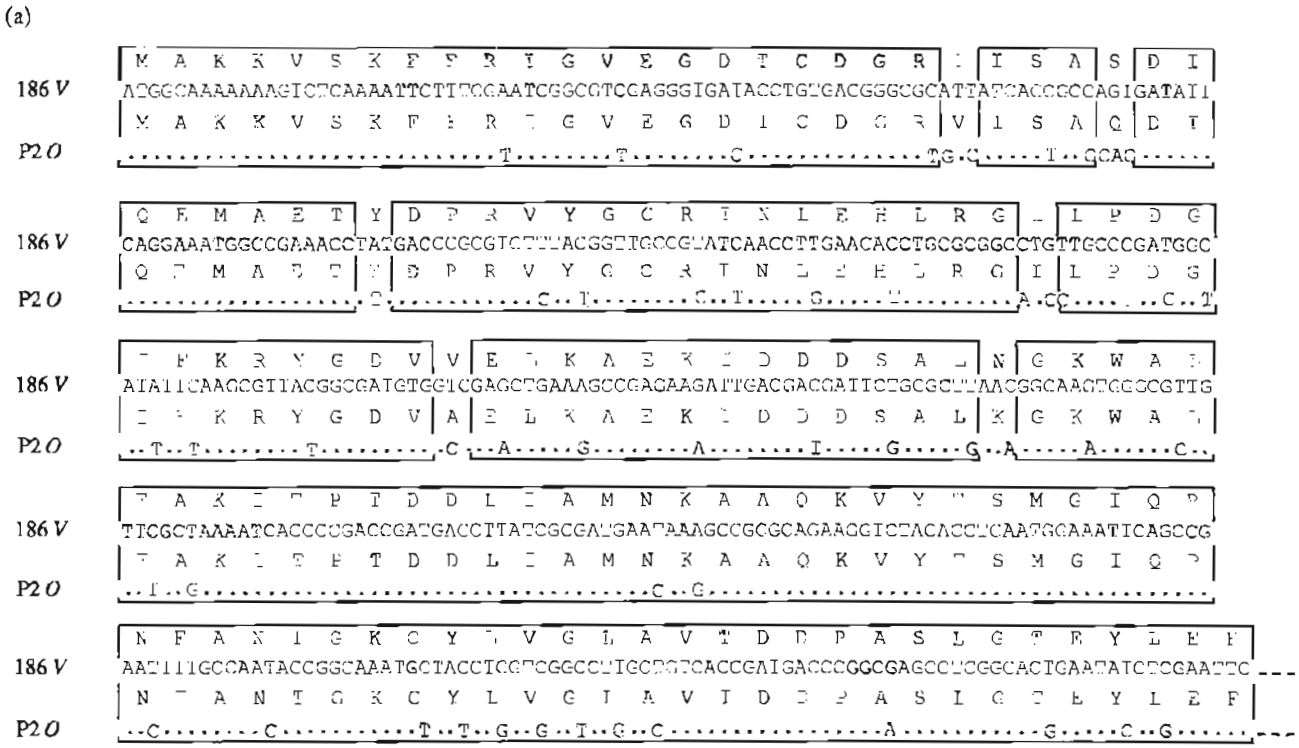
A potential helix-turn-helix motif was identified at position 25 in the amino acid sequence of *CP12* (Section 3.3.2(a)). In phage P2, the *P* gene product apppears to be involved in the production of linear (mature) P2 DNA from the closed circular precursor, in conjunction with the *M* gene product (Bowden and Calendar, 1979). The products of genes *P* and *M* are not found in mature heads (Lengyel *et al.*, 1973) and it maybe that the products of genes *P* and *M* are required for the site-specific cleavage of P2 DNA. The product of 186 CP*12* is likely to be identical to that of the P2 *P* gene, extrapolating from the homology in the amino terminal ends of the proteins. Therefore the *CP12* product may be involved in the site-specific cleavage of 186 DNA and this activity may be associated with the potential helix-turn-helix motif identified in the amino acid sequence. It is interesting to note that the lysate from 186 *Wam*15 contained intact head structures under

Figure 3.7 Nucleotide and amino acid comparison of the 186 *V* and *CP12* genes with the P2 genes *O* and *P*.

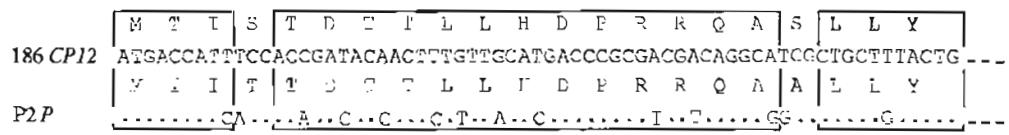
(a) Comparison of the nucleotide and amino acid sequences of 186 gene *V* (Section 3.2.2(a)) and P2 gene *O* (Christie and Calendar, 1983) aligned at the predicted initiation codons. Nucleotides not conserved are indicated. Amino acid homology is indicated by the boxed regions.

(b) Comparison of the nucleotide and amino acid sequences of 186 *CP12* (Section 3.2.2(a)) and the portion of the P2 *P* gene sequenced (Christie and Calendar, 1983). Nucleotide and amino acid homology is as described in the legend to (a).

(a)



(b)



the electron microscope, but these head structures did not appear to contain DNA (Hocking and Egan, 1982c). This would be consistent with the absence of DNA for packaging if the *W* gene product was involved in maturation of 186 DNA. However, it is also possible that the *CP12* product may have a regulatory role in some aspect of phage gene expression, by virtue of the potential helix-turn-helix motif.

The *O* gene of P2 appears to code for a product sized at 30 kD (Lengyel *et al.*, 1973) and the *P* gene of P2 appears to code for a product sized at 65 kD (Bowden and Calendar, 1979). Presuming that the products of 186 genes *V* and *CP12* code for products of a similar size to the corresponding genes in P2, this allowed the prediction of the probable extent of the coding regions for each of the genes on the 186 chromosome. This comparison indicated that 186 gene *V* would probably extend from the predicted initiation codon at 13.1% (Section 3.2.2(a)) to 15.9% on the 186 chromosome and that *CP12* would extend from the predicted initiation codon at 12.6% (Section 3.2.2(a)) to 6.7% on the phage chromosome, assuming that 1% of the 186 chromosome was equivalent to 300 bp.

3.2.5 Promoters.

The opposite directions of the *V* and *CP12* reading frames indicated the presence of divergent transcription that probably initiated between these genes. It was therefore of interest to determine if there were any potential promoters located between these genes which were involved in transcription of these genes.

3.2.5(a) Promoters recognized by unmodified *E. coli* RNA polymerase.

Promoters recognized by unmodified *E. coli* RNA polymerase have been shown to contain highly conserved sequences at the -35, -10 and +1 positions in relation to the initiation sites for RNA synthesis (Siebenlist *et al.*, 1980; Hawley and McClure, 1983; Galas *et al.*, 1985). This has been incorporated into a consensus sequence derived from 112 different promoters (Hawley and McClure, 1983). The extent of homology of a promoter to the consensus has been shown to be related to the strength of the promoter *in vitro* (Mulligan *et al.*, 1984), although sequences outside of the -35, -10 and +1 regions have also been shown contribute significantly to promoter activity in some cases

(Kammerer *et al.*, 1986). The program SCAN (Section 2.34.2) uses a weight matrix derived from the occurrence of each base position of promoters compiled by Hawley and McClure (1983), and was used to assess the presence and strength of possible promoters in the DNA sequence of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region.

Using the SCAN program (Section 2.34.2), no rightward or leftward promoter sequences (scoring above the threshold for each of the promoter spacings) were detected on the entire *Bam*HI-*Eco*RI (10.0% to 13.3%) DNA sequence. By comparison, the early rightward promoter of 186, p_R, which is active *in vitro* (Pritchard and Egan,1985) and *in vivo* (Dodd *et al.*, 1990), scores 67.2 using the SCAN program, 1.2 above the threshold for a spacing of 17 bp (Kalionis *et al.*, 1986a). Thus it was considered unlikely that the *Bam*HI-*Eco*RI (10.0% to 13.3%) region contained any significant promoter recognized by *E. coli* RNA polymerase alone, and that genes *CP12* and *V* were likely to be transcribed from promoters requiring the participation of additional factors for their activity.

3.2.5(b) Prediction of late promoters p_V and p_{12}.

Late transcription in 186 is under the positive control of the protein product of the *B* gene (Finnegan and Egan, 1981). If the positive control of late gene transcription was by a mechanism of promoter activation, then the nucleotide sequence of a 186 late promoter would be expected to differ significantly from the *E. coli* consensus sequence for promoters recognised by RNA polymerase alone. The opposite directions of the *V* and *CP12* reading frames indicated the presence of divergent transcription that possibly initiated between these genes, and consistent with a mechanism of promoter activation, the program SCAN failed to detect any potential promoters recognized by unmodified *E. coli* RNA polymerase in the complete nucleotide sequence of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region (Section 3.2.5(a)).

Phage P2 displays divergent operons *QP* and *ONMLKRS* at a parallel position on its chromosome (Figure 1.1), and the associated promoters have been identified and located between genes *O* and *P* (Christie and Calendar, 1983). As the late control protein B of 186 is known to activate the late promoters of P2 (Hocking and Egan, 1982b), the nucleotide sequence of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region was inspected for any homology with the late promoters p_Q and p_P of P2. Two regions of extensive homology

were detected. One region on the *r*-strand, spanning sequence coordinates 849-778, showed considerable homology to the nucleotide sequence upstream of the transcription initiation site for the P2 *P* mRNA, particularly in the -10 and -35 regions from the transcription initiation point of p_P (Figure 3.8(a); Christie and Calendar, 1983). The second region on the *l*-strand, which spanned sequence coordinates 841-903, showed extensive homology to the nucleotide sequence upstream of the initiation site for transcription for the P2 *O* mRNA (Figure 3.8(b); Christie and Calendar, 1983). The extensive homology upstream of the coding regions for 186 *V* and *CP12* with the P2 late promoters p_O and p_P, and the fact that the 186 late control protein B is able to activate the P2 late promoters (Hocking and Egan, 1982b), indicated that genes *CP12* and *V* of 186 were likely to be transcribed from late promoters equivalent to those identified in P2 and led to the prediction that the *V* operon was transcribed from a rightward promoter p_V initiating at position 909, while *CP12* was predicted to be transcribed leftward from promoter p_{12} initiating at position 798.

3.2.6 Homology of the late promoters p_V and p_{12} with the late promoters of P2.

The nucleotide sequences around the transcript initiation sites for four P2 late promoters have been determined (Christie and Calendar, 1983; 1985). A comparison of the late promoter sequences of P2 and the predicted promoters p_V and p_{12} of 186 is presented in Figure 3.9.

As expected for promoters that are probably activated, the DNA sequence of the late promoters of 186 and P2 also differed significantly from the *E. coli* consensus sequence (Hawley and McClure, 1983). However, consistent with the promoters probably activated by a similar mechanism, several regions of homology were evident, at the positions -10, -35 and -55 upstream of the predicted transcription initiation sites (Figure 3.9)

McClure (1985) concluded that positively controlled promoters in *E. coli* recognized by RNA polymerase utilizing the σ^{70} subunit possessed at least two out of the three most highly conserved bases TA---T in the conserved hexamer sequence about the -10 region (TAtaaT) and at least one of the most highly conserved bases of the TTG sequence in the -35 hexamer (TTGaca). Inspection of the 186 late promoters p_V and p_{12}

Figure 3.8 Comparison of the DNA sequence upstream of the 186 genes *V* and *CP12* with the sequences of the P2 late promoters p_O and p_P.

(a) Comparison of the DNA sequence upstream of the predicted initiation codon for the 186 *V* gene with the DNA sequence of the late promoter for the P2 *O* gene (Christie and Calendar, 1983). The *l*-strand of the sequence of the 186 *V* gene is shown with sequence positions from the *Bam*HI site at 10.0% shown above the DNA sequence. The presumptive initiation codons of the 186 *V* gene (Section 3.2.2(a)) and the P2 *O* gene (Christie and Calendar, 1983) are underlined. Homology is indicated by circles between the DNA sequences. To maximize the homology, spaces have been introduced into the DNA sequence upstream of the 186 *V* gene (between positions 924 and 925). The transcription initiation site for the P2 *O* gene is indicated (+1). The conserved sequences about the -10 and -35 regions in the P2 p_O promoter (Christie and Calendar, 1985) are boxed. The predicted initiation site for transcription of the 186 *V* gene is shown.

(b) Comparison of the DNA sequence upstream of the predicted initiation codon for the 186 *CP12* gene with the DNA sequence of the late promoter for the P2 *P* gene (Christie and Calendar, 1983). The *r*-strand of the DNA sequence upstream of the 186 *CP12* gene is shown, with sequence positions from the *Bam*HI site at 10.0% shown above the sequence. To maximize the homology (as indicated by circles), a space has been introduced into the DNA sequence of the P2 p_P promoter (adjacent to postion 829). The presumptive initiation codons (Section 3.2.2(a); Christie and Calendar, 1983) and transcription start site for the P2 p_P promoter (Christie and Calendar, 1983) are as indicated in the legend to (a). The predicted transcription initiation site of the 186 *CP12* gene is shown.

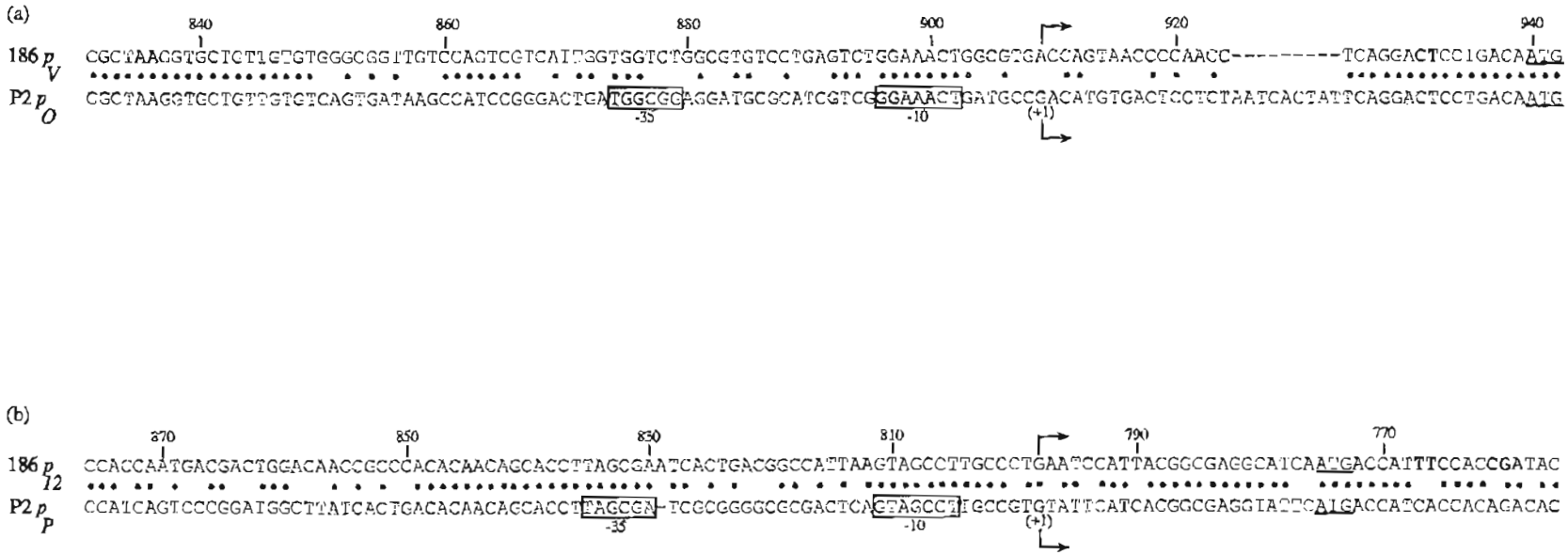
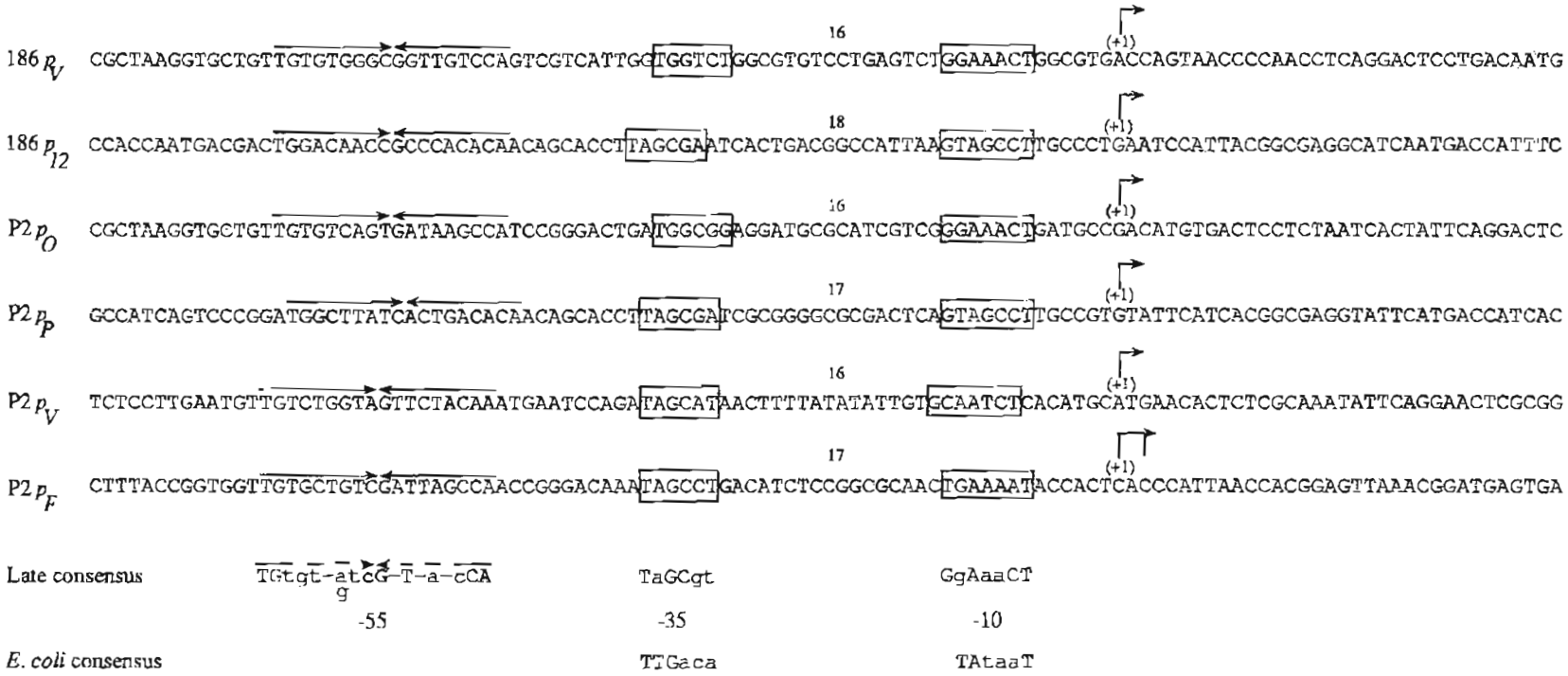

(a)
840
860
880
900
920
940
186 p_V
P2 p_O
-35
-10
(+1)
(b)
870
850
830
810
790
770
186 p_{12}
P2 p_P
-35
-10
(+1)

Figure 3.9 Comparison of the late promoter sequences of 186 and P2.

Nucleotide sequences of the predicted 186 late promoters p_V and p_{12} (Section 3.2.5(b)) and the P2 late promoters (Christie and Calendar, 1985) are shown aligned at the transcription initiation sites. The transcription initiation sites for the P2 late promoters (Christie and Calendar, 1985) and the predicted transcription initiation sites for the 186 late promoters p_V and p_{12} (Section 3.2.5(b)) are shown (+1). Boxed regions indicate sequences in the late promoters showing homology and are positioned about the -10 and -35 regions as originally identified by Christie and Calendar (1985). The number of intervening bases between the conserved sequences at the -10 and -35 positions are shown. Dale *et al.* (1986) identified an inverted repeat structure centered at the -56 position in the P4 p_{sid} promoter and noted homology with the P2 late promoters in the corresponding region (-55 position). These inverted repeat structures are indicated by converging arrows. Inverted repeat structures in the 186 late promoters p_V and p_{12} (Section 3.2.5(b)) are also shown. The conserved sequences about the -10, -35 and -55 positions have been incorporated into a late promoter consensus sequence as shown below the promoter sequences. Upper case letters represent nucleotides present in 6/6 or 5/6 of each sequence position. Lower case letters represent nucleotides present in 4/6 or 3/6 of each sequence position. The *E. coli* promoter consensus sequence (Hawley and McClure, 1983) is also shown for comparison, with upper case letters representing highly conserved nucleotides, lower case letters less well conserved nucleotides.


186 p_V CGCTAAGGTGCTGTTGTGTGGGCGGTTGTCCAGTCGTCATTGGTGGTCTGGCGTGTCCTGAGTCTGGAAACTGGCGTGACCAGTAACCCCAACCTCAGGACTCCTGACAATG
16
(+1)
186 p_{12} CCACCAATGACGACTGGACAACCGCCCACACAACAGCACCTTAGCGAATCACTGACGGCCATTAAGTAGCCTTGCCCTGAATCCATTACGGCGAGGCATCAATGACCATTTC
18
(+1)
P2 p_O CGCTAAGGTGCTGTTGTGTCAGTGATAAGCCATCCGGGACTGATGGCGGAGGATGCGCATCGTCGGGAAACTGATGCCGACATGTGACTCCTCTAATCACTATTCAGGACTC
16
(+1)
P2 p_P GCCATCAGTCCCGGATGGCTTATCACTGACACAACAGCACCTTAGCGATCGCGGGGCGCGACTCAGTAGCCTTGCCGTGTATTCATCACGGCGAGGTATTCATGACCATCAC
17
(+1)
P2 p_V TCTCCTTGAATGTTGTCTGGTAGTTCTACAAATGAATCCAGATAGCATAACTTTTATATATTGTGCAATCTCACATGCATGAACACTCTCGCAAATATTCAGGAACTCGCGG
16
(+1)
P2 p_F CTTTACCGGTGGTTGTGCTGTCGATTAGCCAACCGGGACAAATAGCCTGACATCTCCGGCGCAACTGAAAATACCACTCACCCATTAACCACGGAGTTAAACGGATGAGTGA
17
(+1)
Late consensus TGtgt-atcG-T-a-cCA TaGCgt GgAaaCT
-55 -35 -10
E. coli consensus TTGaca TAtaaT

and the P2 late promoters (Christie and Calendar, 1985) revealed that in the -10 region, three of the four P2 late promoters and the 186 late promoters p_V and p_{12} shared with the *E. coli* consensus sequence the highly conserved T residue at the -7 position (TAtaaT) and the highly conserved A residue at the -11 position (TAtaaT) (Figure 3.9). In the -35 region, regions of homology were evident between the late promoters and in these sequences limited homology to the *E. coli* -35 consensus could be identified, positioned 16-18 bp upstream of the conserved -10 regions in these promoters (Figure 3.9). In these regions the first T and the third G residues were conserved with the hexamer *E. coli* consensus sequence (TTGaca). The lack of homology of the late promoters with the *E. coli* consensus sequence and the fact that the late promoters conformed to the characteristics defined by McClure (1985), indicated that the 186 and P2 late promoters were likely to be positively regulated promoters which required the participation of the σ^{70} subunit of RNA polymerase.

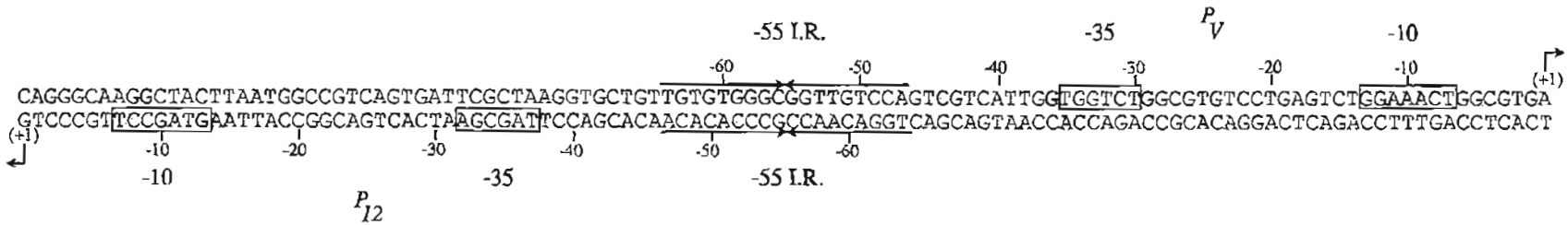
Dale *et al.* (1986) identified an inverted repeat structure at the -56 position in the P4 *sid* promoter and noted homology in the corresponding position (-55) in the P2 late promoters (Figure 3.9). The P4 *sid* promoter is activated by the P4 δ and P2 Ogr proteins (Dale *et al.*, 1986; Keener *et al.*, 1988) and this inverted repeat element has been shown to be essential for activation of the *sid* promoter by the δ gene product (van Bokkelen *et al.*, 1990). This inverted repeat element was also identified at the same position in the 186 late promoters p_V and p_{12} (Figure 3.9), and therefore may represent a sequence involved in the activation of the late promoters. This element has been postulated to be the DNA-binding site for the late control proteins in order to activate transcription from the late promoters (Halling, 1989).

The arrangement of the divergent 186 late promoters p_V and p_{12} was found to be such that both promoters shared the same inverted repeat element at the -55 position with respect to the predicted transcription initiation sites for both promoters (Figure 3.10). This arrangement has also been noted for the corresponding p_O and p_P promoters of P2 (Halling, 1989) and suggests that these divergent promoters in 186 and P2 may utilize this same element in the control of their transcription.

It was concluded from the lack of sequence homology of the late promoters of 186 and P2 with the *E. coli* consensus that the late promoters were unlikely to be recognised by

Figure 3.10 Arrangement of the inverted repeat element in the 186 late promoters p_V and p_{12}.

Nucleotide sequence of the intergenic region between 186 genes *V* and *CP12* showing the positioning of the inverted repeat element with respect to the predicted transcription initiation sites of promoter p_V and p_{12} (Section 3.2.5(b)). Predicted transcription initiation sites are indicated (+1). Nucleotide positions upstream of each transcription initiation site are indicated. The inverted repeat element (shown by the converging arrows) is located at the -55 position from the transcription initiation site of each promoter.


-55 I.R.
-35
P_V
-10
-60
-50
-40
-30
-20
-10
(+1)
CAGGGCAAGGCTACTTAATGGCCGTCAGTGATTCGCTAAGGTGCTGTTGTGTGGGCGGTTGTCCAGTCGTCATTGGTGGTCTGGCGTGTCCTGAGTCTGGAAACTGGCGTGA
GTCCCGTTCCGATGAATTACCGGCAGTCACTAAGCGATTCCAGCACAACACACCCGCCAACAGGTCAGCAGTAACCACCAGACCGCACAGGACTCAGACCTTTGACCTCACT
(+1)
-10
-20
-30
-40
-50
-60
-10
-35
-55 I.R.
P_{12}

E. coli RNA polymerase alone and that the control of late transcription was most likely to be by promoter activation, with an accessory factor (e.g. 186 B protein) required for transcription initiation to occur. For a similar reason, control of late transcription by antitermination of a constitutive transcript (and therefore a promoter recognized by *E. coli* RNA polymerase alone) was considered unlikely.

3.2.7 The *B box.*

The 186 B protein appears to repress its own transcription, either directly or indirectly (Kalionis *et al.*, 1986b). Two copies of an inverted repeat were found overlapping the *B* promoter (Figure 3.11) and Kalionis *et al.* (1986b) suggested that these inverted repeats may be the binding site for the factor present in the lysogenic state which represses *B* gene transcription. Three of the four late promoters of P2 (p_Q, p_P and $p_V p_Q$) were found to contain a highly conserved ten nucleotide sequence centered approximately 25 nucleotides downstream of the transcription initiation site (Figure 3.11; Christie and Calendar, 1985), which corresponded to one arm of the inverted repeats found overlapping the *B* promoter. This ten nucleotide sequence has been termed the *B box* (Kalionis *et al.*, 1986b).

Sequences showing homology to the *B box* were found four times in the DNA sequence downstream of the transcription initiation site for the 186 late promoter p_V (Figure 3.4(a)), but not in the DNA sequence downstream of the leftward late promoter p_{12}. However, the sequences identified downstream of p_V were located considerably further downstream than the *B box* elements found in the P2 late promoters (Figure 3.11). Since the 186 B protein is capable of activating both the 186 and P2 late promoters (Finnegan and Egan, 1981; Hocking and Egan, 1982c), and also in repressing its own transcription (Kalionis *et al.*, 1986b), the *B box* may have a role in the control of transcription from the late promoters of 186 and P2. The role of the *B box* in the activation of the late promoter p_V is investigated in Chapter 4.

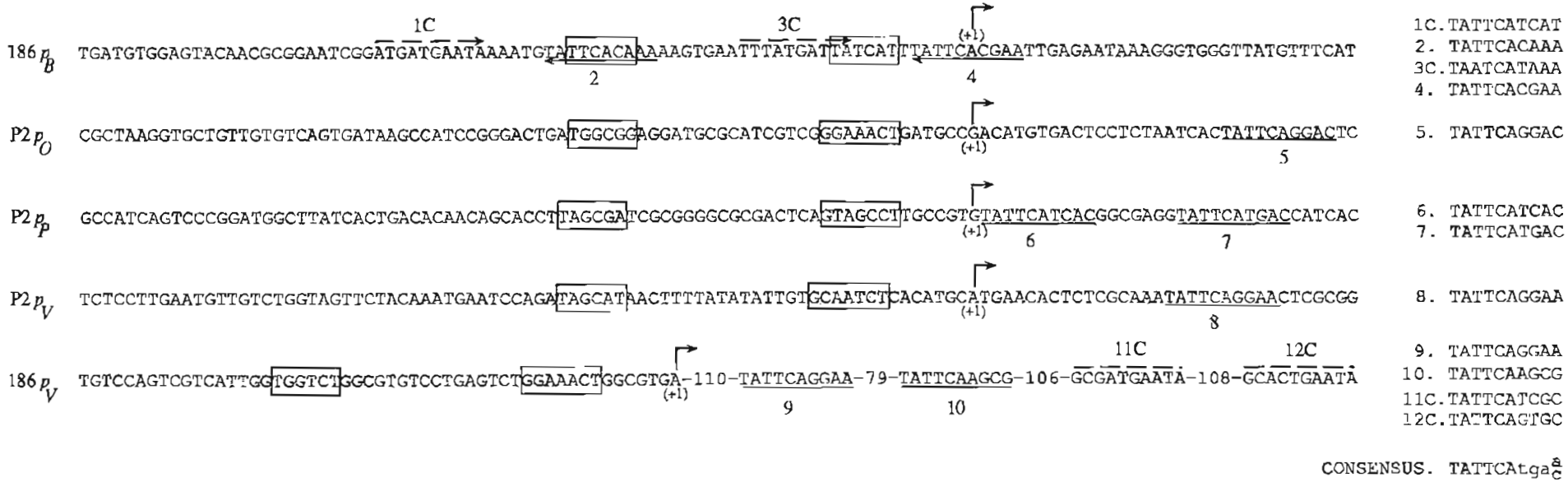
3.2.8 Termination of transcription.

DNA sequences that signal the termination of transcription in *E. coli* belong to two major classes, the *rho*-independent and the *rho*-dependent terminators (Rosenberg and

Figure 3.11 The *B box*.

Promoter of the 186 *B* gene (Kalionis *et al.*, 1986b) and the late promoters of the P2 genes *O*, *P* and *V* (Christie and Calendar, 1985) are presented aligned at the proposed mRNA start-point of each gene (+1). The -10 and -35 regions of the 186 *B* promoter (Kalionis *et al.*, 1986b) and the P2 late promoters (Christie and Calendar, 1985) are boxed. The inverted repeat structures overlapping the *B* promoter are indicated by converging arrows (Kalionis *et al.*, 1986b). One arm of these inverted repeats has been found downstream of three of the four late promoters (underlined regions 5-8), and has been designated the *B box* (Kalionis *et al.*, 1986b). The 186 late promoter p_V is presented with the *B box* sequences identified downstream of the predicted transcription initiation site (regions 9, 10, 11C and 12C; Section 3.2.7). Intervening nucleotides are indicated. *B box* sequences found in the left strand of the sequence are indicated by solid lines (regions 2,4, 5-10) and sequences on the complementary strand are indicated by dashed lines (regions 1C, 3C, 11C and 12C).

A consensus sequence for the *B box* sequences identified is presented to the right. Upper case letters indicate nucleotides present in 12/12 or 11/12 of the sequence postions. Lower case letters indicate sequences present in 5 or more of the sequences.


186 p_B
TGATGTGGAGTACAACGCGGAATCGGATGATGAATAAAATGTATTCACAAAAAGTGAATTTATGATTATCATTTATTCACGAATTGAGAATAAAGGGTGGGTTATGTTTCAT
1C
2
3C
4
(+1)
P2 p_O
CGCTAAGGTGCTGTTGTGTCAGTGATAAGCCATCCGGGACTGATGGCGGAGGATGCGCATCGTCGGGAAACTGATGCCGACATGTGACTCCTCTAATCACTATTCAGGACTC
5
(+1)
P2 p_P
GCCATCAGTCCCGGATGGCTTATCACTGACACAACAGCACCTTAGCGATCGCGGGGCGCGACTCAGTAGCCTTGCCGTGTATTCATCACGGCGAGGTATTCATGACCATCAC
6
7
(+1)
P2 p_V
TCTCCTTGAATGTTGTCTGGTAGTTCTACAAATGAATCCAGATAGCATAACTTTTATATATTGTGCAATCTCACATGCATGAACACTCTCGCAAATATTCAGGAACTCGCGG
8
(+1)
186 p_V
TGTCCAGTCGTCATTGGTGGTCTGGCGTGTCCTGAGTCTGGAAACTGGCGTGA-110-TATTCAGGAA-79-TATTCAAGCG-106-GCGATGAATA-108-GCACTGAATA
9
10
11C
12C
(+1)
1C. TATTCATCAT
2. TATTCACAAA
3C. TAATCATAAA
4. TATTCACGAA
5. TATTCAGGAC
6. TATTCATCAC
7. TATTCATGAC
8. TATTCAGGAA
9. TATTCAGGAA
10. TATTCAAGCG
11C. TATTCATCGC
12C. TATTCAGTGC
CONSENSUS. TATTCAtga a/c

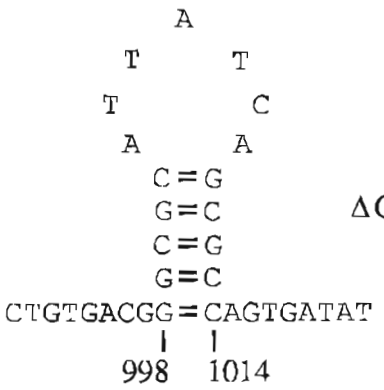
Court, 1979; Adhya and Gottesman, 1978; Galloway and Platt, 1985; Platt, 1986). *rho*-independent terminators are functionally active *in vitro*, without the requirement for additional proteins. They are characterized by a region of dyad symmetry rich in GC base-pairs, which allows the formation of a stable stem-loop structure in the RNA transcript, and a string of consecutive T residues, at which termination of transcription usually occurs (Holmes *et al.*, 1983; Galloway and Platt, 1985; Platt, 1986). The determinants of *rho*-dependent termination in *E. coli* are only partially understood. Several *rho*-dependent terminators have been identified but show few common features (von Hippel *et al.*, 1984; Morgan *et al.*, 1985; Platt, 1986). They are generally, but not always, characterized by a region of GC-rich hyphenated dyad symmetry and usually associated with a stretch of untranslated RNA (100 to 300 nucleotides) that has a low content of G residues (Morgan *et al.*, 1983a,b; Lowery and Richardson, 1977; Adhya *et al.*, 1979; von Hippel *et al.*, 1984; Lau and Roberts, 1985; Galloway and Platt, 1985; Ceruzzi *et al.*, 1985). The region of untranslated RNA has been shown to be required for the interaction of Rho protein with the RNA (Richardson and Macy, 1981; Richardson, 1982; Sharp and Platt, 1984; Ceruzzi *et al.*, 1985; Chen *et al.*, 1986).

Potential *rho*-independent transcription terminators (stem-loop structures) were searched for using the program COMSTR (Section 2.34.3) which searches for stem-loop structures and calculates the ΔG values of the structures using the rules of Tinoco *et al.*, (1973) as modified by Steger *et al.* (1984). Potential transcription terminators were searched for in the nucleotide sequence of the regions of the *Bam*HI-*Eco*RI (10.0% to 13.3%) fragment predicted to be transcribed from the late promoters p_{12} (sequence co-ordinates 798-1 on the *r*-strand;) and p_V (sequence co-ordinates 909-1362 on the *r*-strand). No simple stem-loop structures were found in the region on the *r*-strand of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region predicted to be transcribed from the late promoter p_{12}. Two simple stem-loop structures were found on the *l*-strand, in the region transcribed from the late promoter p_V (Figure 3.12) and were designated Structure 1*l* (sequence coordinates 999-1014) and Structure 2*l* (sequence coordinates 1168-1189). Neither of these structures was consistent with a *rho*-independent terminator, since they were not followed by a long string of consecutive T residues. It was also considered unlikely that they represented *rho*-dependent terminators, since they were both located in a region expected to be translated,

Figure 3.12 Potential stem-loop structures in the *Bam*HI-*Eco*RI (10.0% to 13.3%) region from 186 *cIts*.

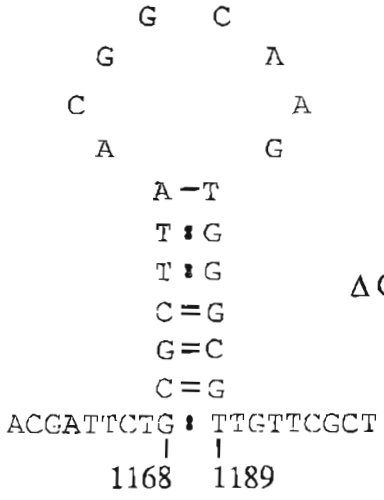
This figure shows the most stable simple stem-loop structures predicted within the *Bam*HI-*Eco*RI (10.0% to 13.3%) region. A threshold value of $\Delta G = -8.0$ was arbitrarily chosen. The stem-loop structures were predicted using the program COMSTR and by dot matrix analysis (Section 2.34). The free energy values were calculated using the rules of Steger *et al.* (1984) and are shown to the right of each structure. The DNA sequence coordinates are listed beneath each structure.

Structure 1*l*



$\Delta G = -8.03$ Kcal/mol

Structure 2*l*



$\Delta G = -8.25$ Kcal/mol

although *rho*-dependent terminators have been observed in translated regions (Szybalski *et al.*, 1983).

Thus no significant termination was expected to occur from either of the late promoters p_{12} or p_V in the *Bam*HI-*Eco*RI (10.0% to 13.3%) region. Since the approach to be taken to identify the phage functions involved in activation of late transcription was to fuse these promoters to a reporter function, it was necessary that no significant termination of transcription occured in the cloned DNA preceding the reporter function. Since there were no identifiable *rho*-independent or *rho*-dependent transcription terminators in the nucleotide sequence downstream of the late promoters p_V or p_{12}, fusion of either promoter to a reporter gene should allow expression of the reporter gene upon activation without any significant transcription termination occurring and possibly masking activation.

3.3 SUMMARY.

Determination of the DNA sequence of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region of 186 and analysis of the DNA sequence revealed that this region is likely to encode two divergent genes, which are not completely contained on the *Bam*HI-*Eco*RI (10.0% to 13.3%) restriction fragment. The rightward gene has been previously identified as gene *V* (Dibbens, 1984). The leftward gene has been designated *CP12*. Both genes show extensive homology to the corresponding head genes of phage P2, *O* and *P* respectively (Christie and Calendar, 1983) and indicates that they have similar roles in phage morphogenesis. A potential helix-turn-helix DNA binding motif was located in the amino acid sequence of *CP12* using the method of Dodd and Egan (1990) and since the *P* gene of P2 is involved in the site-specific cleavage of P2 DNA, the presumptive helix-turn-helix motif in *CP12* may be involved in a similar reaction in the site-specific cleavage of 186 DNA.

The divergent nature of the genes *V* and *CP12* indicated the likelihood of divergent transcription that initiated between the genes. Analysis of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region indicated that there was unlikely to be any promoter recognized by unmodified *E. coli* RNA polymerase on the fragment. However, the DNA sequence upstream of genes *V* and *CP12* showed extensive homology to the corresponding late promoters of P2 (Christie and Calendar, 1985) and genes *V* and *CP12* were predicted to be

transcribed from non-overlapping, divergent late promoters. The lack of homology of the late promoters p_V and p_{12} with the *E. coli* consensus sequence indicated that control of late gene transcription was most likely to be by an activation mechanism rather than an antitermination mechanism. The *in vivo* activities of the late promoters p_V and p_{12} are investigated in Chapter 4.

CHAPTER 4

INVESTIGATION OF THE *IN VIVO* ACTIVITIES OF THE LATE PROMOTERS ENCODED ON *Bam*HI-*Eco*RI (10.0%-13.3%) REGION

CHAPTER 4. INVESTIGATION OF THE *IN VIVO* ACTIVITIES OF THE LATE PROMOTERS ENCODED ON THE *Bam*HI-*Eco*RI (10.0% to 13.3%) REGION.

4.1 INTRODUCTION.

The results presented in Chapter 3 indicated that the *Bam*HI-*Eco*RI (10.0% to 13.3%) region encoded two divergent genes, *CP12* and *V* (Section 3.2.2(a)). The divergent nature of these genes indicated the possibility of divergent transcription, and the DNA sequence upstream of each of these genes was found to show considerable homology to the late promoters of phage P2 (Section 3.2.5(b)). Since the late control protein B of 186 is able to activate the P2 late promoters (Hocking and Egan, 1982b), genes *CP12* and *V* were predicted to be transcribed from non-overlapping, divergent late promoters, designated p_{12} and p_V respectively (Section 3.2.5(b)).

In order to confirm the *in vivo* activities of the late promoters p_V and p_{12} predicted from the DNA sequence (Section 3.2.5(b)), it was necessary to identify the associated transcripts and determine that the sites of initiation of transcription were consistent with the sites of initiation predicted from the DNA sequence (Section 3.2.5(b)). In addition, these transcripts should be present late in lytic infection and be dependent upon the function of the *B* gene product, as demonstrated for late transcription by Finnegan and Egan (1981). Identification of the associated transcripts and location of the 5'-ends can be determined by primer extension analysis of *in vivo* RNA.

To identify the phage functions required for late gene transcription the approach chosen was to clone a 186 late promoter before a reporter gene and then determine the functions necessary to achieve transcription of the reporter gene. Use of the *galK* gene as a reporter function in transcription-fusion systems has been shown to be effective in the analysis of transcriptional regulation (McKenney *et al.*, 1981). Such a system also allows the identification of the *cis*-acting elements in the DNA sequence required for initiation of transcription.

This chapter describes the confirmation of the *in vivo* activities of the rightward late promoter p_V and the leftward late promoter p_{12} by primer-extension analysis and the establishment of an assay system for the activation of a 186 late promoter, by fusion of a region encoding the late promoter p_V to the reporter gene *galK*. One element that has been

implicated in the control of late gene transcription in 186 and P2 is the ten nucleotide sequence referred to as the *B box* (Section 3.2.7; Kalionis *et al.*, 1986b). Using the *galK* system, the role of the *B boxes* in the activation of transcription from the late promoter p_V is investigated.

4.2 RESULTS AND DISCUSSION.

4.2.1 Confirmation of the *in vivo* activity of the rightward late promoter p_V.

In order to confirm the *in vivo* activity of a promoter, it is necessary to identify the associated transcript and determine that the location of the 5'-end of the transcript is consistent with initiation of transcription from the predicted promoter. One method which can be used to identify transcripts and locate the 5'-ends is that of primer extension analysis (McKnight *et al.*, 1981). In this method, a radioactive primer is annealed to total cellular RNA and then extended with AMV reverse transcriptase in the presence of all four unlabelled dNTPs. The primer specifically hybridizes to complementary RNA transcripts and is extended to the 5'-end of the RNA, generating specific extension products that can be accurately sized by comparing the mobility relative to a DNA sequencing ladder on a 6% denaturing polyacrylamide gel (Section 2.33). However, this technique cannot differentiate between 5'-ends due to transcription initiation and the products of RNA processing or degradation.

Primer extension analysis has been used to confirm the *in vivo* activity of the rightward late promoter p_V (Dibbens, 1984). Using a primer specific for *V* gene transcription, transcription from p_V was detectable late in lytic infection (35 minutes after heat induction of a 186 *cIts* prophage) and was shown to be dependent upon the protein product of the *B* gene, as no extension product was detectable upon induction of a 186 *cIts Bam*17 prophage in a non-permissive host (Dibbens, 1984). This was consistent with the requirement of the *B* gene product in late gene transcription as observed by Finnegan and Egan (1981).

However, the primer used to detect the *V* transcript (a *r*-strand 161 base *Hae*III-*Acc*I (sequence coordinates 1072-1252) primer) required extension with reverse transcriptase in excess of 180 bases to the predicted 5'-end of the transcript (Section

3.2.5(b)) and therefore made determination of the precise location of the 5'-end difficult. In order to confirm that the site of initiation of transcription was consistent with that predicted, primer extension analysis using a primer located closer to the predicted 5'-end was required.

As shown in Figure 4.1, the rightward promoter p_V was predicted to initiate transcription at sequence position 909 in the DNA sequence of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region (Section 3.2.5(b)). To locate the site of initiation of transcription from p_V, the *r*-strand of a 82 bp *Oxa*NI-*Hae*II (sequence coordinates 928 to 1009) restriction fragment was chosen as an appropriate primer to map the 5'-end of the *V* transcript, since it was located 19 bases from the predicted site of initiation (Section 3.2.5(b)). This primer was prepared by synthesizing a radioactive, complementary copy (Section 2.30.2) of the *l*-strand M13 clone mEC404 (Section 2.4.2), followed by restriction with *Hae*II and *Oxa*NI and fractionation on a 5% non-denaturing polyacrylamide gel (Section 2.26.2(a)).

The 82 bp *Oxa*NI-*Hae*II restriction fragment with the *r*-strand labelled, was heat-denatured (Section 2.32.5) and used in an extension reaction with AMV reverse transcriptase (Section 2.33) on RNA isolated from a heat-induced 186 *cIts* lysogen of the non-permissive strain E251 (35 minutes after induction), a heat-induced 186 *cIts Bam*17 lysogen of strain E251 (35 minutes after induction), a non-induced E251 186 *cIts* lysogen, and the corresponding non-lysogen 35 minutes after mock heat-induction (Section 2.32.1; Section 2.32.2). The products were analysed on a 6% denaturing gel using a sequencing ladder to size the extension products and the results are shown in Figure 4.2.

Consistent with the primer extension studies originally performed (Dibbens, 1984), a specific extension product was observed on extension on RNA isolated late after heat induction of a 186 *cIts* lysogen (Figure 4.2, track 1). The largest and most predominant extension product obtained on extension on late RNA was sized at 19 bases larger than the unextended primer and corresponded to a 5'-end (from the *Oxa*NI site at sequence coordinate 928) at sequence coordinate 909, which was identical with that predicted for the site of initiation of transcription from p_V (Section 3.2.5(b)). No extension products were detected on extension on RNA isolated 35 minutes after mock heat-induction of the corresponding non-lysogen (Figure 4.2, track 3) and demonstrated that the transcripts detected upon 186 *cIts* prophage induction were specific to the induced lysogen.

Figure 4.1 DNA sequence of the *Bss*HII-*Alu*I (11.9% to 13.8%) region from 186 *c*Its.

The nucleotide sequence of the *Bss*HII-*Alu*I (11.9% to 13.8%) region from 186 *c*Its, showing the predicted amino acid sequences of the proteins encoded in the region. The upper strand represents the *l*-strand (Kalionis and Egan, 1981) and the numbering is from the *Bam*HI site at 10.0% (Section 3.2.1). Ribosome-binding sites (RBS) and predicted initiation codons are boxed (Section 3.2.2(a)). Predicted transcription initiation sites (Section 3.2.5(b)) are indicated (+1). Sequences showing homology to the *B box* sequences (Section 3.2.7; Kalionis *et al.*, 1986b) are underlined. Relevant restriction sites are indicated. The coordinates of the restriction sites used to generate the primers for primer extension analysis described in the text are from the bases constituting the respective 5' and 3'-ends of the primers on the appropriate strand.

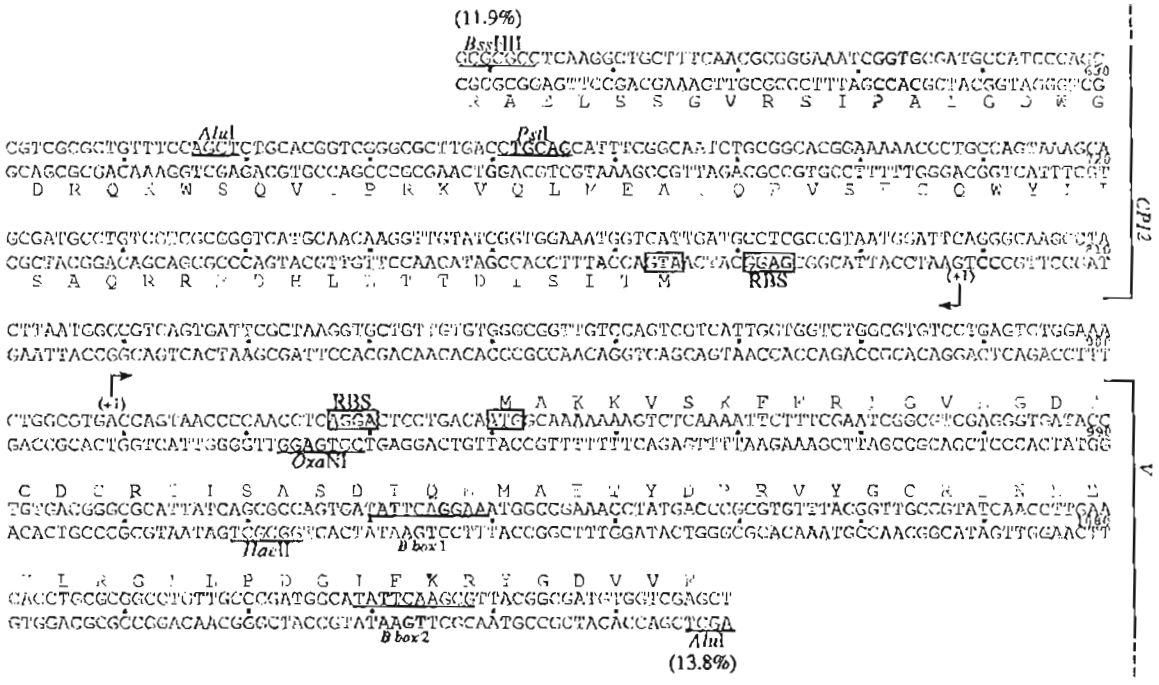

(11.9%)
BssHII
AluI
PstI
RBS
(+1)
RBS
(+1)
OxaNT
NaeII
B box 1
B box 2
AluI
(13.8%)
cpl2
A

Figure 4.2 Confirmation of the *in vivo* activity of the late promoter p_V by primer extension.

RNA was prepared from cultures of E252 [E251(186 *cIts*)], E264 [E251(186 *cIts Bam*17)] and E251, 35 minutes after transferral from 30°C to 39°C as described in Section 2.32. RNA was also prepared from a non-induced culture of E252 at 30°C. The extension reaction was based on the method of McKnight *et al.* (1981) and is described in Section 2.33. The 82 base *Oxa*NI-*Hae*II (sequence coordinates 928 to 1009) primer specific for rightward message was constructed by extension (Section 2.30.2) on the *l*-strand M13 clone mEC404 (Section 2.4.2), followed by restriction with *Hae*II and *Oxa*NI and fractionation on a 5% non-denaturing polyacrylamide gel (Section 2.26.2(a)). This primer ($2.0x10^5$ cpm per reaction) was denatured, annealed to 5 ug of *in vivo* RNA and extended with AMV reverse transcriptase in the presence of all four dNTPs as described in Section 2.33. Electrophoresis was as described in Section 2.26.2(b). The autoradiograph was exposed at room temperature for 12 hours.

Gel Track

1. Primer extension of RNA isolated 35 min after heat-induction of a 186 *cIts* lysogen.

2. Primer extension of RNA isolated 35 min after heat-induction of a 186 *cIts Bam*17 lysogen.

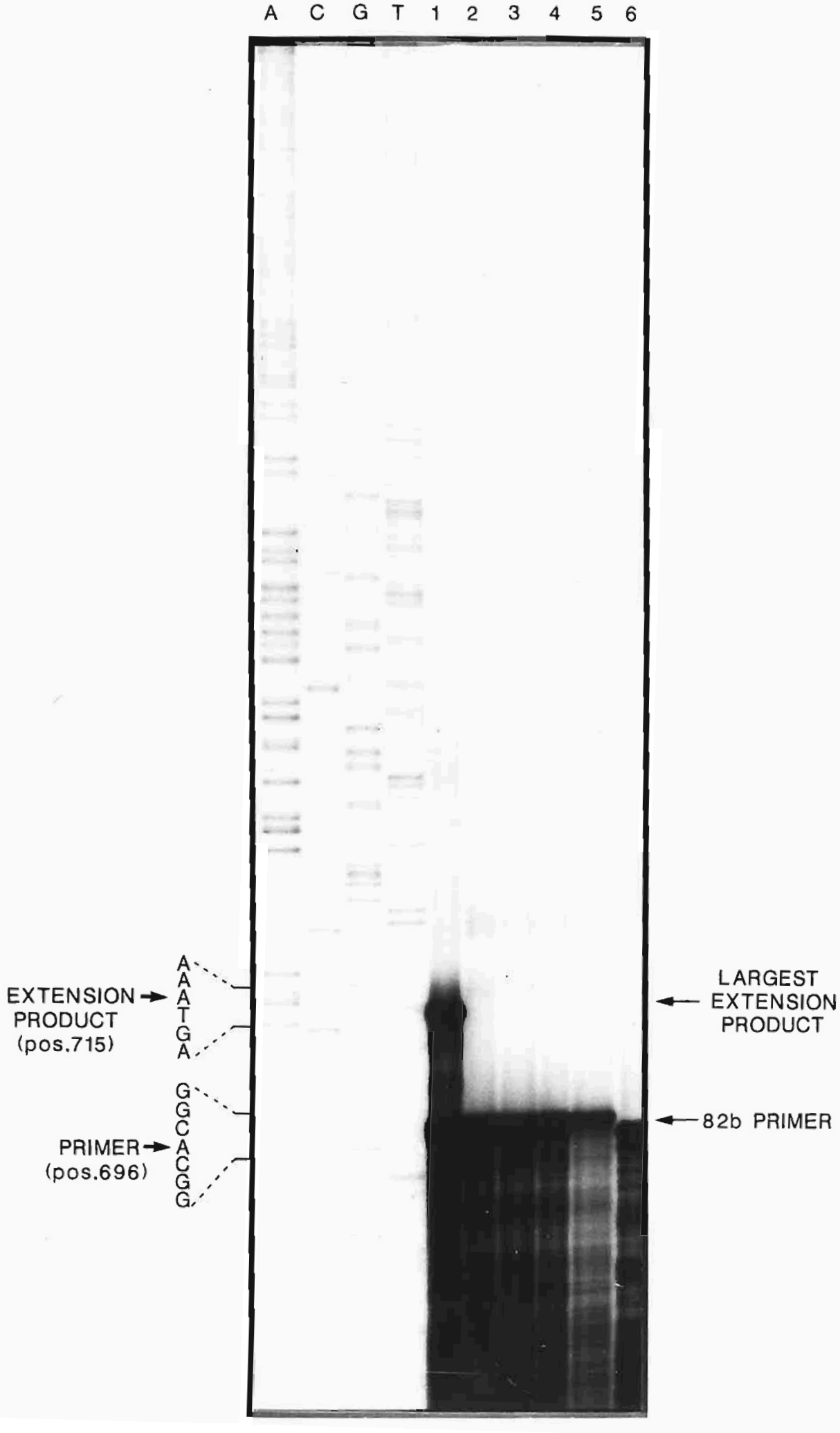
3. Primer extension of RNA isolated 35 min after mock heat-induction of a non-lysogen.

4. Primer extension of RNA isolated from a non-induced 186 *cIts* lysogen.

5. Primer extension of the 82 b *Oxa*NI-*Hae*II primer in the absence of RNA.

6. Identical to track 1 except that AMV reverse transcriptase was omitted from the reaction mix.

A sequencing ladder was included to provide size markers. The sequencing ladder was generated from M13 clone mEC411 (a M13mp18 clone of the *Pst*I-*Eco*RI (12.6% to 13.3%) fragment, blunted and cloned into the *Sma*I site; Section 2.4.2) and corresponds as read to the *l*-strand of the sequence in Figure 3.4(a). As shown on the left of the Figure, the 82 b primer runs at a position corresponding to sequence position 696 (Figure 4.1). The extension product runs at sequence position 715, giving an extension product of 19 bases.


A C G T 1 2 3 4 5 6
EXTENSION PRODUCT (pos.715)
A
A
A
T
G
A
PRIMER (pos.696)
G
G
C
A
C
G
G
LARGEST EXTENSION PRODUCT
82b PRIMER

No extension products were observed for extension on RNA isolated 35 minutes after heat-induction of a 186 *cIts Bam*17 lysogen in the non-permissive host (Figure 4.2, track 2), even upon longer exposure of the autoradiograph (data not shown), and demonstrated that the appearance of late transcription from promoter p_V was dependent upon the protein product of the *B* gene, as was shown for late gene transcription by Finnegan and Egan (1981). The absence of detectable transcription from p_V by primer extension on RNA isolated from a heat-induced 186 *cIts Bam*17 prophage was not a result of a reduction in template number, since 186 DNA replication in a *Bam*17 mutant upon prophage induction is elevated approximately two-fold with respect to the wild-type 35 minutes after prophage induction (Osborne, 1986).

RNA isolated from a non-induced 186 *cIts* lysogen also showed no products upon extension with reverse transcriptase (Figure 4.2, track 4), even upon longer exposure (data not shown), consistent with the absence of late transcription in the prophage state as shown by Finnegan and Egan (1981). However, the level of transcription from the non-induced lysogen is expected to be significantly reduced due to the single copy prophage template, and the inability to detect late transcription from the prophage could reflect this fact.

Verification that the 5'-end of p_V mapped by primer extension was a result of initiation of transcription rather than a result of post-transcriptional processing, by demonstration that the terminal nucleotide is able to be labelled with vaccinia virus guanylyl transferase (Martin and Moss, 1975; Haynes and Rothman-Denes, 1985), was considered unnecessary as the 5'-end mapped by primer extension was consistent with that predicted (Section 3.2.5(b)). The smaller extension products observed upon extension on late RNA after 186 *cIts* prophage induction (Figure 4.2, track 1) were considered to be most likely a result of pausing of reverse transcriptase at secondary structures in the RNA (McKnight *et al.*, 1981) or extension on degradation or processing products of the primary transcript, rather than 5'-ends due to alternative initiation points.

4.2.2 Confirmation of the *in vivo* activity of the leftward late promoter p_{12}.

To confirm that the leftward late promoter p_{12} predicted from the DNA sequence of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region was active *in vivo* (Section 3.2.5(b)), primer extension analysis (Section 2.33) with a radioactive primer specific for leftward message,

was used to detect and map the 5'-ends of leftward transcripts initiating in the region predicted to contain promoter p_{12}.

The leftward promoter p_{12} was predicted to initiate transcription at sequence coordinate 798 (Figure 4.1; Section 3.2.5(b)). The *l*-strand of the 97 base *Bss*HII-*Pst*I (positions 579-675; Figure 4.1) restriction fragment was chosen as an appropriate primer to map the extension products of any leftward transcripts initiating in the region of promoter p_{12}. This primer was prepared by synthesizing a radioactive, complementary copy (Section 2.30.2) of the *r*-strand M13 clone mEC406 (a *r*-strand M13mp18 clone of the *Xmn*I-*Eco*RI (11.3% to 13.3%) fragment, blunted and cloned into the *Sma*I site; Section 2.4.2) followed by restriction with *Bss*HII and *Pst*I and fractionation on a 5% non-denaturing polyacrylamide gel (Section 2.26.2(a)).

The 97 bp *Bss*HII-*Pst*I restriction fragment with the *l*-strand labelled, was heat-denatured (Section 2.32.5) and used in an extension reaction (Section 2.33) on RNA isolated from a heat-induced 186 *cIts* lysogen of the non permissive strain E251 (35 minutes after induction), a heat-induced 186 *cIts Bam*17 lysogen of strain E251 (35 minutes after induction), a non-induced E251 186 *cIts* lysogen, and the corresponding non-lysogen 35 minutes after mock heat-induction. The products were analysed on a 6% denaturing gel, using a sequencing ladder to size the extension products. The results are shown in Figure 4.3.

Extension products were observed on extension on late RNA isolated after heat-induction of a 186 *cIts* prophage (Figure 4.3, track 1). The largest and predominant extension product, a band sized at 123 bases larger than the unextended primer by comparison with a sequencing ladder of known DNA sequence, corresponded to a 5'-end (from the *Pst*I site of the primer at sequence position 675) at sequence position 798, which was identical with that predicted from homology with the transcription initiation sites of the P2 late promoters (Section 3.2.5(b)).

The absence of extension products for extension on RNA isolated 35 minutes after mock heat-induction of a non-lysogen (Figure 4.3, track 4) and on extension on RNA isolated 35 minutes after heat-induction of the corresponding 186 *cIts Bam*17 lysogen (Figure 4.3, track 2), even upon longer exposure of the autoradiograph (data not shown), demonstrated that the transcription detected was phage specific and dependent upon the

Figure 4.3 Confirmation of the *in vivo* activity of the late promoter p_{L2} by primer extension.

RNA was prepared from cultures of E252 [E251(186 *cIts*)], E264 [E251(186 *cIts Bam*17)] and E251, 35 minutes after transferral from 30°C to 39°C as described in Section 2.32. RNA was also prepared from a non-induced culture of E252 at 30°C. The extension reaction was based on the method of McKnight *et al.* (1981) and is described in Section 2.33. The 97 b *Bss*HII-*Pst*I (sequence coordinates 579 to 675) primer specific for leftward message was constructed by extension (Section 2.30.2) on the *r*-strand M13 clone mEC406 (Section 2.4.2), followed by restriction with *Bss*HII and *Pst*I and fractionation on a 5% non-denaturing polyacrylamide gel (Section 2.26.2(a)). This primer (2.0×10^5 cpm per reaction) was denatured, annealed to 5 ug of *in vivo* RNA and extended with AMV reverse transcriptase in the presence of all four dNTPs as described in Section 2.33. Electrophoresis was as described in Section 2.26.2(b). The autoradiograph was exposed at room temperature for 16 hours.

Gel Track

1. Primer extension of RNA isolated 35 min after heat-induction of a 186 *cIts* lysogen.

2. Primer extension of RNA isolated 35 min after heat-induction of a 186 *cIts Bam*17 lysogen.

3. Primer extension of RNA isolated from a non-induced 186 *cIts* lysogen.

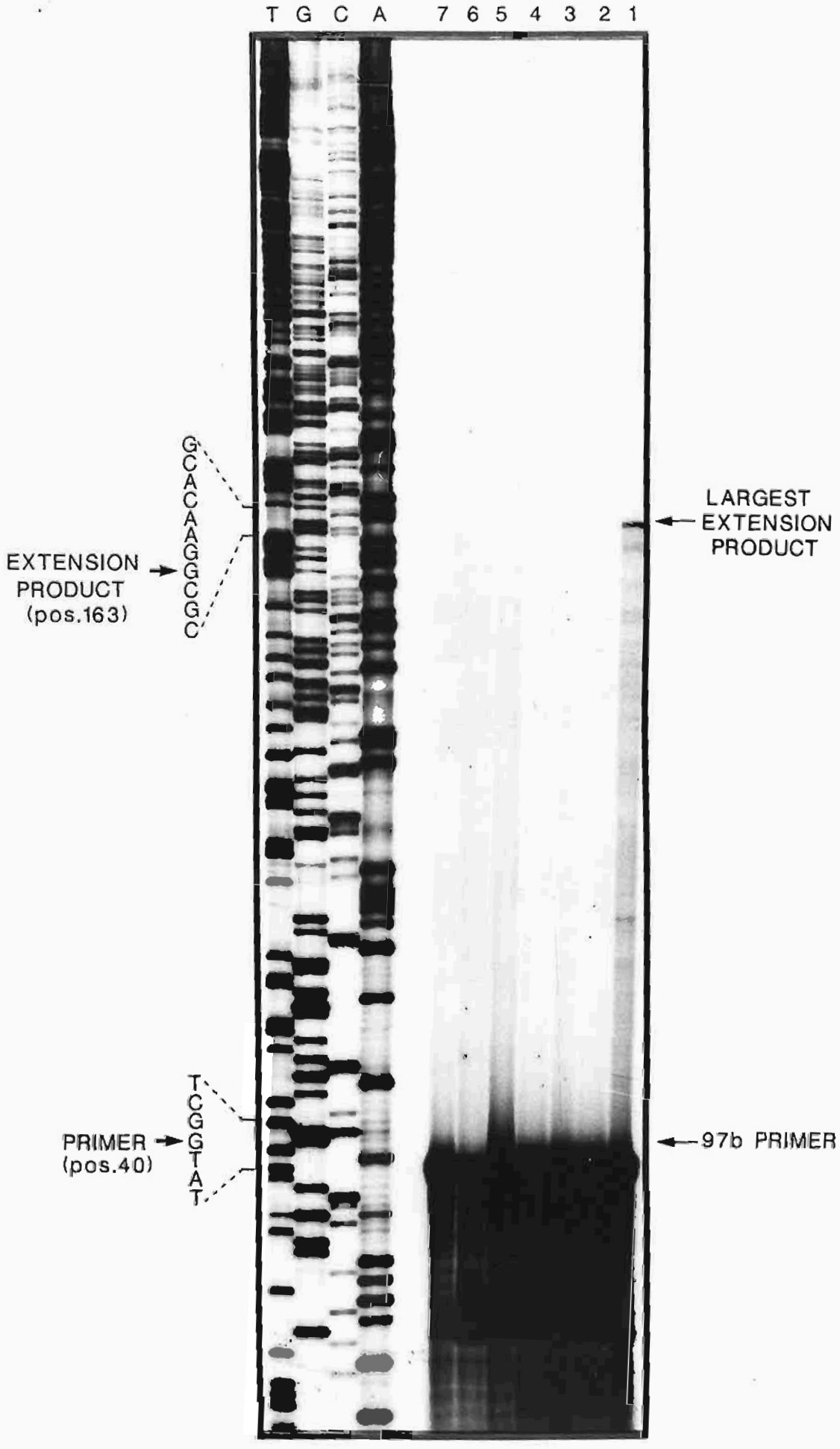
4. Primer extension of RNA isolated 35 min after mock heat-induction of a non-lysogen.

5. Identical to track 1 except that AMV reverse transcriptase was omitted from the reaction mix.

6. Primer extension of the 97 b *Bss*HII-*Pst*I primer in the absence of RNA.

7. Primer alone.

A sequencing ladder was included to provide size markers. The sequencing ladder was generated from M13 clone mEC407 (a M13mp18 clone of the 393 bp *Bam*HI-*Xmn*I (10.0% to 11.3%) fragment) and corresponds as read to the *l*-strand of the sequence in Figure 3.4(a). As shown on the left of the Figure, the 97 b primer runs at a position corresponding to sequence position 40 (Figure 4.1). The extension product runs at sequence position 163, giving an extension product of 123 bases.


T G C A 7 6 5 4 3 2 1
GCACAAGGCGC
EXTENSION PRODUCT (pos.163)
LARGEST EXTENSION PRODUCT
TCGGTAT
PRIMER (pos.40)
97b PRIMER

protein product of the *B* gene. As demonstrated for transcription from p_V (Section 4.2.1), no extension products were observed on extension with reverse transcriptase on RNA isolated from a non-induced 186 *cIts* lysogen (Figure 4.3, track 4), even upon longer exposure (data not shown). This requirement for the *B* gene product and the absence of detectable transcription in the prophage state were consistent with the results of Finnegan and Egan (1981).

Since the 5'-end mapped by primer extension analysis corresponded to that predicted for initiation of transcription from p_{12}, it appeared that the 5'-end mapped by primer extension was a result of initiation of transcription rather than a result of post-transcriptional processing, and demonstration that the 5'-end was due to initiation (as described in Section 4.2.1) was considered unnecessary. The smaller extension products observed upon extension on late RNA after heat-induction of a 186 *cIts* prophage (Figure 4.3, track 1) were considered to be most likely a result of pausing of reverse transcriptase at secondary structures in the RNA (McKnight *et al.*, 1981) or extension on degradation or processing products of the primary transcript, rather than 5'-ends due to alternative initiation points.

It was concluded that both late promoters p_V and p_{12} were active *in vivo* late after prophage induction, and that transcription from both promoters was dependent upon the *B* gene product, consistent with the requirement of the *B* gene product for late gene transcription as demonstrated by Finnegan and Egan (1981). The *B*-dependent positive control of transcription from both late promoters p_V and p_{12} demonstrated by primer extension analysis was considered unlikely to be due to a *B*-dependent mechanism of antitermination of a constitutive leader transcript or a mechanism of *B*-dependent stablilization of a constitutive transcript. The DNA sequence preceding the transcription startpoint for both late promoters showed little homology to the *E. coli* promoter consensus sequence (Section 3.2.5(b)) and indicated positive control of transcription, rather than expression of a constitutive transcript, which is antiterminated or stabilized by B protein late in lytic infection.

It is interesting to note the large difference in the relative proportions of the extension products from promoters p_V and p_{12} obtained with each of the primers used in the primer extension analysis. Obviously the relative levels of the transcripts (as reflected

by the intensities of the extension products) are not directly comparable due to the different primers used and the differences in the extent of extension with reverse transcriptase. However, the significantly greater proportion of transcription from p_V detected, in comparison to transcription detected from p_{12}, would indicate that p_V is considerably more active than p_{12} *in vivo* late in lytic infection, presuming that a difference in messenger stability was not the reason for the observed difference.

4.2.3 Fusion of the late promoters to the *galK* gene.

The late promoters encoded on the *Bam*HI-*Eco*RI (10.0% to 13.3%) region have been shown by primer extension analysis to be active *in vivo* late in lytic infection and to be dependent on the function of the *B* gene for transcription to occur (Section 4.2.1; Section 4.2.2). To identify the functions required for late transcription, the approach taken was to clone the late promoters p_V and p_{12} so as to direct transcription of a reporter gene and then determine the phage functions necessary to achieve transcription of the reporter gene. This method has been used to identify and determine the requirements for activation of transcription in several bacteriophage systems (Guidolin *et al.*, 1989; Margolin *et al.*, 1989)

The level of production of the *galK* gene product in strains containing *galK* vectors (pKO vectors) has been shown to be directly proportional to the *in vivo* strength of several well characterized promoters (McKenney *et al.*, 1981). Expression of the *galK* gene is relatively independent of the RNA structure or stability of the RNA upstream of the gene (McKenney *et al.*, 1981) and therefore allows comparisons to be made between the fusion of different regulatory regions to the *galK* gene. The pKO vectors have been designed to minimize the potential effects of upstream translation, originating from the inserted DNA and reading into the *galK* gene and thereby exerting different effects on *galK* translation efficiency, by the construction of translation stop codons in all three reading frames in the region between the cloning sites of the plasmid and the *galK* coding sequence.

A 490 bp *Alu*I (sequence positions 648-1137) restriction fragment spans the sites of initiation of transcription from both late promoters p_V and p_{12} encoded on the *Bam*HI-*Eco*RI (10.0% to 13.3%) region (Figure 4.1). The rightward late promoter p_V is contained on this *Alu*I restriction fragment with 229 bp downstream of the transcription initiation site

and 261 bp upstream. The leftward late promoter p_{12} is contained on the *Alu*I restriction fragment with 339 bp upstream and 151 bp downstream of the transcription initiation site. In addition, two of the *B boxes* downstream of promoter p_V (positions 1020-1029 and 1109-1118 on the *l*-strand, Figure 4.1; Section 3.2.7) were also contained on this 490 bp *Alu*I restriction fragment. Since the activated promoters so far characterized have not included sequences more than 100 bases from the transcription initiation site (Raibaud and Schwartz, 1984; Busby, 1986), this 490 bp *Alu*I restriction fragment was considered likely to contain the DNA sequence elements necessary for activation of transcription from both late promoters p_V and p_{12}.

To place the *galK* gene under the transcriptional control of promoters p_V and p_{12}, the 490 bp *Alu*I restriction fragment derived from the M13 clone mEC404 (Section 2.4.2) was isolated by polyacrylamide gel electrophoresis (Section 2.26.2(a)) after restriction and visualization of the fragments by end-labelling (Section 2.28.2). This 490 bp *Alu*I fragment was cloned (Section 2.28.3) into the *Sma*I site of the *galK* promoter analysis plasmid pK02 (de Boer, 1984; Section 2.3.1). Recombinants with the inserted DNA in both orientations were identified by diagnostic restriction analysis (data not shown; Section 2.28.3(c)) and this gave constructs such that the *galK* gene was under the transcriptional control of promoter p_V (p_V*galK*, pEC428) or promoter p_{12} (p_{12}*galK*, pEC438).

4.2.4 Properties of the cloned late promoters.

The initial aim in the establishment of a late promoter assay system was to demonstrate that the late promoters cloned onto the multicopy *galK* plasmid pKO2 showed the properties of phage late promoters.

4.2.4(a) Late promoter p_V.

The p_V*galK* construct pEC428 showed no constitutive galactokinase expression in the absence of phage functions, as demonstrated by its white colony phenotype in a $galE^+T^+K^-$ strain (E864) on MacConkey galactose indicator plates (complementation of the defective host *galK* gene by the plasmid *galK* gene when expressed results in a red colony phenotype on indicator plates) and by assay of galactokinase expression (Section

2.18; Table 4.1, pEC428 0.3 units). The absence of constitutive transcription from the late promoter p_V was consistent with the requirement for accessory phage functions for activation of transcription.

The p_V*galK* construct also showed no galactokinase expression upon introduction into a *galE*$^+$*T*$^+$*K*$^-$ strain (E864) lysogenic for 186 *cIts*, as determined by the white colony phenotype on indicator plates and when assayed (Table 4.1, pEC428 -1.4 units). This was consistent with the absence of late gene transcription in the prophage state as observed by Finnegan and Egan (1981), and demonstrated that the 186 functions expressed in the lysogenic state were not sufficient for activation of late transcription, at least from the p_V promoter cloned onto the multicopy plasmid pKO2.

However, prophage induction by temperature shift of a culture lysogenic for a prophage encoding a temperature sensitive CI repressor, resulted in induction of galactokinase expression from the p_V*galK* plasmid after 25 minutes, and this galactokinase activity steeply increased with continued incubation until lysis at 45 minutes (Table 4.1). This activation of galactokinase expression was *B* gene dependent, as induction of a 186 *cIts Bam*17 prophage in the non-suppressing host E864 gave no significant induction of galactokinase activity (Table 4.1). The activation of the p_V*galK* plasmid was also demonstrated to be specific to the cloned promoter, as induction of a 186 *cIts* lysogen carrying the plasmid pK02 gave no significant induction of galactokinase activity (Table 4.1; pKO2, 5 units 35 min after induction).

To give an indication as to the relative activity of the p_V promoter upon prophage induction, the level of transcription from the p_V promoter on plasmid pEC428 was compared with another well characterized promoter. Plasmid pKL600 (Section 2.3.1) contains the wild-type *lac* promoter in an orientation such that it expresses the *galK* gene of a derivative of plasmid pKO2. In host strain E864 (*galK*$^-$, *lacI*$^+$), plasmid pKL600 gave 92 galactokinase units upon IPTG induction and assayed at 37^{o}C (Table 4.1). It was concluded that the level of p_V transcription, measured in the *galK* promoter-analysis system 45 minutes after prophage induction, was at least as efficient as the induced wild-type *lac* promoter.

Late gene transcription in 186 is also dependent upon the function of replication gene *A* (Finnegan and Egan, 1981), which is *cis*-acting (Hocking and Egan, 1982d). The

TABLE 4.1

Transcriptional activity of late promoter p_V[a]

Plasmid Clone[b]	Conditions		Galactokinase activity[c] non-lysogen	(B^+)	(*Bam*)
p_V*galK*	30°C		0.3±0	-1.4±0.1	-
p_V*galK*	39°C	0 min	-	-1.4±0.4	1.0±0.8
		15 min	-	-1.0±0.5	-
		20 min	-	0.5±0.9	-
		25 min	-	4.4±1.9	-
		30 min	-	47 ±12	2.8±1.1
		35 min	-	72 ±15	-
		45 min	-	110 ±6	3.2±1.0
		60 min	-	126 ±10	-
pKO2	39°C	35 min	-	5.0±0.3	-
p_{lac}*galK*	37°C	1 mM IPTG	92 ±4		

Notes to Table 4.1

a. Non-induced cultures of the non-permissive strain E864 (non-lysogen) or the corresponding 186 *cIts* or 186 *cIts Bam*17 lysogens [(B^+) and (*Bam*)] carrying the p_V*galK* plasmid (pEC428) or pKO2 were grown at 30°C in M63 supplemented media (containing ampicillin) to A_{650}=0.4-0.6. Induced cultures were grown at 30°C to A_{650}=0.2 and shifted to 39°C and samples taken with time. Cultures of strain E864 carrying the p_{lac}*galK* plasmid (pKL600) were grown overnight at 37°C in M63 supplemented media (containing ampicillin) and subcultured into the same media containing 1 mM IPTG (final conc.) and grown at 37°C to A_{650}=0.4-0.6. Lysis of the heat-induced 186 *cIts* cultures occurred at 45 min post induction as determined by A_{650} readings.

b. Plasmids and plasmid clones are described in detail in Section 2.3.

c. The galactokinase activity was determined as described in Section 2.18 and are presented as the mean of at least three separate assays presented with the standard deviations. Negative units are a result of regression analysis giving a negative slope for the change in bound counts per time.

ability of the p_V*galK* plasmid to be activated *in trans* upon prophage induction demonstrated that there was no direct requirement for phage replication *in cis* in the activation of the p_V promoter cloned onto the plasmid pKO2. However, this may have been a reflection of the plasmid providing the topological equivalent of a phage replicating template (discussed in detail in Chapter 6). Activation of bacteriophage T4 late transcription requires concurrent viral DNA replication (Riva *et al.*, 1970) and is uncoupled from this requirement when the late promoters are present on a plasmid template (Albright *et al.*, 1988; Jacobs and Geiduschek, 1981).

The p_V promoter on the plasmid pEC428 was concluded to be competent for activation, under positive control by the B protein and to display kinetics consistent with those observed by Finnegan and Egan (1981) for late gene transcription. Thus the p_V*galK* system was chosen as an authentic assay for activation of a 186 late promoter.

4.2.4(b) Late promoter p_{12}.

The p_{12}*galK* construct, plasmid pEC438, showed no constitutive galactokinase expression in the absence of phage functions (Table 4.2, pEC438 0.4 units), as demonstrated for the p_V*galK* construct (Section 4.2.4(a)). Similarly, introduction of plasmid pEC438 into a 186 *cIts* lysogen also showed no activation of galactokinase activity (Table 4.2, pEC438 -2.5 units). However, heat-induction of a 186 *cIts* lysogen carrying the p_{12}*galK* plasmid gave only a small induction of galactokinase activity above that of induction of a 186 *cIts Bam*17 prophage in the non-suppressing host (Table 4.2), in contrast to activation of galactokinase expression upon heat induction of a lysogen carrying the p_V*galK* plasmid (Section 4.2.4(a)).

It was conceivable that the 490 bp *Alu*I restriction freagment used to create the p_{12}*galK* construct pEC438 did not contain all the regulatory signals in the DNA required for activation of promoter p_{12}. To assess this possibility, the entire *Bam*HI-*Eco*RI (10.0% to 13.3%) restriction fragment from 186 *cIts* DNA was cloned (Section 2.28.3) into the *Sma*I site of plasmid pK02 (de Boer, 1984), after end-filling of the terminii of the insert (Section 2.28.2), in an orientation such that the *galK* gene was under the transcriptional control of promoter p_{12}, as determined by diagnostic restriction analysis (data not shown; Section 2.28.3(c)). This construct, plasmid pEC439, with 564 bp of DNA sequence

TABLE 4.2

Transcriptional activity of late promoter p_{12}[a].

Plasmid clone[b]	Conditions		Galactokinase activity[c] non-lysogen	(B^+)	(*Bam*)
$p_{12}galK$	30°C		0.4±0.2	-2.5±0.3	-
$p_{12}galK$	39°C	0 min	-	-2.2±0.4	0.1±0.2
		10 min	-	-1.8±0.3	-
		20 min	-	-1.3±0	-
		25 min	-	0.4±0.2	-
		30 min	-	-0.2±0.1	-
		35 min	-	1.2±0.3	0.3±0.1
		40 min	-	2.8±0.6	0.6±0.1
$p_{12}galK$ (pEC439)	39°C	35 min	-	2.5±0.3	0.5±0.1

Notes to Table 4.2

a. Non-induced cultures of the non-permissive strain E864 (non-lysogen) or the corresponding 186 *cIts* or 186 *cIts Bam*17 lysogens [(B^+) and (*Bam*)] carrying the $p_{12}galK$ plasmids (pEC438 or pEC439) were grown at 30°C in M63 supplemented media (containing ampicillin) to A_{650}=0.4-0.6. Induced cultures were grown at 30°C to A_{650}=0.2 and shifted to 39°C and samples taken with time.

b. The $p_{12}galK$ plasmid pEC438 contains the 490 bp *Alu*I sequence coordinates 648-1137) with the *galK* gene under the transcriptional control of p_{12}. The $p_{12}galK$ plasmid pEC439 contains the entire *Bam*HI-*Eco*RI (10.0% to 13.3%) region with the *galK* gene under the transcriptional control of p_{12}. Both plasmids are described in detail in Section 2.3.2.

c. The galactokinase activity was determined as described in Section 2.18 and are presented as the mean of at least three separate assays presented with the standard deviations. Negative units are a result of regression analysis giving a negative slope for the change in bound counts per time.

upstream of the mapped transcription initiation site for p_{12} and 797 bp downstream, gave only a small activation of galactokinase activity assayed 35 minutes after heat-induction of a 186 *cIts* lysogen (Table 4.2, pEC439 2.5 units), above that for the same construct 35 minutes after heat-induction of a 186 *cIts Bam*17 lysogen (Table 4.2, 0.5 units). It was concluded that the inability to activate promoter p_{12} significantly was unlikely to be due to the absence of any regulatory elements required for transcription. Additionally, it was considered unlikely that the poor activation of promoter p_{12} on pEC438 was due to transcription termination within the cloned restriction fragment, since no structures consistent with *rho*-independent or *rho*-dependent terminators were found in the region of the DNA expected to be transcribed from promoter p_{12} (Section 3.2.8).

Consistent with the poor activation of the late promoter p_{12} on constructs pEC438 and pEC439, primer extension analyses indicated that the p_{12} promoter appeared to be considerably less active *in vivo* late in lytic infection than the divergent promoter p_V (Section 4.2.2). This is contrary to the fact that the corresponding late promoter p_P of P2 appeared to be as active as the divergent promoter p_O late in lytic infection, as determined by S1 protection studies (Christie and Calendar, 1983). The 186 late promoter p_{12} differs from the corresponding P2 late promoter p_P in two respects. Firstly, the conserved sequence in the -35 region of the 186 late promoter p_{12} (5'-TAGCGA-3'; Section 3.2.6), which is identical to that identified in the corresponding promoter p_P of P2, is shifted one base pair upstream of that of the p_P promoter and is the furthest of these conserved sequences in the -35 region from the -10 regions of the late promoters of 186 and P2 (Figure 3.9). It is tempting to speculate whether removal of 1 bp in the intervening sequence may increase its activity *in vivo* to a level comparable to p_V. Secondly, the P2 p_P promoter contains two copies of the *B box* sequence immediately downstream of the transcription initiation site (Section 3.2.7), while no sequence with homology to the *B box* was found downstream of the 186 late promoter p_{12} (Section 3.2.7). Whether the absence of the *B box* sequence in p_{12} is responsible for the low level of activation of transcription remains to be determined.

Alternatively, the low activity of the leftward promoter p_{12} in comparison to the divergent promoter p_V may be a result of the promoter p_{12} competing inefficiently for the components required for activation of transcription with the rightward promoter p_V. The

two promoters do share the same inverted repeat element located at position -55 in both promoters (Section 3.2.6) and this does suggest that this sequence element is utilized by both promoters. The activities of several promoters in divergent promoter systems are increased upon inactivation of one of the promoters (reviewed by Beck and Warren, 1988), and inactivation of the p_V promoter may allow efficient transcription from the divergent promoter p_{12}.

It therefore appeared that the leftward late promoter p_{12} was only poorly activated in the *galK* promoter analysis system and this appeared to be consistent with the low level of transcription detected from this promoter by primer extension analysis (Section 4.2.2). For this reason, the p_{12} promoter in the *galK* system was considered unsuitable for determining the requirements for activation of a 186 late promoter. However, the p_V*galK* construct pEC428 was strongly activated upon prophage induction, had the characteristics of a phage late promoter, and was chosen as the appropriate construct to determine the requirements for activation of a 186 late promoter.

4.2.5 Investigation of the role of the *B boxes* in activation of late transcription.

4.2.5(a) Effect of a 3'-deletion on activation of promoter p_V.

In the region of the *V* gene cloned into the *galK* plasmid to create the p_V*galK* construct, there were two copies of the highly conserved ten nucleotide sequence termed the *B box* (Section 3.2.7; Figure 4.1). This sequence is present as one arm of two inverted repeats found overlapping the *B* promoter (Kalionis *et al.*, 1986b) and is also present downstream of the transcription initiation site of three of the four P2 late promoters (Christie and Calendar, 1985). Since the B protein is involved in repressing its own transcription (Kalionis *et al.*, 1986b) and also in the control of late gene transcription (Finnegan and Egan, 1981; Section 4.2.1; Section 4.2.2; Section 4.2.4(a)), the *B box* may represent a DNA motif involved in the control of transcription at the late promoters and repression at the *B* promoter (Kalionis *et al.*, 1986b). Specifically, this sequence could have a role in the activation of transcription from the late promoters in which it is present or alternatively, it was conceivable that the *B box* could be involved in repressing

transcription from the late promoters in a 186 lysogen, since the same sequence may be involved in repression of the *B* promoter in a 186 lysogen (Kalionis *et al.*, 1986b).

As a first step to determining whether the *B box* sequences were involved in control of transcription from p_V, it was decided to determine if the DNA downstream of the late promoter p_V had any role in the control of transcription from the p_V promoter. To assess if downstream DNA sequences were involved, a deletion derivative of the p_V*galK* plasmid extending from the +18 position was constructed. The *Oxa*NI restriction site (position 924; Figure 4.1) in the cloned 490 bp *Alu*I fragment and the *Hin*dIII site in the polylinker of plasmid pEC428 were utilized to create a 3'-deletion of the p_V promoter in the p_V*galK* construct pEC428. The 3'-deleted plasmid pEC440 was constructed after restriction of plasmid pEC428 with *Oxa*NI and *Hin*dIII (Section 2.25), end-filling of the termini of the large fragment (Section 2.28.2) and subsequent ligation (Section 2.28.3(a)). Correct end-filling and religation resulted in the regeneration of the *Hin*dIII site (and loss of the *Oxa*NI site), as determined by restriction analysis (data not shown) and resulted in the construction of a *galK* plasmid fused to late promoter p_V, with DNA sequence downstream of the +18 position deleted.

Plasmid pEC440 showed no constitutive galactokinase expression in the absence of phage functions (Table 4.3, 0.4 units). Upon introduction into a 186 *cIts* lysogen there was also no evidence of galactokinase expression (Table 4.3, 0.6 units) and demonstrated that it was unlikely that there were any downstream elements which inhibited transcription from the late promoter in the lysogenic state. However, upon induction of a 186 *cIts* prophage by heat, pEC440 showed a significant decrease in galactokinase activity (Table 4.3; 40 minutes, 9.1 units) as compared to the original p_V*galK* plasmid, assayed at the same time after prophage induction (Table 4.3; 40 minutes, 73 units). As expected, the activation of plasmid pEC440 remained *B* gene dependent, as there was no significant activation upon induction of a 186 *cIts Bam*17 lysogen (Table 4.3; 40 minutes, 0.8 units).

This reduction in galactokinase activity was considered unlikely to be due to the creation of a termination structure upon fusion of the *Oxa*NI and the *Hin*dIII sites as no significant secondary structures were found in the DNA sequence downstream of the p_V promoter, created as a result of the fusion of *Oxa*NI and *Hin*dIII sites in the deletion plasmid pEC440, using the program COMSTR (Section 2.34.3). Therefore, deletion of the

TABLE 4.3

Transcriptional activity of a downstream deletion derivative of late promoter p_V[a].

Plasmid clone[b]	Conditions	Galactokinase activity[c] non-lysogen	(B^+)	(*Bam*)
p_V*galK*	30°C	0.3±0	-1.4±0.1	-
p_V*galK*	39°C 40 min	-	73 ±10	-
p_V*del*3'*galK*	30°C	0.4±0.3	0.6±0.1	-
p_V*del*3'*galK*	39°C 40 min	-	9.1±2.7	0.8±0.3

Notes to Table 4.3

a. Non-induced cultures of the non-permissive strain E864 (non-lysogen) or the corresponding 186 *cIts* or 186 *cIts Bam*17 lysogens [(B^+) and (*Bam*)] carrying the p_V*galK* plasmid (pEC428) or the p_V*del*3'*galK* plasmid (pEC440) were grown at 30°C in M63 supplemented media (containing ampicillin) to A_{600}=0.4-0.6. Induced cultures were grown at 30°C to A_{650}=0.2 and shifted to 39°C and samples taken with time.

b. The p_V*galK* plasmid pEC428 is described in detail in Section 2.3.2. The p_V*del*3'*galK* plasmid (pEC440) contains DNA deleted downstrean of the +18 position of the 186 late promoter p_V and is described in Section 2.3.2.

c. The galactokinase activity was determined as described in Section 2.18 and are presented as the mean of at least three separate assays presented with the standard deviations. Negative units are a result of regression analysis giving a negative slope for the change in bound counts per time.

DNA downstream of the +18 position appeared to reduce the level of transcription initiating from p_V and indicated the existence of downstream sequence elements that were involved in activation of transcription of p_V. To date, no sequence elements downstream of the transcription initiation site in activated promoters have been identified (Raibaud and Schwartz, 1984; Busby, 1986) and this represents a potentially novel situation in the activation of prokaryotic transcription.

4.2.5(b) Fusion of *B box* deletion derivatives to the *galK* gene.

The sequence of the *B box* is present twice in the DNA sequence of the 490 bp *Alu*I restriction fragment used to create the p_V*galK* plasmid pEC428 (Figure 4.1; on the *l*-strand at positions 1020-1029, 1109-1118). Deletion of the DNA downstream of the +18 position (pEC440) demonstrated a significant reduction in galactokinase expression as compared to the original p_V*galK* construct late after prophage induction (Section 4.2.5(a)). To determine if the *B boxes* had any role in the activation of transcription from p_V, derivatives of the p_V promoter with specific deletions in either or both B boxes present were constructed. Deletions that maintained the reading frame of the *V* gene were chosen to avoid any alteration to translation of the inserted DNA, since change in the frame of translation could conceivably result in *rho*-dependent termination of transcription within the cloned restriction fragment.

Site-directed oligonucleotide mutagenesis (Section 2.31) of the M13 clone mEC405 (Section 2.4.2; a *r*-strand clone of the 490 bp *Alu*I restriction fragment cloned into the *Sma*I site of M13mp18) using the 30-mer oligonucleotide 620 (Section 2.5) created clone mEC408, with an in-frame 12 bp deletion of the first *B box* (deletion positions 1018-1029; Figure 4.1) and mutagenesis using the second 30-mer oligonucleotide 619 (Section 2.5) created clone mEC409 with an in-frame 12 bp deletion of the second *B box* (deletion positions 1107-1119; Figure 4.1). The deletions were confirmed by DNA sequence analysis (Section 2.29) using the 17-mer oligonucleotide 621 (Section 2.5) as a primer to confirm the DNA sequence of the first deletion mutant mEC408 (data not shown), and by the use of the first 30-mer oligonucleotide 620 as a primer to confirm the DNA sequence of the second deletion mutant mEC409 (data not shown). Subsequent oligonucleotide mutagenesis of clone mEC408 with the second 30-mer oligonucleotide

619 allowed the creation of mutant mEC410 with both *B boxes* deleted, the DNA sequence of which was confirmed using the 17-mer (oligo. 621) and first 30-mer (oligo. 620) oligonucleotides as sequencing primers (data not shown).

The original M13mp18 clone mEC405 was spanned by *Eco*RI and *Hin*dIII sites in the polylinker, and these sites were utilized to clone the p_V promoter and the deleted *B box* derivatives into the *Eco*RI and *Hin*dIII sites of plasmid pK02 (Section 2.28.3), in an orientation such that the promoter p_V transcribed the *galK* gene. The 541 bp *Eco*RI-*Hin*dIII fragments were isolated by polyacrylamide gel electrophoresis (Section 2.26.2(b)) and cloned into the *Eco*RI and *Hin*dIII sites of plasmid pKO2 to yield plasmids pEC441 (p_V*galK*), pEC442 (p_V*del1galK*), pEC443 (p_V*del2galK*) and pEC444 (p_V*del1+2galK*) as demonstrated by diagnostic restriction analysis (data not shown; Section 2.28.3(c)).

4.2.5(c) Investigation of the role of the *B boxes* in activation of p_V transcription.

To determine if the *B boxes* downstream of the late promoter p_V had any role in the control of p_V transcription, the galactokinase expression of the *B box* deletion derivatives of the p_V promoter fused to the galactokinase gene (Section 4.2.5(b)) before and after prophage induction was determined.

None of the *B box* deletion plasmids showed any significant galactokinase activity upon introduction into a 186 *cIts* lysogen as determined by their white colony phenotype in a *galK*$^-$ strain on indicator plates and by galactokinase assay (Table 4.4; pEC441 0.3 units; pEC442 0.6 units; pEC443 1.6 units; pEC444 0.7 units). This demonstrated that the *B boxes* had no role in acting to inhibit transcription from p_V in the lysogenic state.

Upon prophage induction by heat, each of the *B box* deletion derivatives assayed 30 minutes after prophage induction gave equivalent levels of galactokinase activity to the wild-type p_V*galK* plasmid pEC441 (Table 4.4, pEC441 47 units; pEC442 46 units; pEC443 48 units; pEC444 46 units). Therefore, the *B boxes* appeared to have no role in the activation of p_V transcription in the *galK* promoter analysis system, presuming that the nature of the plasmid template did not decouple the normal requirement for these sequence elements, and it appeared that they had no role in the activation of p_V transcription.

The absence of any effect on activation of transcription from p_V also indicated that the reduction in activation of p_V transcription seen for the 3'-deletion plasmid pEC440 was

TABLE 4.4

Transcriptional activity of *B box* deletion derivatives of late promoter p_V[a].

Plasmid clone[b]	Conditions	Galactokinase activity[c] (B^+)
$p_V$galK	30°C	0.3±0.1
$p_V$galK	39°C 30 min	47 ±4.8
$p_V$del1galK	30°C	0.6±0.1
	39°C 30 min	46 ±6.5
$p_V$del2galK	30°C	1.6±0.3
	39°C 30 min	48 ±2.0
$p_V$del1+2galK	30°C	0.7±0.3
	39°C 30 min	46 ±4.2

Notes to Table 4.4

a. Non-induced cultures of the 186 *cIts* lysogen [(B^+)] of the non-permissive strain E864 carrying the *$p_V$galK* plasmid (pEC441), the *$p_V$del1galK* plasmid (pEC442), the *$p_V$del2galK* plasmid (pEC443) or the *$p_V$del1+2galK* plasmid (pEC444) were grown at 30°C in M63 supplemented media (containing ampicillin) to A_{650}=0.4-0.6. Induced cultures were grown at 30°C to A_{650}=0.2 and shifted to 39°C and samples taken with time.

b. The *$p_V$galK* plasmid (pEC441), the *$p_V$del1galK* plasmid (pEC442) with the first *B box* downstream of p_V deleted, the *$p_V$del2galK* plasmid (pEC443) with the second *B box* downstream of p_V deleted and the *$p_V$del1+2galK* plasmid (pEC444) with both *B box* elements downstream of p_V deleted, are described in detail in Section 2.3.2.

c. The galactokinase activity was determined as described in Section 2.18 and are the average of two separate assays.

not due to the removal of the *B boxes*, but rather due to unidentified sequence elements. Identification of the sequence elements responsible for the reduction in p_V activity on deletion of sequence from the +18 position, should be possible by the generation of a set of progressive downstream deletions of p_V (prepared, for example by successive deletion with BAL 31 exonuclease as described in Maniatis *et al.*, 1982) fused to the *galK* gene.

4.3 SUMMARY.

The late promoters p_V and p_{12} predicted from the DNA sequence (Section 3.2.5(b)) of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region were shown to be active *in vivo* late in lytic infection by primer extension analysis. The 5'-ends located were consistent with the initiation sites predicted from homology with the P2 late promoters p_O and p_P (Section 3.2.5(b)). Transcription from both promoters was found to be dependent upon the protein product of the *B* gene, as no transcription was detectable from either promoter upon heat induction of a 186 *cIts Bam*17 lysogen. Transcription from p_V and p_{12} was also not detectable by primer extension analysis in an uninduced 186 lysogen. The requirement for the *B* gene product and the absence of late gene transcription in the prophage state were consistent with the results of Finnegan and Egan (1981) for late gene transcription. Primer extension analysis also indicated that the leftward late promoter p_{12} was considerably less active *in vivo* late after prophage induction than the rightward promoter p_V.

In order to identify the functions required for late gene transcription, the late promoters p_V and p_{12} were cloned so as to direct transcription of the reporter gene *galK*. The p_V*galK* construct was not active in the absence of phage functions, demonstrating the requirement for accessory phage functions to achieve activation of transcription. The p_V*galK* construct was also not active in a 186 lysogen, consistent with the absence of late gene transcription of the prophage state as observed by Finnegan and Egan (1981). The p_V*galK* construct demonstrated *B* gene dependent activation of transcription upon prophage induction with kinetics consistent with those observed by Finnegan and Egan (1981) for late gene transcription. Thus the p_V promoter on plasmid pEC428 had all the characteristics of a 186 late promoter and was chosen as an authentic assay for late transcription.

However, the leftward late promoter p_{12} was found not to be significantly active upon prophage induction in the *galK* promoter analysis system. To assess the possibility that the 490 bp *Alu*I (positions 648-1137) fragment did not contain all the regulatory signals in the DNA required for activation of late promoter p_{12}, the entire *Bam*HI-*Eco*RI (10.0% to 13.3%) region was cloned into the *galK* analysis plasmid pK02, with *galK* under transcriptional control of late promoter p_{12}. However this p_{12}*galK* construct also showed no significant improvement in activation. The lack of substantial activation in the *galK* analysis system appeared to be consistent with primer extension studies, which indicated that promoter p_{12} was considerably less active late in infection than promoter p_V.

The role of the *B boxes* in activation of p_V transcription was investigated using the p_V*galK* analysis system. Deletion of the DNA sequence downstream of the +18 position gave a significant reduction in activation of transcription from p_V late after prophage induction. To assess if the loss of the *B boxes* were responsible for this reduction in transcription, deletion derivatives of the p_V*galK* plasmid with either or both *B boxes* were constructed. However, these *B box* deletion derivatives had no effect on activation of transcription from p_V in the *galK* system, and the role of the *B boxes* in control of late gene expression remains to be determined. However, the significant reduction in p_V activity upon deletion of DNA sequences downstream of position +18 indicated the existence of downstream elements involved in the activation of late transcription from this promoter, and represents a potentially novel situation in the activation of prokaryotic transcription.

RESULTS: SECTION II

CHAPTER 5

ROLE OF THE *B* GENE IN THE ACTIVATION OF LATE TRANSCRIPTION IN PHAGE 186

CHAPTER 5. ROLE OF THE *B* GENE IN THE ACTIVATION OF LATE TRANSCRIPTION IN PHAGE 186.

5.1 INTRODUCTION.

The results presented in Chapter 4 demonstrated that the late promoter p_V on the p_V*galK* plasmid pEC428 was activated upon prophage induction, under positive control by the B protein and displayed kinetics consistent with those observed for late gene transcription by Finnegan and Egan (1981). For these reasons, the late promoter p_V on the p_V*galK* plasmid was chosen as an authentic assay for the activation of a 186 late promoter.

One of the major aims of this work was to identify the phage functions required directly for activation of late gene transcription and the approach chosen was to determine the phage functions sufficient to activate galactokinase expression from the p_V*galK* plasmid. Late gene transcription in 186 has been shown to be dependent upon the functions of genes *B* and *A* (Finnegan and Egan, 1981) and a third function, normally under CI repressor control, may also be required for late gene transcription (Section 1.5.7; Kalionis *et al.*, 1986b). Kalionis *et al.* (1986b) postulated that gene *B* encodes a protein directly involved in activating transcription at the late promoters. Gene *A* mutants are defective in phage replication (Hocking and Egan, 1982d; Sivaprasad *et al.*, 1990) and the role of gene *A* in late transcription is expected to be indirect, and to reflect a requirement for increased gene dosage and/or a requirement for a replicating template in the activation of late transcription (Kalionis *et al.*, 1986b).

The first step in identifying the phage functions required directly to activate late transcription, was to determine if the *B* gene product alone would activate galactokinase expression of the p_V*galK* plasmid. Activation of p_V transcription at a level comparable to that assayed late after prophage induction would indicate that the B protein is the only phage function required directly to activate late transcription. Absence of activation (or poor activation) would indicate the requirement of other phage functions in activation of late transcription, which were not identified in the genetic screening for essential genes (Hocking and Egan, 1982c,d).

This chapter describes the construction of a plasmid directing the expression of functional B protein, and the demonstration that the B protein is the only phage factor required for activation of p_V transcription in the *galK* promoter analysis system.

5.2 RESULTS AND DISCUSSION.

5.2.1 Construction of a *B* expression plasmid.

In order to determine if the B protein was the only phage function sufficient to activate galactokinase expression from the p_V*galK* plasmid, it was necessary to supply B protein to the p_V*galK* plasmid and determine if activation of galactokinase expression occurred. This approach required the construction of a plasmid for the expression of functional B protein which was compatible with the p_V*galK* plasmid. A similar strategy has been used in bacteriophage Mu to identify the phage functions required for activation of the late promoters. The Mu C protein has been shown to activate transcription of a Mu late promoter-*lacZ* fusion when supplied *in trans* from a compatible plasmid and was concluded to be the only phage function required for activation of these promoters (Margolin *et al.*, 1989).

In order to supply B protein *in trans* to the p_V*galK* plasmid, it was decided to place the *B* gene under the transcriptional regulation of another, controllable promoter. Transcription of the *B* gene appears to be under negative autogenous control, either directly or indirectly (Kalionis *et al.*, 1986b) and placing the *B* gene under the control of another, controllable promoter would remove the possiblity of direct autogenous control and therefore lowered expression. Placing the *B* gene under the control of another promoter would also allow expression of the *B* gene to be regulated.

Plasmid pEC612 is a vector constructed in this laboratory for the high level expression of native proteins (as opposed to fusion proteins) and is described in Section 2.3.1. Vector pEC612 is derived from the R1-derived replicon pOU61 (Larsen *et al.*, 1984) and therefore should be compatible with the ColE1-related replicon of pKO2 (de Boer, 1984). Transcription of cloned gene products in pEC612 is from the *lacI* repressible, hybrid *trc* promoter (p_{trc}) (Brosius *et al.*, 1985). Efficient translation of cloned gene products is achieved by construction of a *Nde*I site (5'-CATATG-3') at the initiating

methionine of the gene product to be expressed and cloning into the unique *Nde*I site of the vector. A strong ribosome binding site and AT rich spacer direct translation from the cloned initiation codon. Plasmid pEC612 is maintained at one copy per host chromosome at 30°C and displays uncontrolled replication above 42°C. High level expression of cloned gene products is achieved by transferral of the plasmid from 30°C to high temperature (42°C) and addition of the gratuitous inducer IPTG (in a host encoding *lacI* repressor).

To express the *B* gene product in plasmid pEC612, it was necessary to construct a *Nde*I site at the initiating methionine codon. As shown in Figure 5.1, the *B* gene has been identified from the DNA sequence of the *Pst*I-*Xho*I (65.5% to 67.6%) fragment and was predicted to code for a protein of 72 amino acids (Kalionis *et al.*, 1986b). Primer extension analysis located the 5'-end of the *B* gene transcript and translation of the DNA sequence revealed that there is only one in-frame methionine codon upstream of the *Bam*57 mutation encoded on this transcript (Kalionis *et al.*, 1986b). This methionine is preceded by a possible ribosome binding site (Kalionis *et al.*, 1986a) and is most likely to represent the *in vivo* translation initiation codon of the *B* gene (Figure 5.1).

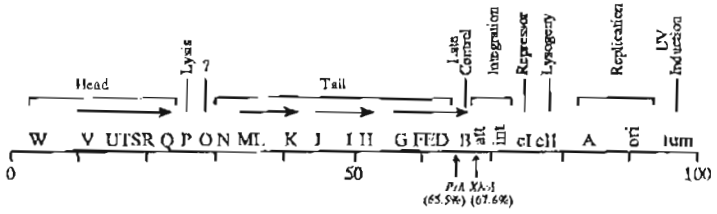
The M13mp19 clones mEC650 and mEC651 contain the 1.9kb *Hin*dIII-*Xho*I (61.3% to 67.6%) restriction fragments, derived from 186 *cIts* and 186 *cIts Bam*17 DNA respectively, with *Nde*I sites created at the predicted initiation codon of the *B* and *Bam*17 genes (Osborne, 1986; Figure 5.1). The *Nde*I sites were created by site-directed oligonucleotide mutagenesis of the corresponding M13mp19 clones mEC648 and mEC649 (Section 2.4.2), containing the 1.9kb *Hin*dIII-*Xho*I (61.3% to 67.6%) restriction fragment, cloned into the *Hin*dIII and *Sal*I sites of the M13 vector (Osborne, 1986). To clone the wild-type *B* gene and the *Bam*17 derivative into plasmid pEC612, the *Sma*I site in the M13 polylinker of clones mEC650 and mEC651 were utilized in conjunction with the *Nde*I site at the predicted initiation codons. Restriction of RF DNA of the M13 clones mEC650 and mEC651 (Section 2.22) with *Nde*I and *Sma*I (Section 2.25) resulted in the production of 355 bp *Nde*I-*Sma*I restriction fragments, which were isolated after non-denaturing gel electrophoresis (Section 2.26.2(a); Section 2.27.2) and subsequently cloned into the *Nde*I site and the *Not*I (end-filled; Section 2.28.2) site of plasmid pEC612 to give the *B* expression plasmid pEC434 ($pp_{trc}B^+$) and the corresponding *Bam*17 expression plasmid

Figure 5.1 Location and DNA sequence of the *Pst*I-*Xho*I (65.5% to 67.6%) region from 186 *cIts*.

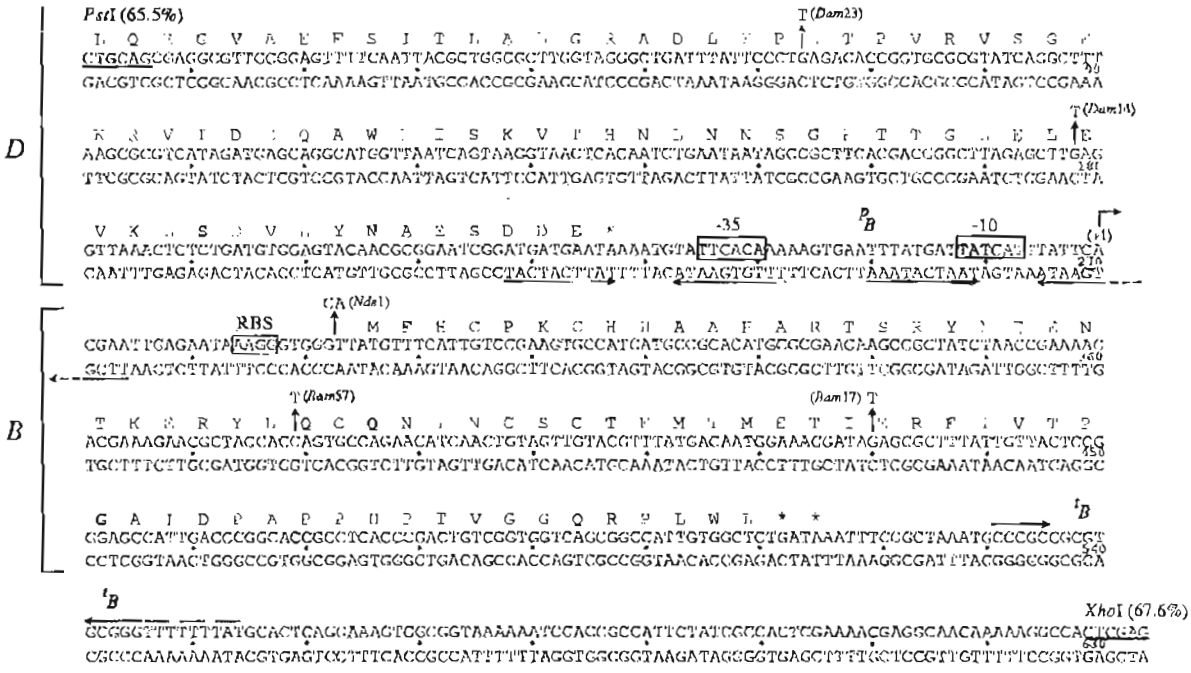
(a) Physical and genetic map of phage 186. The functions of the genes are shown and are as determined by Hocking and Egan (1982c), and the genetic and physical mapping are as previously described (Hocking and Egan, 1982a; Finnegan and Egan, 1979). The location of the *Pst*I-*Xho*I (65.5% to 67.6%) region is shown. Two of the four alleles of the *D* gene (*Dam*23 and *Dam*14) map to the right of the *Pst*I site at 65.5%, as do both alleles of the *B* gene (*Bam*57 and *Bam*17) (Finnegan and Egan, 1979). Transcription units based on polarity effects (Hocking and Egan, 1982c) are represented by arrows. *In vivo* transcription studies indicated that the *B* gene is encoded on the *GFED* transcription unit (Kalionis *et al.*, 1986b).

(b) DNA sequence of the *Pst*I-*Xho*I (65.5% to 67.6%) region from wild-type and from amber mutants in the *B* gene (*Bam*57 and *Bam*17) and in the *D* gene (*Dam*23 and *Dam*14). The open reading-frames of the *D* and *B* genes are shown, as are the base changes associated with the amber mutations in the *D* and *B* genes. The ribosome-binding site (RBS), the 5'-terminus of the *B* transcript (+1) and the associated promoter (*p*) and terminator (*t*) are indicated. The base-pair changes used to create the *Nde*I site at the predicted initiation codon of the *B* gene (Osborne, 1986) are also shown.

(a)


Lysis
?
Late Control
Integration
Repressor
Lysogeny
Replication
CIV Induction
Head
Tail
W V U T S R Q P O N M L K J I H G F E D B att int cI cII A ori tum
0
50
100
PstI (65.5%)
XhoI (67.6%)

(b)


D
B
PstI (65.5%)
T (Dam23)
T (Dam114)
-35
-10
(+1)
CA (NdeI)
RBS
T (Dam57)
(Dam17) T
XhoI (67.6%)

pEC435 ($pp_{trc}Bam$), as determined by diagnostic restriction analysis (data not shown; Section 2.28.3(c)).

5.2.2 Plasmid pEC434 directs the expression of functional B protein.

In order to determine whether the *B* expression plasmid pEC434 produced functional B protein, the ability of the plasmid pEC434 to complement a 186 *cIts Bam*17 phage in infection was tested by efficiency of plating. In a non-suppressing host (E538; Table 2.1) carrying the *B* expression plasmid pEC434, 186 *cIts Bam*17 phage plated with slightly better efficiency than on the isogenic suppressing (*supE*) host E540 (Table 5.1; eop 1.32). However, the efficiency of plating of 186 *cIts Bam*17 phage on host E538 carrying the *Bam*17 expression plasmid pEC435 was less than 1.5×10^{-4} (Table 5.1) with respect to the plating on the isogenic suppressing host E540 and demonstrated that the cloned *B* gene product expressed from plasmid pEC434 was functional in complementation, activating late transcription of the 186 *cIts Bam*17 phage *in trans*.

Expression of B protein from plasmid pEC434 did not appear to be detrimental to normal 186 infection, (as could result if premature expression of the late genes occurred during infection), since wild-type 186 phage plated on E538 carrying plasmid pEC434 at the same efficiency and with the same plaque size as the plating on the isogenic *supE* host E540 (Table 5.1; 1.08). However, both 186 *cIts Bam*17 and wild-type 186 phage demonstrated a distinct clear plaque phenotype on E538 carrying the *B* expression plasmid pEC434, but showed a normal turbid phenotype on the same strain carrying the *Bam* expression plasmid pEC435 (Table 5.1). It appeared that expression of B protein from plasmid pEC434 lowered the frequency of lysogeny of the infecting phage for some unknown reason. As discussed in Section 6.2.1(a), this clear plaque phenotype appears to be a consequence of increased expression of B protein being lethal to a 186 prophage.

The *B* gene product as cloned into plasmid pEC612 should produce a protein of 72 amino acids and a molecular weight of 8.3 kd (Kalionis *et al.*, 1986a). In the non-supp ressing F'*lacI*^{q} host JM105, plasmid pEC434 directed the synthesis of a distinct protein band of approximately 7 kd in size, 2 hours after temperature and IPTG induction (42°C, 1 mM IPTG), which was absent upon induction of the *Bam*17 expression plasmid pEC435 after 2 hours in the non-suppressing host JM105 or upon induction of the parent

TABLE 5.1

Plating efficiencies and phenotypes of 186 wild-type and 186 *Bam* on strains carrying *B* expression plasmid pEC434[a].

Plasmid clone[b]	Efficiency of plating[c]		Plaque size and phenotype	
	186^+	186 *Bam*	186^+	186 *Bam*
$pp_{trc}B^+$	1.08	1.32	large/ clear	large/ clear
$pp_{trc}Bam$	N.D.	$<1.5\times10^{-4}$	large/ turbid	large/ turbid

Notes to Table 5.1

a. Cultures of the non-permissive strain E538 carrying expression plasmids $pp_{trc}B^+$ (pEC434) or $pp_{trc}Bam$ (pEC435) were grown at 30°C in TB broth containing ampicillin to A_{600}=0.8. Dilutions of 186 wild-type (186^+) phage and 186 *cIts Bam*17 (186 *Bam*) phage were plated with cells with each plasmid onto TB plates (Section 2.16.1). Phage 186 strains are described in Section 2.2. Plates were incubated overnight at 30°C and the number of plaques and their phenotypes noted.

b. Plasmid clones $pp_{trc}B^+$ (pEC434) and $pp_{trc}Bam$ (pEC435) are described in detail in Section 2.3.2.

c. The efficiency of plating is expressed as the ratio of the number of plaques obtained on strains carrying the plasmid clones to the number of plaques obtained on the isogenic *supE* strain E540 carrying the parent plasmid pEC612 (Section 2.3.1). N.D. designates not determined.

plasmid pEC612 in the same host (R. Tiwari, personal communication). Therefore, plasmid pEC434 was concluded to direct the expression of a 7 kd protein, which most probably represented the *B* gene product.

5.2.3 Construction of a kanamycin resistant derivative of the p_V*galK* plasmid.

Since the *B* expression plasmid pEC434 was functional in complementing the defective *B* gene of a 186 *cIts Bam*17 phage (and therefore functional *in trans* in activating late gene transcription), it was necessary to determine whether B protein alone supplied from pEC434 was sufficient to significantly activate galactokinase expression from the p_V*galK* plasmid. However, to supply B protein from the plasmid pEC434 to the p_V*galK* plasmid, it was necessary to construct a derivative of the p_V*galK* plasmid pEC428 with a selectable antibiotic marker other than that conferring resistance to ampicillin, since the vector pEC612 and the derivatives pEC434 and pEC435 also conferred resistance to ampicillin. Modification of the antibiotic marker of the p_V*galK* plasmid was chosen, rather than modification of pEC612 and its derivatives, since the entire nucleotide sequence of pEC428 was known (Section 3.2.1; de Boer, 1984) and the genetic content was well established. However, this cloning strategy was dependent on the absence of transcription from the inserted DNA fragment reading into the *galK* gene of the plasmid.

Plasmid pKC7 carries the kanamycin resistance gene from transposon Tn5 (Rao and Rogers, 1979; Auerswald *et al.*, 1981; Mazodier *et al.*, 1985), conveniently bound by *Eco*RI and *Sma*I sites. The 1.3 kb *Eco*RI to *Sma*I restriction fragment of plasmid pKC7 was isolated from plasmid DNA (Section 2.21.2) after restriction (Section 2.25) by agarose gel electrophoresis (Section 2.26.1; Section 2.27.1(b)). Kanamycin resistant derivatives of the vector pKO2 (de Boer, 1984) and the p_V*galK* plasmid pEC428 (Section 2.3.2) were constructed by cloning (Section 2.28.3) the 1.3 kb *Eco*RI to *Sma*I restriction fragment of plasmid pKC7 into the large *Eco*RI to *Xmn*I fragment of pKO2 and pEC428. This yielded plasmid pEC432, an ampicillin sensitive, kanamycin resistant derivative of vector pKO2, and plasmid pEC433, an ampicillin sensitive, kanamycin resistant derivative of plasmid pEC428, the identity and orientation of both plasmids confirmed by diagnostic restriction analysis (data not shown; Section 2.28.3(c)).

Both plasmids pEC432 and pEC433 had a white colony phenotype in the *galK*$^-$ host E864 on MacConkey indicator plates at 37^{o}C, and demonstrated that no transcription originating from the cloned *Eco*RI to *Sma*I restriction fragment from pKC7 was reading into the *galK* gene of the plasmids. Thus plasmids pEC432 and pEC433 were chosen as suitable kanamycin derivatives to determine the requirements for activation from the late promoter p_V.

5.2.4 B protein is the only phage function required for activation of p_V transcription.

To determine whether B protein alone was sufficient to activate galactokinase expression from the p_V*galK* plasmid pEC433 (Section 5.2.3), the *B* expression plasmid pEC434 (Section 5.2.1) was transformed into the *galK*$^-$ strain E864 (*lacI*$^+$) carrying the p_V*galK* plasmid pEC433. In the presence of the *B* expression plasmid pEC433, under inducing conditions (42^o and 1mM IPTG, 2 hours), the p_V*galK* plasmid pEC433 produced high levels of galactokinase activity (Table 5.2; 420 units, 45 minutes post induction), the level of galactokinase expression from the p_V*galK* plasmid pEC433 reaching a maximum 45 minutes after induction of the *B* expression plasmid pEC434 and presumably reflected maximal activation of the p_V promoter. The induction of galactokinase activity was entirely *B*-dependent, as demonstrated by the absence of induction of galactokinase expression with the use of the *Bam*17 expression plasmid pEC435 under the same inducing conditions in the non-suppressing host E864 (Table 5.2; 4.5 units, 45 minutes post induction). This activation was also specific to the cloned restriction fragment carrying the p_V promoter, as the parent plasmid pEC432 gave no significant galactokinase activity after induction of the *B* expression plasmid pEC434 (Table 5.2; 5.2 units, 45 minutes post induction). The *B* gene product alone was concluded to be able to significantly activate p_V transcription and therefore the *B* gene product at high concentration was the only phage function directly required for activation of p_V transcription in the *galK* promoter analysis system.

The maximum level of galactokinase activity observed (Table 5.2; 420 units, 45 minutes post induction) was approximately 4 fold higher than that observed in the prophage induction assay (Section 4.2.3; Table 4.1; 110 units, 45 minutes post induction). The high level of galactokinase activity from the p_V*galK* plasmid observed after induction

TABLE 5.2

Activation of p_V transcription by B protein[a].

Expression plasmid[b]	Conditions		Galactokinase activity[c]
$pp_{trc}B^+$	30°C		76.0±9.1
$pp_{trc}Bam$	30°C		0.7±0.5
$pp_{trc}B^+$	42°C/IPTG	15 min	187 ±18
		30 min	241 ±13
		45 min	420 ±9
		60 min	432 ±21
$pp_{trc}Bam$	42°C/IPTG	45 min	4.5±1.2
$pp_{trc}B^+$	30°C/F'$lacI^q$		32 ±3.7

Notes to Table 5.2

a. Cultures of the non-permissive strain E864 carrying the $p_V galK$ plasmid (pEC433), and either the *B* expression plasmid $pp_{trc}B^+$ (pEC434) or the *Bam* expression plasmid $pp_{trc}Bam$ (pEC435) were grown at 30°C in M13 minimal media supplemented with the appropriate antibiotics. Non-induced cultures were grown to A_{650}=0.3-0.6. For induction of expression plasmid pEC434 and pEC435, the cultures were grown at 30°C in M13 minimal media supplemented with antibiotics to A_{650}=0.2, transferred to 42°C and IPTG addd to 1 mM final concentration and samples were taken at the indicated times. The galactokinase activity of the $p_V galK$ plasmid (pEC433) in the presence of the $pp_{trc}B^+$ plasmid (pEC434) and the $lacI^q$ gene on an episome was performed in strain JM105. Strain JM105 has a $galK^+$ genotype and the galactokinase units presented represent the units above that recorded for JM105 carrying the $p_V galK$ plasmid (pEC433) and plasmid pEC612 (Section 2.3.1).

b. Plasmids and plasmid clones are described in detail in Section 2.3.

c. The galactokinase activity was determined as described in Section 2.18, and are the mean of three separate experiments, with the standard deviations shown.

of the *B* expression plasmid pEC434, in comparison to that recorded after prophage induction, was considered to be due to the combination of the higher concentration of B protein from the plasmid pEC434 and the absence of late promoters competing for B protein with p_V, when the source of B protein was the expression plasmid rather than the induced prophage.

The uninduced *B* expression plasmid also gave high levels of galactokinase activity from the p_V*galK* plasmid pEC433 in strain E864 (Table 5.2; 76.1 units at 30°C). This high level of galactokinase activity observed from the p_V*galK* plasmid in the presence of uninduced *B* expression plasmid was not expected. The plasmid pEC434 should be maintained at single copy at 30°C and the *trc* promoter expressing the *B* gene should be repressed by the *lac* repressor encoded by the host (E864) *lacI*$^+$ gene. Under these conditions, the level of B protein produced from pEC434 was expected to be low, and superficially was expected to mimic the level of B protein produced in a 186 lysogen, due to the autogenous control of *B* gene transcription (Kalionis *et al.*, 1986b). Under conditions of repression of the *trc* promoter of pEC434, the p_V*galK* plasmid was not expected to be activated, as there is no evidence of activation of transcription of the p_V*galK* plasmid in a lysogen (Section 4.2.4(a)). However, Lanzer and Bujard (1988) reported that the related, hybrid *tac* promoter was less repressible by *lac* repressor than the *lac* promoter. In those studies, high levels of *lac* repressor were produced from a plasmid clone and the *tac* promoter was found to be repressed only 50 fold under these conditions, compared to over 900 fold repression for the wild-type *lac* promoter. The *tac* promoter was also found to have significant activity under conditions of high *lac* repressor concentration (6% of the activity of the fully derepressed *lac* promoter). Therefore the related *trc* promoter was expected to be poorly repressed by the single copy *lacI* gene of strain E864. It was also anticipated from the poor repression of the *tac* promoter at high *lac* repressor concentration as observed by Lanzer and Bujard (1988), that the *trc* promoter would not be completely repressed at high *lac* repressor concentration. Consistent with this, the presence of the *lacI*^{q} gene on an episome to increase the level of *lac* repressor (Muller-Hill *et al.*, 1968) did not dramatically improve the repression at the *trc* promoter, as the p_V*galK* plasmid pEC433 in host JM105 (F'*lacI*^{q}) gave 32.0 units of galactokinase activity at 30°C in the presence of the *B* expression plasmid pEC434 (Table 5.2).

Despite the poor repressibility of the *trc* promoter expressing the *B* gene from plasmid pEC434, the *B* gene product at high concentration was concluded to be the only phage function required directly for activation of p_V transcription in the *galK* promoter analysis system and it is considered likely that the B protein will also be the only phage function required to directly activate transcription from the remaining late promoters of 186.

Since the B protein has been shown to be the only phage function directly required to activate p_V transcription on the plasmid pEC433, this raised the question as to why there is no detectable galactokinase expression from p_V on the p_V*galK* plasmid pEC428 in a 186 lysogen (Section 4.2.4(a)), despite the fact that the *B* gene is transcribed in this state and the B protein produced is active in repressing its own transcription (Kalionis *et al.*, 1986b). This superficially suggested that the level of B protein in the lysogenic state may be insufficient for activation of p_V transcription. The autogenous control of *B* gene transcription in the lysogenic state could be the reason for the inability of B protein to activate the p_V*galK* plasmid, as it may result in levels of B protein insufficient for late activation. Alternatively, the inability of B protein to activate p_V transcription in a lysogen could be due to the expression of a function in the lysogen which directly antagonises B protein activating late transcription. The reason for the inability of the *B* gene product to activate late gene transcription in the lysogenic state is investigated in Chapter 7.

Late gene transcription has been shown to be dependent upon the functions of genes *B* and *A* (Finnegan and Egan, 1981). Kalionis *et al.* (1986b) considered the possibility that another function apart from genes *B* and *A* was required for activation of late transcription, and this function was expected to be under *c*I repressor control. However, the results presented in this Chapter demonstrated that the *B* gene is the only phage function required to directly activate late transcription and the reasons that led Kalionis *et al.* (1986b) to entertain the possibility of a third function for activation of late gene trancription are discussed in Section 8 1.1(a). The role of gene *A* in activation of late transcription is expected to be indirect, and to reflect a requirement for increased gene dosage and/or a requirement for a replicating template in order for B to activate late transcription (Kalionis *et al.*, 1986b). The role of replication in activation of late transcription is investigated in Chapter 6.

5.3 SUMMARY.

The aim of this chapter was to identify the phage functions required to directly activate 186 late transcription. The first step was to determine if B protein alone was sufficient to activate late transcription and this was investigated by determining the ability of B protein alone to activate p_V transcription on the p_V*galK* plasmid to a level comparable to that after prophage induction. This approach required the construction of a *B* expression plasmid so as to supply B protein *in trans* to the p_V*galK* plasmid.

The *B* expression plasmid constructed, plasmid pEC434, was shown to express functional *B* gene product *in vivo*, as the *B* gene product expressed from the plasmid was able to complement a 186 *Bam*17 phage in infection. Expression of B protein from plasmid pEC434 was shown to activate the p_V*galK* plasmid *in vivo* in the absence of phage functions, and it was concluded that B protein at high concentration was the only phage function required directly for activation of p_V transcription in the *galK* promoter analysis system.

Since B protein was shown to be the only phage function required directly for activation of p_V transcription, this raised the question as to why p_V transcription on the p_V*galK* plasmid was not activated in a 186 lysogen (Section 4.2.4(a)), even though B protein is present and active in repressing its own transcription in this state (Kalionis *et al.*, 1986b). This indicated that the concentration of B protein in the lysogenic state was probably insufficient to activate the p_V promoter, and it appeared that B protein at high concentration was required for activation. However the existence of a lysogenic factor which directly antagonises *B*-dependent activation has not been excluded. The nature of the autogenous control of *B* gene transcription and the inability of B protein to activate late transcription in the lysogenic state are investigated in Chapter 7.

Late gene transcription in 186 has been shown to be dependent upon the functions of genes *B* and *A* (Finnegan and Egan, 1981). The work presented in this chapter demonstrated that the *B* gene was the only phage gene required directly for activation of late transcription. The following chapter investigates the indirect role of the *A* gene in activation of late transcription.

CHAPTER 6

INVESTIGATION OF THE ROLE OF REPLICATION IN ACTIVATION OF 186 LATE TRANSCRIPTION

CHAPTER 6. INVESTIGATION OF THE ROLE OF REPLICATION IN ACTIVATION OF 186 LATE TRANSCRIPTION.

6.1 INTRODUCTION.

The results presented in Chapter 5 demonstrated that the *B* gene product was the only phage function directly required for activation of late transcription, since B protein alone was able to efficiently activate p_V transcription on the p_V*galK* plasmid. However, the appearance of late gene transcription after prophage induction is also dependent upon the function of gene *A* (Finnegan and Egan, 1981). Gene *A* mutants are defective in phage replication (Hocking and Egan, 1982d; Sivaprasad *et al.*, 1990) and this raised the question as to the role of the replication gene *A* in the activation of 186 late transcription.

The role of gene *A* in activating late transcription has been shown to be indirect, since late gene transcription occurred in the absence of the *A* gene product in a phage-plasmid hybrid, in which replication was under plasmid control (Kalionis *et al.*, 1986b). Kalionis *et al.* (1986b) concluded that the role of replication in activation of late transcription was at two levels. Firstly, replication was predicted to provide a template topology which was necessary for activation of late transcription. Secondly, a consequence of replication would be an increase in template numbers. This was expected both to increase late gene dosage and to titrate the finite pool of *B* gene repressor present in the lysogen. Increase in late gene dosage was anticipated to be important for efficient morphogenesis and titration of the *B* gene repressor was expected to result in the induction of transcription of the *B* gene, which was probably necessary for efficient expression of the late functions.

Thus the role of replication in activation of late transcription was expected to be the provision of a topological requirement to effect *B* activation, to increase gene dosage and to derepress *B* gene transcription. The aim of this chapter was to determine if replication had any role in providing a specific template topology and also to investigate the role of replication in induction of *B* gene transcription in order to activate late transcription.

6.2 RESULTS AND DISCUSSION.

6.2.1 Investigation of a topological requirement for activation of p_V transcription.

The ability of the B protein alone to activate p_V transcription on the p_V*galK* plasmid (Section 5.2.4) may have reflected the fact that the plasmid provided the topological equivalent of a phage replicating template. A topological requirement was entertained by Kalionis *et al.*(1986b) because of the apparent inability of B protein from a superinfecting, immunity-insensitive phage to rescue a late function from a prophage (Hocking and Egan, 1982a), in spite of the fact that the superinfecting phage could replicate and presumably provide excess *B* function. Gene *A* does not act *in trans* (Hocking and Egan, 1982c), and the possibility existed that the non-replicating state of the prophage was one barrier to B activating the late promoters of the prophage.

To determine if B protein at high concentration would activate late transcription from a prophage, the ability of the B protein expressed from the *B* expression plasmid pEC434 (Section 5.2.1) to activate p_V transcription of a prophage was determined, by hybridization of total RNA to a radiolabelled DNA probe specific for *V* gene transcription.

6.2.1(a) Transformation of a 186 lysogen with the *B* expression plasmid pEC434.

The first step in providing excess B protein to a 186 prophage was to transform a 186 lysogen with the *B* expression plasmid pEC434 (Section 2.15.2(b)). However, it was found that the *B* expression plasmid pEC434 was unable to transform a 186 lysogen of host E251 (transformation efficiency of less than 1 transformant/ug of DNA; Table 6.1). This inability to transform a 186 lysogen was dependent upon functional B protein, as the *Bam* expression plasmid pEC435 transformed a 186 lysogen of the non-permissive host E251 with the same efficiency as the parent vector pEC612 (Table 6.1; 2.5×10^5 transformants/ug). The inability of the *B* expression plasmid to transform a lysogen was demonstrated to be due to the presence of the prophage, and not a lethal consequence of *B* expression itself, as the *B* expression plasmid pEC434 was able to transform the corresponding E251 non-lysogen (Table 6.1; 1.1×10^5 transformants/ug).

It appeared that inadequate control of the *trc* promoter of pEC434 by the host encoded *lacI* repressor of E251 was leading to sufficient *B* expression to activate

TABLE 6.1

Transformation efficiencies of plasmid clones encoding the *B* or *Bam* genes.

Culture[a]	Transforming plasmid[b]	Transformation efficiency[c]
non-lysogen	$pp_{trc}B^+$	$1.1x10^5$
(186 *cIts*)	$pp_{trc}B^+$	<1
	$pp_{trc}Bam$	$2.5x10^5$
	pp_{trc}	$2.5x10^5$
(186 *cIts*) p*lacI*^{q}	$pp_{trc}B^+$	$4.7x10^4$
	$pp_{trc}Bam$	$2.0x10^5$

Notes to Table 6.1.

a. Cultures of the non-permissive strain E251 and the corresponding 186 *cIts* lysogen (E252) were grown at 30°C in M13 minimal media (supplemented with tetracycline for the p*lacI*^{q} plasmid pMC7) until A_{600}=0.4-0.6, competent cells prepared and transformed (Section 2.15.2) with CsCl purified DNA in varying amounts from 1 ng to 1 ug (Section 2.21.2). Transformation mixes were allowed to express antibiotic for 2 hours at 30°C in M13 minimal media and plated (0.1 ml) onto M13 minimal plates supplemented with the appropriate antibiotics. Plates were grown at 30°C for 36 hours and the number of transformants scored. The p*lacI*^{q} plasmid pMC7 is described in Section 2.3.1.

b. The $pp_{trc}B^+$ plasmid pEC434 and the $pp_{trc}Bam$ plasmid pEC435 are described in Section 2.3.2. The parent plasmid pp_{trc} (pEC612) is described in Section 2.3.1.

c. Transformation efficiency is expressed as the total number of transformants obtained per ug of transformed DNA.

transcription of the prophage late genes with lethal consequences. This was consistent with the fact that the *B* expression plasmid pEC434 in the presence of a single copy of the *lacI* gene was also found to considerably activate the p_V promoter of the p_V*galK* plasmid pEC433 (Section 5.2.4). The lethality of *B* expression plasmid pEC434 to a 186 prophage in the absence of *lacI*^{q} control was also consistent with the observation that 186 formed clear plaques on a strain carrying the *B* expression plasmid pEC434 (Section 5.2.2). Presumably, the inability to stably maintain a 186 lysogen in the presence of uncontrolled *B* expression from pEC434 would result in clear plaques, as observed in Section 5.2.2.

Increasing the concentration of *lacI* repressor by use of the multicopy plasmid pMC7 carrying the *lacI*^{q} gene (Calos, 1978), allowed the successful transformation of a 186 lysogen with the plasmid pEC434 (Table 6.1; 4.7×10^4 transformants/ug), albeit with slightly reduced efficiency as compared to the *Bam* expression plasmid pEC435 (Table 6.1; 2.0×10^5 transformants/ug). This ability to transform a lysogen did not appear to be due to loss of the prophage or a plasmid mutation inactivating the *B* gene. This was drawn from the fact that the pEC434 transformants of the 186 lysogen remained lysogenic for 186, as determined by their immunity to 186 *c*I$^-$ phage and sensitivity to 186 *vir*1 phage by cross streaking (Section 2.15.1), and plasmid pEC434 DNA extracted from these transformants was still functional in activation of late transcription, as determined by the ability of plasmid DNA isolated from these transformants (Section 2.21.1) to activate the p_V*galK* plasmid in a *galK*$^-$ host (E863) on MacConkey indicator plates and the inability of plasmid DNA to transform a 186 lysogen (data not shown).

Additionally, the *trc* promoter of plasmid pEC434 did not appear to be completely repressed by the multicopy *lacI*^{q} clone pMC7, since 186 *cIts Bam*17 phage plated on strain E251, carrying the *B* expression plasmid pEC434 and plasmid pMC7, with an efficiency of plating of 0.67 (Table 6.2). However, the plaques formed by 186 *cIts Bam*17 in the presence of plasmid pMC7 had a very small phenotype (Table 6.2). This lowered transformation efficiency and small plaque size were concluded to be a result of *lacI*^{q} repression, since in the presence of 1 mM IPTG the efficiency of plating increased to 1.24 and the plaques showed a large, clear phenotype (Table 5.2). It was concluded that even in the presence of high *lacI* concentration from plasmid pMC7, there was still sufficient *B* expression to allow complementation of a 186 *Bam* phage in infection. As discussed

TABLE 6.2

Plating efficiencies of 186 *Bam* phage on strains carrying *B* expression plasmid pEC434 and *lacI*^{q} clone pMC7[a].

Plasmid clone[b]	Conditions	Efficiency of plating[c]	Plaque size
p*ptrcB*$^+$	30°C	1.36	large
p*ptrcB*$^+$ p*lacI*^{q}	30°C	0.67	very small
	30°C/1 mM IPTG	1.24	large

Notes to Table 6.2

a. Cultures of the non-permissive strain E251 carrying the *B* expression plasmid p$p_{trc}B^+$ (pEC434) or E251 carrying plasmid p$p_{trc}B^+$ (pEC434) and the p*lacI*^{q} plasmid (pMC7) were grown at 30°C in M13 minimal media supplemented with the appropriate antibiotics to A_{600}=0.8. Dilutions of 186 *cIts Bam*17 phage were plated with E251 carrying expression plasmid p$p_{trc}B^+$ (pEC434) onto M13 minimal plates supplemented with the appropriate antibiotics (Section 2.16.1). Dilutions of 186 *cIts Bam*17 were also plated with E251 carrying plasmid p$p_{trc}B^+$ (pEC434) and p*lacI*^{q} plasmid (pMC7) onto M13 minimal plates supplemented with the appropriate antibiotics and 1 mM IPTG (final concentration). Phage 186 *cIts Bam*17 is described in Section 2.2. Plates were incubated overnight at 30°C and the number of plaques and their phenotypes noted.

b. The p*lacI*^{q} plasmid pMC7 is described in detail in Section 2.3.1 and the *B* expression plasmid clone p$p_{trc}B^+$ (pEC434) is described in detail in Section 2.3.2.

c. The efficiency of plating is expressed as the ratio of the number of plaques obtained on strains carrying the plasmid clones to the number of plaques obtained on the *supE* strain E540 carrying the parent plasmid pEC612 and the p*lacI*^{q} plasmid pMC7.

previously (Section 5.2.4), the related *tac* promoter is also poorly repressed at high repressor concentration (Lanzer and Bujard, 1988). Nevertheless, the reduction in *B* expression was sufficient to allow transformation of a 186 lysogen, and demonstrated that a 186 lysogen could tolerate a low level of *B* expression from plasmid pEC434.

6.2.1(b) p_V transcription from a 186 prophage.

In order to determine whether B protein was able to activate p_V transcription from a prophage, a 186 *cIts* lysogen carrying the *B* expression plasmid pEC434 and the *lacI*^{q} plasmid pMC7, was induced by the addition of 1 mM IPTG (final concentration) and total RNA isolated two hours after IPTG induction (Sections 2.32.1 and 2.32.2). To quantitate the amount of p_V transcription, the level of hybridization of a radiolabelled probe specific for *V* transcription to RNA was quantitated using RNA slot blot analysis, as described in Section 2.32.4. The probe specific for *V* gene transcription was a *r*-strand, 325 base *Oxa*NI-*Acc*I (sequence positions 928-1252; Figure 3.4(a)) probe prepared by synthesizing a radiolabelled, complementary strand of the M13 clone mEC404 (a M13mp9 clone of the *Bam*HI-*Eco*RI (10.0% to 13.3%) restriction fragment cloned into the *Bam*HI and *Eco*RI sites; Section 2.4.2) as described in Section 2.31.2.

The results of the RNA slot blot analysis are shown in Figure 6.1 and the level of hybridization quantitated by laser densitometric scanning presented in the legend to Figure 6.1. Substantial p_V transcription was detected with the *V* specific probe upon induction of *B* gene expression from plasmid pEC434 with IPTG in a 186 *cIts* lysogen (Figure 6.1, tracks 3 and 4). No transcription was detectable upon induction with IPTG of a 186 *cIts* lysogen of the non permissive strain E251 carrying the *Bam* expression plasmid pEC435 (Figure 6.1, track 5) and demonstrated that the transcription detected upon induction of pEC434 was *B* gene specific. No transcription was detectable upon induction of a non-lysogen (E251) carrying plasmids pMC7 and pEC434 with IPTG (Figure 6.1, track 6), demonstrating that the transcription detected was specific to the prophage.

Further, it appeared that p_V transcription on the prophage was being strongly activated, as the extent of transcription from the prophage was considered substantial, being greater than 3% of that found 35 minutes after heat-induction of a 186 *cIts* prophage

Figure 6.1 (cont.)

Quantitation of slot blot analysis of *pV* transcription.

Culture[a]	Conditions[b]	Level of Hybridization per ug of total RNA[c]
(186 *cIts*)	39°C 35 min (1)	2.7×10^6
(186 *cIts*) *placI*^{q} p*p*$_{trc}$*B*$^+$	30°C (2)	2.8×10^3
	30°C 120 min after IPTG induction (3,4)	9.7×10^4
(186 *cIts*) *placI*^{q} p*p*$_{trc}$*Bam*	30°C 120 min after IPTG induction (5)	Not Detectable
placI^q p*p*$_{trc}$*B*$^+$	30°C 120 min after IPTG induction (6)	Not Detectable

a. Cultures and plasmids are described above. The *placI*^{q} plasmid pMC7, the p*p*$_{trc}$*B*$^+$ plasmid pEC434 and p*p*$_{trc}$*Bam* plasmid pEC435 are described in detail in Section 2.3.

b. Corresponding tracks in the autoradiograph are indicated in brackets.

c. The level of hybridization represents the total peak area obtained by laser densitometric scanning of autoradiographs of RNA slot blot hybridized with the *V* specific probe as described in Section 2.32.4(c). Not detectable represents no signal above background. Results presented are the average of separate, duplicate RNA preparations hybridized with the probe.

Figure 6.1 Level of p_V transcription from a 186 prophage.

Slot blot analysis of the level of hybridization of a *V* specific probe to RNA prepared from cultures of the non-permissive strain E251 and the corresponding 186 *cIts* lysogen (E252), carrying the *lacI*^{q} clone pMC7 (Section 2.3.1) and either the *B* expression plasmid (p$p_{trc}B^+$) pEC434 or the *Bam* expression plasmid (p$p_{trc}Bam$) pEC435 (Section 2.3.2). Cultures were grown in M13 minimal media supplemented with the appropriate antibiotics at 30°C. For non-induced cultures, the cultures were grown at 30°C to A_{600}=0.8 and RNA prepared (Sections 2.32.1 and 2.32.2). For heat-induction of E251 (186 *cIts*), the cultures were grown at 30°C to A_{600}=0.8 and transferred to 39°C and RNA prepared 35 minutes after induction (Sections 2.32.1 and 2.32.2). For IPTG induction of cultures, the cultures were grown at 30°C to A_{600}=0.2, IPTG added to 1 mM final concentration and RNA was prepared 2 hours after IPTG induction.

Glyoxal denatured RNA (Section 2.32.3) was analysed by slot blot analysis (Section 2.32.4) with a 325 b *Oxa*NI-*Acc*I (sequence coordinates 928-1252, Figure 3.4(a)) probe ($4.2x10^6$ cpm) specific for *V* transcription, prepared from mEC404 (Section 2.31.2; Section 2.4.2). The position of the probe on the DNA sequence of the *Bam*HI-*Eco*RI (10.0% to 13.3%) region is shown below the autoradiograph. RNA from each culture (2 ug, 1 ug, 0.5 ug, 0.25 ug, 0.125 ug and 0.0625 ug) was loaded in the order shown. The filter was autoradiographed at -80°C with an intensifying screen for 4 days.

RNA samples:

1. RNA from a 186 *cIts* lysogen 35 minutes after heat-induction.

2. RNA from a non-induced 186 *cIts* lysogen carrying the *lacI*^{q} plasmid pMC7 and the *B* expression plasmid (p$p_{trc}B^+$) pEC434.

3. RNA from an IPTG-induced 186 *cIts* lysogen carrying the *lacI*^{q} plasmid pMC7 and the *B* expression plasmid (p$p_{trc}B^+$) pEC434, 2 hours post induction.

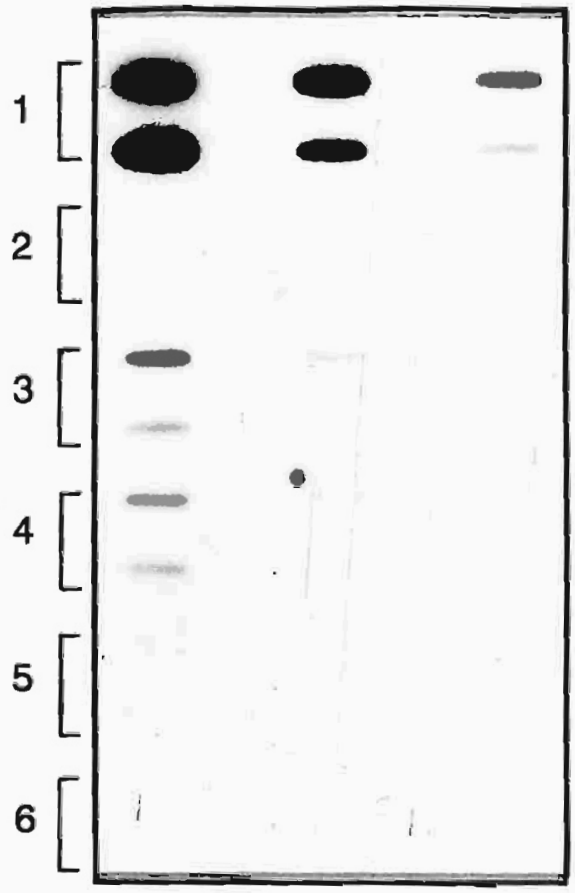
4. Identical to RNA sample 3.

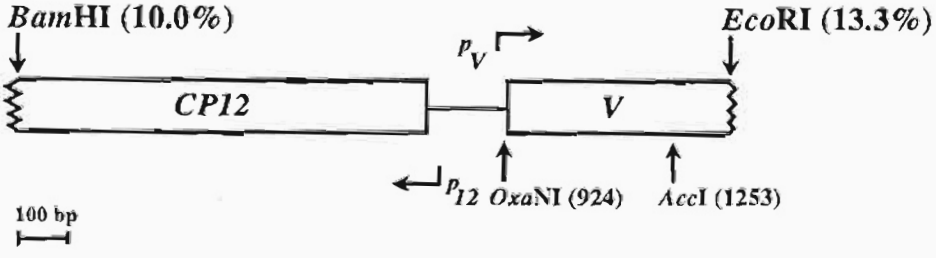
5. RNA from an IPTG-induced 186 *cIts* lysogen carrying the *lacI*^{q} plasmid pMC7 and the *Bam* expression plasmid (p$p_{trc}Bam$) pEC435, 2 hours post induction.

6. RNA from an IPTG-induced non-lysogen carrying the *lacI*^{q} plasmid pMC7 and the *B* expression plasmid (p$p_{trc}B^+$) pEC434, 2 hours post induction.

The order of the spots in the figure is labelled incorrectly and should read:

2 ug	0.5 ug	0.125 ug
1 ug	0.25 ug	0.0625 ug


1
2
3
4
5
6
2 ug
1 ug
0.5 ug
0.25 ug
0.125 ug
0.0625 ug


BamHI (10.0%)
EcoRI (13.3%)
p_V
CP12
V
p_{12}
OxaNI (924)
AccI (1253)
100 bp

(Figure 6.1, track 1) and was therefore approximately equal to the relative number of templates available per cell in each case.

Transcription from the p_V promoter on the prophage in the presence of uninduced *B* expression plasmid was just detectable (Figure 6.1, track 2) and was approximately 30 fold lower than the level of hybridization detected after induction with IPTG. Since no p_V transcription was detectable upon induction of a 186 *cIts* lysogen carrying the *Bam* expression plasmid pEC435 with IPTG (Figure 6.1, track 5), the transcription in the presence of uninduced *B* expression plasmid was most probably due to incomplete repression of the *trc* promoter. This was consistent with the ability of *B* expression plasmid to complement 186 *Bam* phage in the presence of the *lacI*^{q} clone pMC7 (Section 6.2.1(a)).

6.2.1(c) Increased *B* expression does not cause prophage induction.

It was possible that induction of *B* expression from plasmid pEC434 resulted in induction of the prophage for an unknown reason. In such a case, the p_V transcription detected may have been from an excised, replicating phage template and not from a prophage as expected. To determine whether induction of *B* expression from plasmid pEC434 did result in prophage induction, phage assays (Section 2.16.1) were performed on lysogenic cultures carrying plasmids pEC434 after IPTG induction (Table 6.3).

Induction of a wild-type 186 lysogen carrying the *lacI*^{q} clone pMC7 and the *B* expression plasmid pEC434 with IPTG resulted in no increase in phage titre four hours after IPTG induction, above the non-induced equivalent culture (Table 6.3). Similarly, the phage titre after IPTG induction of a lysogen carrying plasmids pMC7 and the *Bam* expression plasmid pEC435 was equivalent to that after IPTG induction of a lysogen with the *B* expression plasmid (Table 6.3). It therefore appeared that increased B protein was not resulting in prophage induction.

However, it may also have been possible that *B* expression from pEC434 was causing prophage induction, but the phage was blocked in productive infection after prophage excision, due to the fact that increased B protein activated the late genes prematurely. To determine if increased *B* expression was detrimental to productive infection after prophage induction, the ability of 186 to be induced by UV irradiation under

TABLE 6.3

Increased *B* expression does not cause prophage induction.

Culture	Titre of uninduced culture	IPTG-induced Culture		IPTG-induced UV-induced Culture	
		Titre of	Increase	Titre	Increase
(186$^+$)[p*lacI*^{q},pp$_{TRC}$*B*$^+$]	2.4x 10^7	1.8x 10^7	0.8 x	1.7x 10^9	94 x
(186$^+$)[p*lacI*^{q},pp$_{TRC}$*Bam*]	1.7x 10^7	1.0x 10^7	0.7 x	1.1x 10^9	110 x

Notes to Table 6.3

a. Cultures of 186 wild-type lysogen (186$^+$) of the non-permissive strain E251 (E635) carrying the p*lacI*^{q} plasmid pMC7 and either the *B* expression plasmid plasmid p$p_{trc}B^+$ (pEC434) or the *Bam* expression plasmid pp_{trc}*Bam* (pEC435) were grown at 30°C in M13 minimal medium supplemented with ampicillin (25 ug/ml) and tetracycline (15 ug/ml) to A_{600}=0.8. For IPTG induction, cultures were diluted into M13 minimal media supplemented with antibiotics and 1 mM IPTG (final concentration) and incubated at 30°C for 4 hours, treated with chloroform and assayed for free phage (Section 2.16.1). Uninduced cultures were treated identically, but without the addition of IPTG. For concurrent UV and IPTG induction, the cultures were induced with 1 mM IPTG (final concentration) for 15 min, centrifuged and the bacteria resuspended in 10 mM $MgSO_4$. Portions were transferred to glass Petri dishes to depth of 1 mm and UV-irradiated at 15 J/m^2 (254 nm), as described in SEction 2.17. Bacteria were then diluted into minimal media containing the appropriate antibiotics and 1 mM IPTG and incubated at 30°C, samples taken after 4 hours and treated with chloroform and assayed for free phage (Section 2.16.1). The p*lacI*^{q} plasmid pMC7, the *B* expression plasmid p$p_{trc}B^+$ (pEC434) and the *Bam* expression plasmid pp_{trc}*Bam* (pEC435) are described in detail in Section 2.3.

b. Phage titres are given in pfu/ml. Increase in phage titres is expressed as the ratio of pfu/ml afer induction to pfu/ml without induction.

conditions of increased *B* expression from plasmid pEC434 was tested. Phage 186 is known to be UV inducible from the prophage state (Woods and Egan, 1974). Induction of *B* expression from plasmid pEC434 was shown not to affect productive development following UV irradiation, as the burst size following UV irradiation and IPTG induction of a lysogen with the *B* expression plasmid was 94 (Table 6.3), only slightly reduced compared to the burst from a 186 lysogen carrying the *Bam* expression plasmid under the same conditions (110; Table 6.3).

It was concluded that increased *B* expression did not result in prophage induction and that the p_V transcription detected upon IPTG induction of the *B* expression plasmid in a 186 lysogen was most likely to be occurring from a 186 prophage. It was therefore concluded that the ability of B protein to activate late transcription from a prophage as efficiently as vegetative late promoters, after taking the difference in template numbers into account, demonstrated that there was no requirement of a specific template topology in order for B to activate the late promoters of a prophage.

6.2.2 Increased *B* gene dosage activates p_V transcription.

The results presented in Sections 6.2.1(b) and 6.2.1(c) demonstrated that there was no role for a specific template topology in the activation of late transcription. This raised the question as to the role of replication in the activation of late transcription, as there is no evidence of late gene transcription in the absence of phage replication (Finnegan and Egan, 1981). Presumably the role of replication in activation of late transcription was related in some manner to the increase in template number that would result as a consequence of replication.

An increase in template number necessarily means an increase in dosage of both the late genes and the *B* gene. An increase in the dosage of the late genes is expected to be important for efficient morphogenesis. However, the effect of an increase in the dosage of the *B* gene on late gene transcription is not known. Transcription of the *B* gene appears to be autoregulated (Kalionis *et al.*, 1986b). The B protein present in an uninduced lysogen was unable to activate p_V transcription on the p_V*galK* plasmid (Section 4.2.4(a)), apparently reflecting the single copy status of the *B* gene. One of the roles of replication in activation of late transcription may be to increase *B* gene dosage and thereby activate late

transcription by the consequent increase in the concentration of free B protein, as the concentration of the gene product of an autogenously repressed gene is expected to increase with increasing gene dosage (discussed in detail in Appendix II). It was therefore necessary to determine if an increase in copy number of the *B* gene, as would occur due to phage replication, was sufficient to efficiently activate late transcription. To gain this increase in *B* gene dosage, the *B* gene under the control of its own promoter was cloned onto the multicopy p_V*galK* plasmid (Section 4.2.3) and the activation of p_V transcription by the *B* gene in multicopy was investigated.

The 2.2 kb *Sac*I (63.0% to 70.2%) restriction fragment encodes the *B* gene and its promoter. The only source of *B* gene transcription on this restriction fragment is expected to be that from the *B* promoter, since the presumptive late promoter which transcribes the *B* gene late in infection initiates to the left of the *Hin*dIII site at 61.3% (Kalionis *et al.*, 1986b). The 2.2 kb *Sac*I (63.0% to 70.2%) restriction fragment from 186 *cIts* DNA was isolated by agarose gel electrophoresis (Section 2.26.1.; Section 2.27.1) and cloned into the *Sac*I site of the plasmid pK02 and the p_V*galK* plasmid pEC428, in an orientation such that *B* gene transcription was directed towards the *bla* gene, as confirmed by diagnostic restriction analysis of plasmid DNA (Section 2.28.3(c)), to give the plasmids pEC429 and pEC430, respectively. In a similar manner the 2.2 kb *Sac*I (63.0% to 70.2%) restriction fragment from 186 *cIts Bam*17 DNA was cloned into the *Sac*I site of the p_V*galK* plasmid pEC428, to give plasmid pEC431, with the same orientation of the inserted DNA as plasmids pEC429 and pEC430.

The presence of the *B* gene on the p_V*galK* plasmid (pEC430) was sufficient to induce high levels of galactokinase activity (Table 6.4; 416 units, pEC430). This activation was dependent on the cloned *B* gene product, as the *Bam* gene cloned onto the plasmid showed no galactokinase activity in the non-suppressing host E864 (Table 6.4; -0.8 units, pEC431). The activation was also specific to the cloned p_V promoter, as demonstrated by the absence of activation of galactokinase expression from the plasmid pEC429 without the cloned p_V promoter (Table 6.4; -0.7 units, pEC429). This absence of galactokinase expression from pEC429 also demonstrated the absence of leftward transcription (with respect to the genetic map) originating from the cloned 2.2 kb *Sac*I (63.0% to 70.2%) fragment and reading into the *galK* gene.

TABLE 6.4

Activation of p_V transcription by increase in *B* gene dosage[a].

Plasmid construct[b]	Galactokinase activity[c]
$B^+p_Bp_V galK$	416 ±11
$Bamp_Bp_V galK$	-0.8±0.6
$B^+p_B galK$	-0.7±0.9

Notes to Table 6.4

a. Cultures of the non-permissive strain E864 carrying the $B^+p_Bp_V galK$ plasmid (pEC430), the $Bamp_Bp_V galK$ plasmid (pEC431) or the $B^+p_B galK$ plasmid (pEC429) were grown at 37°C in M63 supplemented media with ampicillin (50 ug/ml) to A_{650}=0.4-0.6, samples taken and the galactokinase activity determined.

b. Plasmid clones are described in detail in Section 2.3.2.

c. The galactokinase activity was determined as described in Section 2.18, and are the average of two separate experiments.

The level of galactokinase activity produced from pEC430 was equivalent to the maximal activation of the p_V promoter on the $p_V galK$ plasmid pEC433 observed in the presence of the *B* expression plasmid pEC434 under inducing conditions (Section 5.2.4) and demonstrated that sufficient B protein was produced from the cloned *B* gene under its own control when present in multiple copy to maximally activate p_V transcription. It was concluded that the *B* gene in multicopy was able to efficiently activate p_V transcription and demonstrated that the provision of increased *B* gene dosage was one of the roles of replication in efficiently activating late transcription.

6.2.3 Elevated copies of the *B* operator sequence activate p_V transcription in a 186 lysogen.

The *B* gene appears to autogenously repress its own transcription. Two inverted repeat sequences were found overlapping the *B* promoter, and these may be the binding site for the presumptive repressor of *B* gene transcription (Kalionis *et al.*, 1986b). Kalionis *et al.* (1986b) proposed that replication would have the effect of derepressing *B* gene transcription, since the rapid increase in template numbers would provide a corresponding increase in operator concentration which was expected to titrate the pool of the repressor of *B* gene transcription, at least transiently, which may be all that was needed to initiate a cascade of events. This derepression of *B* gene transcription would therefore augment the increase in *B* gene transcription due to increased *B* gene dosage (Section 6.2.2). Such an effect of derepression of *B* gene transcription is presumably observed by the ability of a superinfecting, immunity-insensitive phage mutant in the *B* gene (eg 186 *vir*2 *Bam*17) to rescue *B* gene function from a prophage by complementation (Hocking and Egan, 1982a), since the single prophage copy of the *B* gene when repressed is unable to activate late transcription (Section 4.2.4(a)), and it was surmised that some level of derepression must be occurring on superinfection in order to activate late transcription.

Increased *B* operator concentration may therefore activate late transcription by derepressing *B* gene transcription. To determine if an increase in the concentration of the *B* operator had the ability to derepress *B* gene transcription and thereby activate late transcription, excess *B* operator DNA was introduced into a 186 lysogenic cell, by way of

cloning onto the multicopy $p_V galK$ plasmid, and the activation of p_V transcription investigated.

As a first experiment to determine if increased operator sequences in a 186 lysogen would result in activation of p_V transcription, the source of operator DNA was the 2.2 kb *Sac*I (63.0% to 70.2%) restriction fragment encoding the *Bam*17 gene, as cloned onto the $p_V galK$ plasmid to give plasmid pEC431 (Section 6.2.2). This plasmid produced no detectable galactokinase activity in a non-permissive, non-lysogenic host (Section 6.2.2; Table 6.2). However, introduction of pEC431 into a 186 *cIts* lysogen resulted in induction of galactokinase activity (Table 6.5; 113 units), which was dependent upon a functional prophage copy of the *B* gene, as introduction of pEC431 into a 186 *cIts Bam*17 lysogen demonstrated almost complete loss of galactokinase activity (Table 6.5; 4.0 units).

To localize this effect specifically to the presumptive *B* operator region and to rule out any possibility of activation due to leakage of the amber mutation of the *Bam* gene on the multicopy plasmid pEC431, a restriction fragment containing only the *B* promoter/presumptive-operator region was cloned into the $p_V galK$ plasmid. A *Nde*I site has been created at the predicted initiation codon of the *B* gene in the M13 clone mEC650 as described in Section 5.2.1, and this *Nde*I site (sequence position 294) was used to separate the *B* promoter/operator region from the coding sequence of the *B* gene in conjunction with the *Pst*I site at 65.5%. The *Pst*I to *Nde*I (65.5% to 66.5%) restriction fragment was isolated by non-denaturing polyacylamide gel electrophoresis (Section 2.27.1(b)) after restriction (Section 2.25) and blunting of the termini (Section 2.28.2), to give a fragment of 290 bp. This restriction fragment was cloned into the blunted *Sac*I (Section 2.28.2) sites of plasmid pKO2 and the $p_V galK$ plasmid pEC428. Clones with the *B* promoter oriented towards the *bla* gene were selected by their white colony phenotype in strain E863 ($galK^-$) on MacConkey indicator plates. Clones with the opposite orientation of the inserted DNA gave red colonies, due to the expression of the *galK* gene from the cloned *B* promoter. Restriction analysis (data not shown; Section 2.28.3(c)) of plasmid DNA preparations confirmed the identity and orientation of the operator-*galK* plasmid pEC436 ($o_B galK$) and the operator-$p_V galK$ plasmid pEC437 ($o_B p_V galK$).

Plasmid pEC437 produced no detectable galactokinase activity in a non-lysogen (Table 6.5; -0.6 units). As was observed for plasmid pEC431, introduction of plasmid

TABLE 6.5

Derepression of prophage *B* gene by excess operator.

Culture[a]	Plasmid clone[b]	Galactokinase activity[c]
non-lysogen	$Bamp_Bp_VgalK$	-0.8±9.1
(186 *cIts*)	$Bamp_Bp_VgalK$	113 ±5.2
(186 *cIts Bam*17)	$Bamp_Bp_VgalK$	4.0±1.1
non-lysogen	o_Bp_VgalK	-0.6±0.2
(186 *cIts*)	o_Bp_VgalK	34.4±3.2
(186 *cIts Bam*17)	o_Bp_VgalK	0.3±0.1
(186 *cIts*)	o_BgalK	-0.1±0.4

Notes to Table 6.5

a. Cultures of the non-permissive strain E864 (non-lysogen) and the corresponding 186 *cIts* and 186 *cIts Bam*17 lysogens carrying the $Bamp_Bp_VgalK$ plasmid (pEC431), the operator-p_VgalK plasmid o_Bp_VgalK (pEC437) or the operator-*galK* plasmid o_BgalK (pEC436) were grown at 30°C in M63 supplemented media with ampicillin (50 ug/ml) to A_{650}=0.4-0.6, samples taken and the galactokinase activity determined (Section 2.18).

b. Plasmids clones are described in detail in Section 2.3.2.

c. The galactokinase activity was determined as described in Section 2.18, and are the mean of at least three separate experiments, with the standard deviations shown.

pEC437 into a 186 *cIts* lysogen led to the induction of galactokinase activity (Table 6.5; 34.4 units), which was dependent upon a functional prophage copy of the *B* gene, as introduction of the operator-p_V*galK* plasmid pEC437 into a 186 *cIts Bam*17 lysogen gave no detectable galactokinase activity in the non-suppressing host E864 (Table 6.5; 0.3 units, pEC437). The galactokinase activity was also dependent upon the cloned late promoter p_V, as the operator sequences cloned onto the parent plasmid pKO2 (pEC436) showed no detectable galactokinase activity in a 186 *cIts* lysogen (Table 6.5; -0.1 units, pEC436). The appearance of *galK* activity was also dependent upon the presence of the cloned operator sequence, since the p_V*galK* plasmid has been shown previously not to be active in a 186 lysogen (Section 4.2.4(a)). It was also noted that the level of galactokinase activity from pEC437 in a 186 lysogen (34.4 units; Table 6.5) was reduced three fold compared to that assayed after plasmid pEC431 was introduced into a lysogen (113 units; Table 6.5). The reason for this difference when the source of operator DNA was the *Bam*17 gene is not known, but may have been a result of leakage of the *Bam* gene on plasmid pEC431 contributing to the pool of B protein or may have been due to other DNA sequences downstream of the *Nde*I site at 66.5% that are involved in titration of the repressor of *B* gene transcription.

The results presented demonstrated that the single prophage copy of the *B* gene was able to activate late transcription in the presence of multiple operator sequences. This demonstrated the likely presence of regulatory DNA sequences on this restriction fragment. Presumably, the multiple *B* operator sequences on the p_V*galK* plasmid derepressed the *B* gene of the prophage sufficiently to activate late transcription. [Although it was also possible that the presence of the cloned operator sequences on the p_V*galK* plasmid had an effect *in cis* to allow activation of the p_V promoter in a lysogen]. This ability of excess operator sequences to induce the prophage copy of the *B* gene sufficiently to activate late transcription also suggested that replication had the ability to derepress *B* gene transcription and also suggested that replication may provide the initial event in activating late gene transcription during lytic development, by way of derepressing *B* gene transcription, at least transiently, which may be all that is required to initiate a cascade of events.

6.3 SUMMARY.

The aim of this chapter was to determine the role of replication in the activation of 186 late gene transcription. The role of phage replication in activation was expected to be indirect and to reflect a requirement of a replicating template for B to activate late transcription and the provision of increased *B* gene dosage (Kalionis *et al.*, 1986b).

A topological requirement for late transcription was entertained by Kalionis *et al.* (1986b) since a superinfecting, immunity-insensitive phage could not rescue late functions from a prophage (Hocking and Egan, 1982a). Increased B protein was shown to activate p_V transcription of a prophage, by quantitation of RNA using a *V* specific probe and slot blot analysis. This ability of B protein to activate late transcription on a prophage demonstrated that there was no role for a specific template topology provided by phage replication, in order for late gene transcription to occur.

Since there was shown to be no topological requirement in activation of late transcription, the role of replication was presumably related to the requirement for an increase in template numbers. A consequence of replication would be the provision of increased *B* gene dosage, and despite the fact that *B* gene transcription appears to be autoregulated (Kalionis *et al.*, 1986b), an increase in gene dosage should result in increased levels of free B protein (Appendix II). To investigate whether an increase in *B* gene dosage did activate late transcription, the *B* gene under its own control was cloned onto the multicopy p_V*galK* plasmid. Increased *B* gene dosage was demonstrated to strongly activate p_V transcription and it was concluded that one of the roles of replication in activation of late transcription was the provision of increased *B* gene dosage.

The increase in *B* gene dosage as a result of replication also had the potential to derepress *B* gene transcription, by virtue of the increase in the number of *B* operator sequences titrating the pool of the repressor of *B* gene transcription (Kalionis *et al.*, 1986b). Kalionis *et al.* (1986b) identified potential operator sites overlapping the *B* promoter. To investigate whether an increase in the concentration of the presumptive *B* operator could titrate prophage *B* gene transcription and therefore activate late transcription, a restriction fragment carrying the presumptive operator sites was cloned onto the multicopy p_V*galK* plasmid. This resulted in activation of p_V transcription when the construct was introduced into a 186 lysogen (Section 6.2.3) and presumably the

presence of the operator sites was able to derepress *B* gene transcription sufficiently for the activation of p_V transcription. Thus one of the roles of replication in activating late transcription was concluded to be the derepression of *B* gene transcription by the newly replicated operator sequences, sufficient to activate late transcription, with the replicative increase in *B* gene dosage a necessary event to amplify the initial effect of derepression. This also indicated that replication may provide the initial event in activating late gene transcription, by the derepression of *B* gene transcription due to the titration of the B gene repressor.

The ability of excess operator sequences to induce transcription of the single copy of the *B* gene on the prophage, sufficient for activation of late transcription, also suggested as a general principle, that replication could be used to control expression of an autogenously controlled gene, since replication should be able to derepress transcription by titration of the repressor, at least transiently, which may be all that is needed to initiate a cascade of events.

The inability of the *B* expression plasmid pEC434 to transform a 186 lysogen indicated the lethality of increased *B* expression to a 186 lysogen and provided the basis for the potential isolation of host mutants specifically blocked in activation of late transcription. The lethality of increased *B* expression from plasmid pEC434 in a 186 lysogen is investigated as a method for the isolation of host mutants blocked in late activation in Appendix I.

RESULTS: SECTION III

CHAPTER 7

INVESTIGATION OF THE AUTOGENOUS REPRESSIVE CONTROL OF B GENE TRANSCRIPTION

CHAPTER 7. INVESTIGATION OF THE AUTOGENOUS REPRESSIVE CONTROL OF *B* GENE TRANSCRIPTION.

7.1 INTRODUCTION.

B protein has been shown to be the only phage function required directly to activate transcription of the late promoter p_V (Section 5.2.4). However, there is no evidence of activation of late transcription in the lysogenic state either from *in vivo* hybridization studies (Finnegan and Egan, 1981) or from studies with the p_V*galK* plasmid (Section 4.2.4(a)). This is despite the fact that B protein was presumably present and active in repressing its own transcription in the lysogenic state (Kalionis *et al.*, 1986b). This raised the question as to why the prophage copy of the *B* gene was not active in activation of late transcription. This absence of late gene transcription in the lysogen may have been due to an insufficient concentration of B protein for activation, a result of the autogenous control of *B* gene transcription, or alternatively could have been due to the involvement of a function expressed in the lysogenic state, which directly antagonized B protein activating late transcription.

The aim of this chapter was to investigate whether the inability of B protein to activate late transcription in the lysogenic state was a direct consequence of a low B protein concentration due to autogenous repressive control of *B* gene transcription or due to the expression of a lysogenic function which directly antagonized the ability of B protein to activate late transcription.

This approach required as a first step the identification of the phage functions directly involved in the repression of *B* gene transcription. In its simplest form, the repressive control of *B* gene transcription would be direct, with the B protein behaving as a classical repressor at the *B* promoter (Kalionis *et al.*, 1986b). Alternatively, indirect control could be through a host or another phage function, with the expression or activity of that protein under *B* gene control. Identification of the phage functions involved in repression of *B* gene transcription will demonstrate if the autogenous control of transcription is direct or indirect. The ability of the *B* gene to activate late transcription under conditions of repression of transcription would indicate that the autogenous control was not responsible for the inability of a prophage copy of the *B* gene to activate late

transcription. Alternatively, the inability of the *B* gene to activate late transcription under conditions of repression of transcription would indicate that the autogenous control was likely to be responsible for the inability of a prophage copy of the *B* gene to activate late transcription.

This chapter describes the identification of the phage functions required to repress *B* gene transcription and the demonstration that the inability of the prophage *B* gene to activate late transcription is likely to be a consequence of the repressive control of *B* gene transcription.

7.2 RESULTS AND DISCUSSION.

7.2.1 Investigation of repression of the *B* promoter fused to the *galK* gene.

7.2.1(a) Fusion of the *B* promoter to the *galK* gene.

To identify the phage functions repressing *B* gene transcription *in vivo*, the *B* promoter was cloned before the reporter gene *galK* and the phage functions necessary to achieve repression of transcription of the *galK* gene determined.

In the cloning procedure described in Section 6.2.3, the 290 bp *Pst*I-*Nde*I (65.5% to 66.5%) restriction fragment encoding the *B* promoter (Kalionis *et al.*, 1986b) was cloned into the *galK* vector pKO2. This was achieved by cloning the 290 bp *Pst*I-*Nde*I (65.5% to 66.5%) restriction fragment, the termini blunted by treatment with the Klenow fragment (Section 2.28.2), into the *Sac*I (blunted) site of plasmid pKO2. Plasmid constructs with the *B* promoter transcribing the *galK* gene were evident in the *galK*$^-$ strain E863, since they had a red colony phenotype on MacConkey indicator plates (as opposed to the white colony phenotype of the constructs with the *B* promoter oriented towards the *bla* gene of the plasmid, as demonstrated for pEC436). Diagnostic restriction analysis on plasmid DNA (Section 2.28.3(c)) confirmed the identity and orientation of one such p_B*galK* construct (data not shown), which was designated pEC445.

The *B* promoter of plasmid pEC445 gave 247 *galK* units at 30^{o}C when assayed in the *galK*$^-$ host E864. The activity of the p_B promoter on plasmid pEC445 was considered to be significant, as it was approximately three-fold as active as the wild-type *lac* promoter

on pKL600 in the absence of *lacI* repression (90.2 units at 37°C; Section 4.2.4(a)). Therefore the *B* promoter on plasmid pEC445 was concluded to be significantly active *in vivo* in the absence of phage functions and the first step in studying control of *B* gene transcription with the $p_B galK$ construct was to determine whether the B protein alone, when expressed at high levels from an expression plasmid, was able to repress the *B* promoter on the $p_B galK$ construct.

7.2.1 (b) Investigation of direct repression of the $p_B galK$ plasmid by B protein alone.

To investigate whether B protein alone was capable of repressing transcription from the *B* promoter, B protein was supplied *in trans* from the *B* expression plasmid pEC434 to the $p_B galK$ construct and the activity of the *B* promoter determined.

To supply B protein from plasmid pEC434 to the $p_B galK$ plasmid required the modification of the antibiotic resistance of the $p_B galK$ plasmid pEC445, since both plasmids conferred resistance to ampicillin. In a similar manner to the construction of plasmids pEC432 and pEC433, which are kanamycin resistant derivatives of plasmid pKO2 and the $p_V galK$ plasmid pEC428 (Section 5.2.3), a kanamycin resistant derivative of the $p_B galK$ plasmid pEC445 was created by cloning (Section 2.28.3) the 1.3 kb *Eco*RI to *Sma*I restriction fragment (encoding the kanamycin resistance gene) from pKC7 (Rao and Rogers, 1979) into the large *Eco*RI to *Xmn*I fragment of the $p_B galK$ plasmid pEC445. This gave the kanamycin resistant derivative pEC446, the identity of which was confirmed by diagnostic restriction analysis of plasmid DNA (data not shown; Section 2.28.3(c)).

The $p_B galK$ plasmid pEC446 was transformed into the non-permissive, F'$lacI^q$ strain JM105 carrying either the *B* expression plasmid pEC434, the equivalent *Bam*17 expression plasmid pEC435 or the parent vector pEC612 (Section 2.3.1). Under conditions of $lacI^q$ repression of the *B* expression plasmid pEC434, there was no evidence of repression of transcription from the p_B promoter, as plasmid pEC446 gave equivalent galactokinase activity in the presence of the *B* expression plasmid pEC434 (242 units at 30°C; Table 7.1) as it did in the absence of B protein (i.e. in the presence of the parent plasmid pEC612) under the same conditions (253 units at 30°C; Table 7.1).

Under conditions of induction of *B* expression from plasmid pEC434 (42°C and 1 mM IPTG), there was some evidence of a reduction in galactokinase activity from the

TABLE 7.1

Investigation of repression of the *B* promoter by B protein alone.

Culture[a]	Conditions		Galactokinase activity[b]
$p_B galK$ pp_{trc}	30°C		253±18
$p_B galK$ $pp_{trc}B^+$	30°C		242±12
$p_B galK$ pp_{trc}	42°C/IPTG	120 min	320±23
$p_B galK$ $pp_{trc}B^+$	42°C/IPTG	120 min	190±23
$p_B galK$ $pp_{trc}Bam$	42°C/IPTG	120 min	263±27

Notes to Table 7.2

a. Cultures of the non-permissive strain JM105 ($galK^+$ F' $lacI^q$) carrying the $p_B galK$ plasmid (pEC446), and either the *B* expression plasmid $pp_{trc}B^+$ (pEC434), the *Bam* expression plasmid $pp_{trc}Bam$ (pEC435) or the pp_{trc} plasmid (pEC612) were grown at 30°C in M13 minimal media supplemented with the appropriate antibiotics. Non-induced cultures were grown to A_{650}=0.3-0.6. For induction of expression from plasmids, the cultures were grown at 30°C in M13 minimal media supplemented with antibiotics to A_{650}=0.2, transferred to 42°C and IPTG added to 1 mM final concentration and samples were taken at the indicated times. Strain JM105 has a $galK^+$ genotype and the galactokinase units presented represent the units above that recorded for JM105 carring the pp_{trc} plasmid (pEC612) (Section 2.3.1). The $p_B galK$ plasmid (pEC446), the $pp_{trc}B^+$ plasmid (pEC434) and the $pp_{trc}Bam$ plasmid pEC435 are described in Section 2.3.2, and the pp_{trc} plasmid (pEC612) is described in Section 2.3.1.

b. The galactokinase activity was determined as described in Section 2.18, and are the mean of at least three separate experiments, with the standard deviations shown.

p_B*galK* plasmid pEC446 (190 units; Table 7.1), as compared to the p_B*galK* plasmid in the presence of the *Bam*17 expression plasmid pEC435 (263 units; Table 7.1) or the parent plasmid pEC612 under the same conditions (320 units; Table 7.1). This reduction in galactokinase activity observed from pEC446 upon induction with the *B* expression plasmid pEC434 was not a non-specific effect of increased *B* expression, as the wild-type *lac* promoter cloned into the kanamycin resistant derivative of plasmid pKO2 (pEC432), showed no such reduction in galactokinase activity under the same conditions (data not shown). However, the extent of repression of *B* gene transcription on the p_B*galK* plasmid on induction of the *B* expression plasmid pEC434 did not parallel that observed in the lysogenic state for the difference in transcript levels between *B* and *Bam*17 genes.[The degree of repression of *B* gene transcription in the lysogenic state has been estimated at approximately ten fold, as determined by quantitation of transcript levels in *B* and *Bam*17 lysogens using slot blot analysis of total RNA with a *B*-specific probe (Section 7.2.2).] It was therefore uncertain whether the B protein supplied *in trans* and at high concentration was able to significantly repress the p_B promoter on the multicopy *galK* plasmid. Since repression of the *B* gene was observed in the prophage state (Kalionis *et al.*, 1986b), it was decided to investigate control of *B* gene transcription under conditions that reproduced the single copy status of the prophage *B* gene, by cloning the entire *B* gene onto a plasmid maintained at single copy.

7.2.2 Investigation of autogenous repression of transcription of the *B* gene in single copy.

To reproduce the single copy status of the *B* gene present in a lysogen, the entire *B* and *Bam*17 genes were cloned onto a plasmid vector maintained at one chromosomal equivalent per cell and the *in vivo* transcript levels determined to investigate whether there was any evidence of repression of *B* gene transcription. The absence of repression by B protein under such conditions would suggest that control of *B* gene transcription was likely to be indirect, with control by other phage functions expressed by the prophage.

Plasmid pOU61 is a R1-derived replicon, the copy number of which is under temperature control (Larsen *et al.*, 1984; Section 2.3.1). At 30°C plasmid pOU61 is maintained at one chromosomal equivalent per cell and plasmid stability is maintained through the presence of the partitioning locus *parB* (Rasmussen *et al.*, 1987). To clone the

B and *Bam*17 genes into pOU61, the unique *Bam*HI site on the plasmid was employed. The source of the DNA encoding the *B* and *Bam*17 genes were M13mp18 clones mEC412 and mEC413, respectively (Section 2.4.2). These M13 clones were constructed so as to allow the *B* and *Bam*17 genes to be isolated using a *Bam*HI restriction fragment, to facilitate cloning into the *Bam*HI site of pOU61. Clones mEC412 and mEC413 were constructed from the M13mp19 clones mEC648 and mEC649 (Section 2.4.2), by cloning the 1.4 kb *Sac*I restriction fragment from these clones (utilizing the *Sac*I site at 63.0% and the *Sac*I site in the M13mp19 polylinker) into the *Sac*I site of M13mp18. Complementarity tests (Section 2.28.4(g)) allowed the identification of clones mEC412 and mEC413, which contained the *r*-stand of the 1.4 kb *Sac*I (63.0% to 67.6%) restriction fragment encoding the *B* and *Bam*17 genes respectively, flanked on both sides by *Bam*HI sites.

The 1.4 kb *Bam*HI restriction fragments encoding the wild-type *B* gene and the *Bam*17 gene were isolated from RF DNA (Section 2.22) of clones mEC412 and mEC413 by agarose gel electorphoresis (Section 2.26.1) and cloned into the *Bam*HI site of plasmid pOU61 (Section 2.28.3). This yielded plasmids pEC447 and pEC448 respectively, with the *B* and *Bam*17 genes cloned in an orientation such that transcription of the *B* and *Bam*17 genes was directed towards the *bla* gene of pOU61 (Section 2.28.3(c)).

To determine whether the cloned *B* gene product on plasmid pEC447 was able to repress its own transcription, the level of *B* gene transcription from the single copy plasmids pEC447 and pEC448 was determined using slot blot analysis of total RNA and hybridization to a probe specific for *B* gene transcription. Total RNA was prepared (Sections 2.32.1 and 2.32.2) from cultures of the non-permissive strain E251 carrying the single copy *B* plasmid pEC447 or the corresponding *Bam*17 plasmid pEC448. Total RNA was also prepared from non-induced cultures of 186 *cIts* and 186 *cIts Bam*17 lysogens of strain E251, for comparison with the levels of *B* and *Bam*17 transcription in the lysogenic state. The *B*-specific probe used was a 350 bp *Nde*I-*Bam*HI (66.5% to 67.6%; *Bam*HI is in the M13 polylinker) restriction fragment with the *r*-strand labelled, prepared by synthesizing a labelled, complementary copy (Section 2.30.1) of the M13 clone mEC648 (Section 2.4.2). The results of the slot blot analysis are shown in Figure 7.1 and the level of transcription quantitated by laser densitometry presented in the legend to Figure 7.1.

Figure 7.1. (cont.)

Quantitation of slot blot analysis of *B* gene transcription.

Culture[a]	Conditions[b]	Level of Hybridization per ug of total RNA[c]
(186 *cIts*)	39°C 35 min (1)	$1.3x10^6$
(186 *cIts*)	30°C (2)	$1.8x10^3$
(186 *cIts Bam*17)	30°C (3)	$1.7x10^4$
pOU61B^+	30°C (4,5)	$8.2x10^4$
pOU61*Bam*	30°C (6,7)	$5.0x10^4$
pOU61	30°C (8)	Not detectable

a. Cultures are as described above. The single copy *B* plasmid pOU61B^+ (pEC447), the single copy *Bam* plasmid pOU61*Bam* (pEC448) and the parent plasmid pOU61 are described in detail in Section 2.3.

b. Corresponding tracks in the autoradiograph are indicated in brackets.

c. The level of hybridization represents the total peak area obtained by laser densitometric scanning of autoradiographs of RNA slot blot hybridized with the *B* specific probe as described in Section 2.32.4(c). Not detectable represents no signal above background. Results presented are the average of duplicate RNA preparations hybridized with the probe.

Figure 7.1 Level of *B* gene transcription from single copy plasmid clones.

Slot blot analysis of the level of hybridization of a *B* specific probe to RNA prepared from single copy *B* and *Bam* clones, in comparison to the level of hybridization from B^+ and *Bam* lysogens. RNA was prepared from cultures of 186 *cIts* and 186 *cIts Bam*17 lysogens of the non-permissive strain E251 (E252 and E264 respectively). Non-induced cultures were grown in L broth at 30°C to A_{600}=0.8 and RNA prepared (Section 2.32). For heat-induction of E251 (186 *cIts*), the cultures were grown at 30°C to A_{600}=0.8 and transferred to 39°C and RNA prepared 35 minutes after induction (Section 2.32). RNA was also prepared from cultures grown in L broth at 30°C of the non-permissive strain E251 carrying the single copy *B* clone pEC447, the single copy *Bam* clone pEC448 or the parent plasmid pOU61.

Glyoxal denatured RNA (Section 2.32.3) was analysed by slot blot analysis (Section 2.32.4) with a 350 b *Nde*I-*Bam*HI (66.5% to 67.6%; *Bam*HI is in the M13 polylinker) probe (2.8×10^6 cpm) specific for *B* transcription prepared from mEC648 (Section 2.31.2; Section 2.4.2). The position of the probe on the DNA sequence of the *Pst*I-*Xho*I (65.5% to 67.6%) region of clone mEC648 is shown below the autoradiograph. Duplicate samples of RNA were loaded from each culture (5 ug unless otherwise stated). The filter was autoradiographed at -80°C with an intensifying screen for 7 days.

RNA samples:

1. RNA (1 ug) from a 186 *cIts* lysogen 35 min after heat-induction.

2. RNA from a non-induced 186 *cIts* lysogen.

3. RNA from a non-induced 186 *cIts* *Bam*17 lysogen.

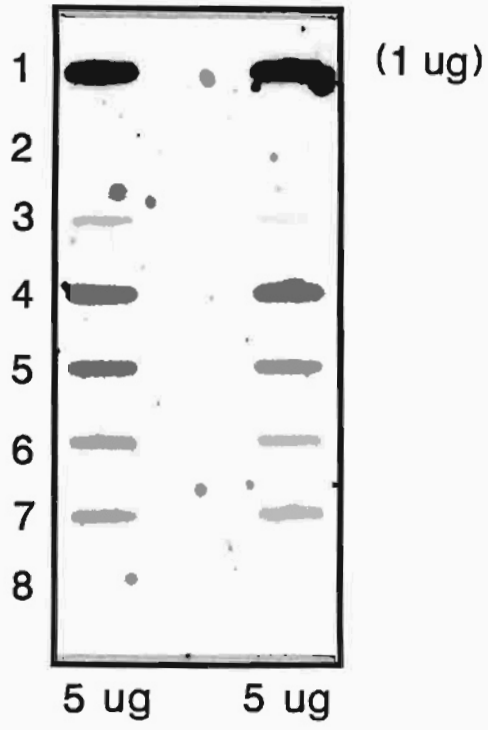
4. RNA from a culture carrying the single copy *B* plasmid pEC447.

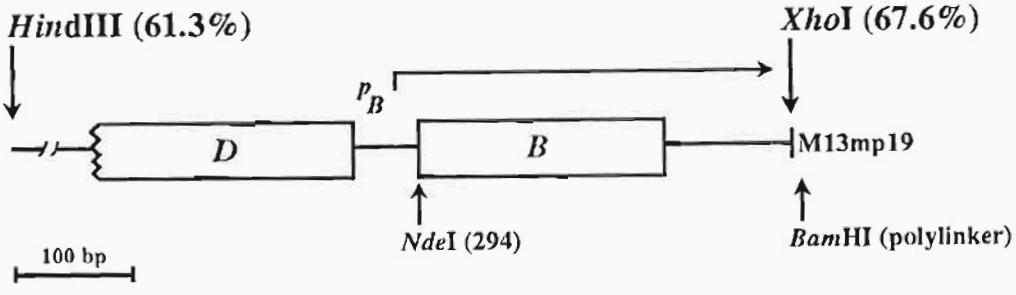
5. As described for 4.

6. RNA from a culture carrying the single copy *Bam* plasmid pEC448.

7. As described for 6.

8. RNA from a culture carrying the single copy parent plasmid pOU61


1
2
3
4
5
6
7
8
(1 ug)
5 ug
5 ug


HindIII (61.3%)
XhoI (67.6%)
P_B
D
B
M13mp19
NdeI (294)
BamHI (polylinker)
100 bp

In the lysogenic state, the transcript encoding the *Bam*17 gene (Figure 7.1, track 3) was found to be increased approximately 10 fold (by laser densitometric analysis of the autoradiograph, as shown in the legend to Figure 7.1) over the transcript encoding the wild-type *B* gene (Figure 7.1, track 2), consistent with the results of Kalionis *et al.* (1986b). However, no such difference in the level of transcription was observed between the *B* gene on the single copy plasmid pEC447 (Figure 7.1, tracks 4 and 5) and the corresponding *Bam*17 gene on plasmid pEC448 (Figure 7.1, tracks 6 and 7). The level of *B* gene transcription detected from the single copy *B* plasmid pEC447 was found to be increased 1.7 fold over the level of transcription detected from the *Bam*17 plasmid pEC448 (legend to Figure 7.1) and it was concluded that there was no evidence of direct autogenous control by the *B* gene product in this system.

Additionally, the level of transcription detected from the single copy *B* plasmid pEC447 was found to be elevated approximately five fold over the level of transcription detected from the 186 *cIts Bam*17 lysogen, even though both genes should have been at an equivalent dosage of one per chromosome. This elevated level of *B* gene transcripion from the single copy plasmid clone of the *B* gene may have been due to plasmid-derived transcription reading into the cloned DNA and contributing to the transcription of the *B* gene. In this regard the transcriptional pattern of pOU61 is poorly defined (Larsen *et al.*, 1984). Alternatively the *B* promoter on the single copy plasmid may have been more active than in a chromosomal context of a prophage, since many prokaryotic promoters are known to be affected by factors such as DNA supercoiling (Pruss and Drlica, 1989; Albright *et al.*, 1988). It was also conceivable that the level of transcription detected from a 186 *cIts Bam*17 lysogen may not have represented a situation in which *B* gene transcription was fully derepressed, and that the increased level of transcription detected from plasmids pEC447 and pEC448 was due to the fact that the single copy *B* plasmid represented a situation where transcription from p_B was fully derepressed (discussed in Section 7.2.4).

Nevertheless, it was concluded that there was no evidence of direct autogenous control by the *B* gene product even when present in single copy. Although the possibility existed that the level of plasmid-derived transcription (reading into the cloned *B* and *Bam* genes) from pOU61 was such that it masked repression of *B* gene transcription on the

single copy plasmid, it was decided to determine if the repression of *B* gene transcription was in fact indirect, with some factor present in a 186 lysogen and under *B* gene control, required for repression of transcription from the *B* promoter, by introducing the single copy *B* plasmid into a lysogen and investigating repression of transcription.

7.2.3 Evidence for indirect control of *B* gene transcription and mapping of the presumptive repressor.

7.2.3(a) Evidence for indirect control.

To determine if a lysogenic function was involved in repression of *B* gene transcription, the single copy *B* plasmid was introduced into a lysogen, and as a first step, the repression of *B* gene transcription investigated by determining whether the *B* gene on the single copy *B* plasmid was able to activate transcription from the p_V*galK* plasmid.

Consistent with the absence of repression of transcription of the *B* gene on the single copy plasmid pEC447, it was found that B protein expressed from this plasmid, when maintained at single copy (30°C), was still capable of activating p_V transcription on the p_V*galK* plasmid pEC433 in a non-lysogen (34.9 units, Table 7.2). As expected, this activation remained *B* gene dependent as the single copy *Bam* plasmid pEC448 did not activate p_V transcription (1.3 units, Table 7.2). However, when the single copy *B* plasmid pEC447 was introduced into a 186 *cIts* lysogen it was found that p_V transcription on the p_V*galK* plasmid pEC433 was no longer activated (1.3 units; Table 7.2). It was concluded that the *B* gene on plasmid pEC447 was not competent for activation of late transcription in a 186 lysogen, presumably due to either repression of *B* gene transcription on the single copy plasmid pEC447 or to the expression of some function which blocked *B*-dependent activation of late transcription directly.

7.2.3(b) Mapping of the lysogenic factor blocking *B*-dependent activation of late transcription.

The B protein expressed from the single copy *B* plasmid pEC447 was found to be capable of activating the p_V*galK* plasmid pEC433 in a P2 lysogen of the non-permissive *galT*$^-$*K*$^-$ strain E536 (20.5 units; Table 7.3), although it was slightly reduced with respect to

TABLE 7.2

Activation of p_V transcription by the *B* gene in single copy.

Culture[a]	Galactokinase activity[b]
p_V*galK* pOU61B^+	34.9±6.1
p_V*galK* pOU61*Bam*	1.3±0.2
(186 *cIts*) p_V*galK* pOU61B^+	1.3±0.4

Notes to Table 7.3

a. Cultures of the non-permissive strain E864 or the corresponding 186 *cIts* lysogen carrying the p_V*galK* plasmid (pEC433), and either the single copy *B* plasmid pOU61B^+ (pEC447) or the single copy *Bam* plasmid pOU61*Bam* (pEC448) were grown at 30°C in M63 supplemented media with the appropriate antibiotics to A_{650}=0.3-0.6, samples taken and the galactokinase activity determined. The p_V*galK* plasmid pEC433, the single copy *B* plasmid pOU61B^+ (pEC447) and the single copy *Bam* plasmid pOU61*Bam* (pEC448) are described in detail in Section 2.3.2.

b. The galactokinase activity was determined as described in Section 2.18, and are the mean of three separate experiments, with the standard deviations shown.

TABLE 7.3

Activation of p_V transcription by the *B* gene in single copy in P2 and Hy5 lysogens.

Culture[a]	Galactokinase activity[b]
p_V*galK* pOU61B^+	35.8±1.6
(P2 *C5*) p_V*galK* pOU61B^+	20.5±2.7
(P2 *C5*) p_V*galK* pOU61*Bam*	1.6±0.3
(Hy5 *cIts*) p_V*galK* pOU61B^+	1.1±0.4

Notes to Table 7.4

a. Cultures of the non-permissive strain E536 (*galK⁻T⁻*) or the corresponding Hy5 *cIts* (E857) or P2 *C5 nip* (E813) (*C5* is a temperature-sensitive repressor; Bertani, 1968; *nip* improves yield of phage after thermal induction; Calendar *et al.*, 1972) lysogens carrying the p_V*galK* plasmid (pEC433), and either the single copy *B* plasmid pOU61B^+ (pEC447) or the single copy *Bam* plasmid pOU61*Bam* (pEC448) were grown at 30°C in M63 supplemented media with the appropriate antibiotics to A_{650}=0.3-0.6, samples taken and the galactokinase activity determined. The p_V*galK* plasmid pEC433, the pOU61B^+ plasmid pEC447 and the pOU61*Bam* plasmid pEC448 are described in detail in Section 2.3.2.

b. The galactokinase activity was determined as described in Section 2.18, and are the average of two separate experiments.

the galactokinase activity in the corresponding non-lysogen E536 (35.8 units; Table 7.3), and demonstrated that a P2 lysogen did not express the same inhibitor function as a 186 lysogen. The activation of the p_V*galK* plasmid in the P2 lysogen remained *B*-dependent, as the corresponding *Bam* plasmid pEC448 was unable to activate the p_V*galK* plasmid in the same P2 lysogen (1.6 units; Table 7.3). The hybrid phage Hy5 contains the right-hand end of the 186 genome from 65.4% and the left-hand end of the P2 genome (Hocking and Egan, 1982b). The single copy *B* plasmid pEC447 was found to be unable to activate galactokinase expression of the p_V*galK* plasmid pEC433 in a Hy5 *cIts* lysogen of E536 at 30°C (1.1 units; Table 7.3) and demonstrated that the lysogenic function blocking *B*-dependent of activation of late transcription was located to the right of 65.4% on the 186 chromosome.

The 186 p_L promoter which is active in the lysogenic state and transcribes genes *cI*, *int* and *CP69* (Dodd *et al.*, 1990; Kalionis *et al.*, 1986a) is located to the right of 65.4% on the 186 chromosome. It was possible that the factor blocking the ability of the *B* gene of plasmid pEC447 to activate late transcription was located on this transcript initiating from the p_L promoter. The major control region of 186 is conveniently bounded by *Xho*I (67.6%) and *Bgl*II (79.6%) restriction sites and encodes genes *cI*, *int* and *CP69* under control of the lysogenic promoter p_L and genes *apl*, *cII*, *fil* and *dhr* under control of the early lytic promoter p_R (Kalionis *et al.*, 1986a; Richardson *et al.*, 1989; Richardson and Egan, 1989; Dodd *et al.*, 1990). The *Xho*I-*Bgl*II (67.6% to 79.6%) region is shown diagrammatically in Figure 7.2. To determine if the 3.6 kb *Xho*I-*Bgl*II (67.6% to 79.6%) region expressed the function blocking activation of late transcription, this restriction fragment was cloned onto plasmid pACYC184, which is a p15A-derived replicon conferring resistance to chloramphenicol and tetracycline (Chang and Cohen, 1978; Section 2.3.1), and the ability of this plasmid clone to block *B*-dependent activation of transcription was tested. The 3.6 kb *Xho*I-*Bgl*II (67.6% to 79.6%) restriction fragment was isolated from wild-type 186 DNA and cloned into the *Sal*I and *Bam*HI sites of plasmid pACYC184 to yield plasmid pEC449 (Section 2.28.3(c)).

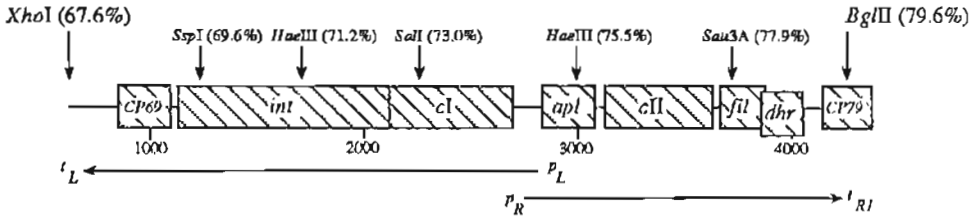
Plasmid pEC449 was first shown to be active in expressing the *cI* gene product, since this plasmid conferred immunity to 186 *cI*10 phage as determined by cross-streaking (Section 2.15.1). Transformation of plasmid pEC449 into the *galK*⁻ strain E864 carrying

Figure 7.2 Diagrammatic representation of the 3.6 kb *Xho*I-*Bgl*II (67.6% to 79.6%) region and the restriction fragments cloned into plasmid pACYC184.

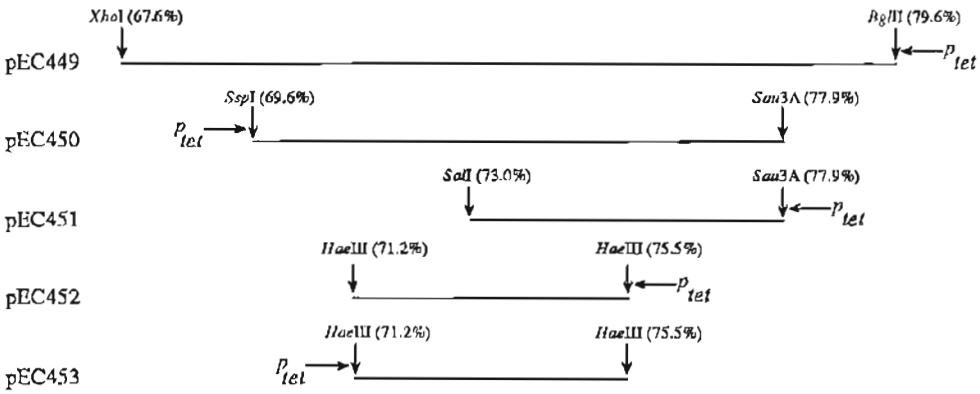
(a) Representation of the predicted coding regions and transcriptional control signals as determined for the *Xho*I-*Bgl*II (67.6% to 79.6%) region (Kalionis *et al.*, 1986a; Richardson *et al.*, 1989). Functions of the genes are described in Sections 1.3.1 and 1.4.2. Genes of unknown function are designated *CP* followed by the chromosomal coordinate approximating the predicted initiation codon of the gene. Promoters (*p*) and termination signals (*t*) are indicated. The restriction sites used to clone fragments into plasmid pACYC184 and their chromosomal positions are shown. DNA sequence co-ordinates from the *Pst*I site at 65.5% are shown below the map.

(b) Restriction fragments derived from the *Xho*I-*Bgl*II (67.6% to 79.6%) region cloned into plasmid pACYC184, with their plasmid construct designation shown to the left. The orientation of the insert DNA in plasmid pACYC184 is indicated by the direction of transcription from the p_{tet} promoter of the plasmid, and gives an indication as to the genes expressed from this promoter on the plasmid.

(a)


XhoI (67.6%)
SspI (69.6%)
HaeIII (71.2%)
SalI (73.0%)
HaeIII (75.5%)
Sau3A (77.9%)
BglII (79.6%)
CP69
int
cI
apl
cII
fil
dhr
CI79
1000
2000
3000
4000
t_L
p_L
p_R
t_{R1}

(b)


XhoI (67.6%)
BglII (79.6%)
pEC449
p_{tet}
SspI (69.6%)
Sau3A (77.9%)
pEC450
p_{tet}
SalI (73.0%)
Sau3A (77.9%)
pEC451
p_{tet}
HaeIII (71.2%)
HaeIII (75.5%)
pEC452
p_{tet}
HaeIII (71.2%)
HaeIII (75.5%)
pEC453
p_{tet}

the single copy *B* plasmid pEC447 and the p_V*galK* plasmid pEC433 resulted in white colonies on the appropriate MacConkey indicator plates, while the parent plasmid pACYC184 gave red colonies on transformation (Table 7.4). Determination of the galactokinase activities confirmed the colony phenotypes. The presence of the plasmid clone pEC449 prevented activation of p_V transcription by plasmid pEC447 (1.6 units; Table 7.4), while the parent plasmid pACYC184 had no effect on the ability of B protein from pEC447 to activate p_V*galK* transcription (40.1 units; Table 7.4). It was therefore concluded that the function blocking activation of late transcription was located on the *Xho*I-*Bgl*II (67.6% to 79.6%) region. The function was also considered likely to be located on the lysogenic transcript initiating from the p_L promoter, since the p_R promoter on the clone is repressed by CI protein (Dodd *et al.*, 1990), and therefore the genes *c*I, *int* and *CP69* were of immediate interest.

To locate the inhibitor gene on plasmid pEC449, smaller restriction fragments contained within the *Xho*I-*Bgl*II (67.6% to 79.6%) region were cloned into plasmid pACYC184 and their ability to express the function blocking activation of p_V transcription was tested by transforming (Section 2.15.2) these plasmid clones into strain E864, carrying the single copy *B* plasmid pEC447 and the p_V*galK* plasmid pEC433, and observing the phenotype on MacConkey indicator plates. The various restriction fragments cloned are shown in Figure 7.2, and the respective phenotypes of these clones, on transformation into E864 carrying pEC447 and pEC433, are shown in Table 7.4.

The 2.5 kb *Ssp*I-*Sau*3A (69.6% to 77.9%) restriction fragment encodes genes *c*I, *apl* and *c*II, and portions of genes *int* and *fil* (Figure 7.2). This fragment was isolated from DNA of the *Xho*I-*Bgl*II (67.6% to 79.6%) plasmid clone pEC449 (Section 2.21.2) and cloned (Section 2.28.3) into the *Eco*RV and *Bam*HI sites of pACYC184, to yield plasmid pEC450 (Section 2.28.3(c)). This plasmid, which conferred immunity to 186 *c*I10 phage (Section 2.15.1), was found to express the inhibitor function, since it gave white colonies on MacConkey indicator plates upon transformation into strain E864 (*galK*$^-$) carrying plasmids pEC447 and pEC433 (Table 7.4). This demonstrated that gene *CP69* was not the function blocking activation of late transcriptiom and most probably that the *int* gene was not this function, as only a portion of the *int* gene was located on this restriction fragment. It therefore appeared that the *c*I gene product was the likely function, since the p_R

TABLE 7.4

Mapping of the lysogenic factor blocking *B*-dependent activation of p_V transcription[a].

Plasmid	Genes encoded[b]	Galactokinase activity[c]	Plate phenotype
pACYC184	-	40.1±4.0	Red
pEC449	*CP69,int,cI, apl,cII,fil,dhr*	1.6±0.2	White
pEC450	*cI,apl,cII*	-	White
pEC451	*apl,cII*	-	Red
pEC452	*cI*	-	White
pEC453	*cI*	-	White

Notes to Table 7.5

a. For determination of galactokinase activity, cultures of the non-permissive strain E864 carrying the p_V*galK* plasmid (pEC433), the single copy *B* plasmid pOU61B^+ (pEC447) and either the plasmid pACYC184 or the 3.6 kb *Xho*I-*Bgl*II (67.6% to 79.6%) clone pEC449 were grown in M63 supplemented media with kanamycin (50 ug/ml), ampicillin (50 ug/ml) and chloramphenicol (30 ug/ml) at 30°C to A_{650}=0.3-0.6 and the galactokinase activity determined. The p_V*galK* plasmid pEC433, the pOU61B^+ plasmid pEC447, the plasmid pACYC184 and the 3.6 kb *Xho*I-*Bgl*II (67.6% to 79.6%) plasmid clone pEC449 are described in detail in Section 2.3.

For determination of plate phenotypes on transformation, cultures of the non-permissive strain E864 carrying the p_V*galK* plasmid (pEC433) and the single copy *B* plasmid pOU61B^+ (pEC447) were grown in L broth at 30°C to A_{600}=0.4-0.6, transformed (Section 2.15.2) with DNA of one of the plasmids pACYC184, pEC449, pEC450, pEC451, pEC452 or pEC453 and plated onto MacConkey indicator plates supplemented with kanamycin (50 ug/ml), ampicillin (50 ug/ml) and chloramphenicol (30 ug/ml) and incubated at 30°C for 18-24 hours and the plate phenotypes noted.

b. Genes completely contained on the restriction fragments cloned to construct the plasmid clones, based on the sequence analysis of Kalionis *et al.* (1986a) and Richardson *et al.* (1988). Plasmid clones pEC449, pEC450, pEC451, pEC452 and pEC453 are described in detail in Section 2.3.2.

c. The galactokinase activity was determined as described in Section 2.18, and are the average of two separate experiments.

promoter was expected to be repressed in the presence of CI protein (Dodd *et al.*, 1990) and genes *apl* and *c*II not expressed.

The 1.5 kb *Sal*I-*Sau*3A (73.0% to 77.9%) restriction fragment encodes genes *apl* and *c*II and the amino-terminal end of the *c*I gene (Figure 7.2). This restriction fragment from pEC449 DNA was cloned (Section 2.28.3) into the *Sal*I and *Bam*HI sites of pACYC184 to give plasmid pEC451 (Section 2.28.3(c)). Transformation of this plasmid did not confer resistance to 186 *c*I10 infection as determined by cross-streaking (Section 2.15.1), consistent with the *c*I gene product not being active. This plasmid did not encode the function blocking activation of p_V transcription, as transformation of plasmid pEC451 into the test strain gave red colonies on MacConkey galactose plates (Table 7.4), confirming the *c*I gene as the function blocking activation of late transcription by the single copy plasmid pEC447.

The 1.3 kb *Hae*III-*Hae*III (71.2% to 75.5%) restriction fragment encodes the *c*I gene and the 5'-ends of the *int* and *apl* genes (Figure 7.2). This restriction fragment was isolated from plasmid pEC449 DNA and cloned (Section 2.28.3) into the *Eco*RV site of pACYC184. This yielded plasmid pEC452, the orientation of which was such that *c*I transcription was opposed to transcription from p_{tet} of the plasmid (Figure 7.2), and plasmid pEC453, with the inserted DNA in the opposite orientation such that transcription of the *c*I gene was under p_{tet} control (as well as p_L control; Figure 7.2). Both plasmids pEC452 and pEC453 expressed functional CI protein, as determined by their ability to confer resistance to 186 *c*I10 by cross-streaking (Section 2.15.1).

Both plasmids pEC452 and pEC453 gave white colony phenotypes on transformation into the test strain carrying plasmids pEC447 and pEC433 (Table 7.4), indicating that the *c*I gene product was reponsible for the inability of B protein from plasmid pEC447 to activate p_V transcription, presuming that the portion of the *int* gene product contained on the 1.3 kb *Hae*III-*Hae*III (71.2% to 75.5%) restriction fragment was not active. There is the theoretical possibility that the fragment of the *int* gene encoded could be the active component (or some unidentified protein encoded on the fragment) and final proof will come from *in vitro* experiments.

It was concluded that the inability of B protein expressed from the single copy plasmid pEC447 to activate p_V transcription in the lysogenic state was likely to be due to

the presence of the CI protein. The CI protein could either be involved in repressing *B* gene transcription on plasmid pEC447 or directly interfere with activation of p_V transcription.

7.2.4 Repression of *B* gene transciption by CI protein.

To determine if CI protein was involved in repressing transcription from the *B* promoter, the level of transcription from the *B* promoter on the single copy *B* plasmid pEC447 was determined in the presence of CI protein supplied from pEC453. Primer extension analysis was chosen for these transcription studies as it allowed the level of transcription specifically from the *B* promoter to be investigated. Since the *B* gene product is involved in repressing its own transcription in the lysogenic state (Kalionis *et al.*, 1986b), the repression of the promoter of the *Bam*17 gene on plasmid pEC448 by CI protein was also determined to investigate the involvement of the *B* gene product in repression of its transcription.

Total RNA was prepared (Sections 2.31.1 and 2.32.2) from cultures of the non-permissive strain E864 carrying the *c*I expression plasmid pEC453, or the parent plasmid pACYC184, and also carrying either the single copy *B* plasmid pEC447 or the single copy *Bam*17 plasmid pEC448. The primer used to identify transcription from the *B* promoter was the *r*-strand of a 101 bp *Hin*PI-*Hin*PI (sequence coordinates 434-334) restriction fragment, prepared by synthesizing a labelled, complementary copy (Section 2.30.1) of the M13 clone mEC648 (Section 2.4.2), and the extent of this primer is shown in Figure 7.3. Primer extension analysis with this primer should give an extension product of 64 bases for transcription initiating from the p_B promoter (Kalionis *et al.*, 1986b).

Primer extension analysis on RNA prepared from a non-induced 186 *cIts* lysogen of the non-permissive strain E864 (Figure 7.4, track 1) or the corresponding 186 *cIts Bam*17 lysogen (Figure 7.4, track 2), gave a predominant extension product sized at 64 bases by comparison with a DNA sequencing ladder, consistent with initiation of transcription from p_B. The marked increase in *Bam*17 gene transcription over that from the wild-type *B* gene in a lysogen, as reflected by the increase in the intensity of the 64 base extension product , was consistent with the autogenous control of *B* gene transcription demonstrated by Kalionis *et al.* (1986b).

Figure 7.3 Nucleotide sequence of the 186 *B* gene.

Nucleotide sequence of the *Pst*I-*Xho*I (65.5% to 67.6%) region, with the upper strand representing the *l*-strand. Predicted coding frames for the *D* and *B* genes are shown. The base changes associated with the amber mutants in the *D* gene (*Dam*23 and *Dam*14) and the *B* gene (*Bam*57 and *Bam*17) are indicated. Also indicated are the ribosome-binding site (RBS) for the *B* gene, the 5' terminus of the *B* transcript (+1) and the associated promoter and termination signals. Inverted repeat structures overlapping the *B* promoter, as identified by Kalionis *et al.* (1986b), are shown by converging arrows. The base pair changes associated with the construction of a *Nde*I site at the initiating methionine of the *B* gene (Osborne, 1986) are indicated. The *Hin*PI sites used to prepare the radioactive primer for primer extension analysis are shown.

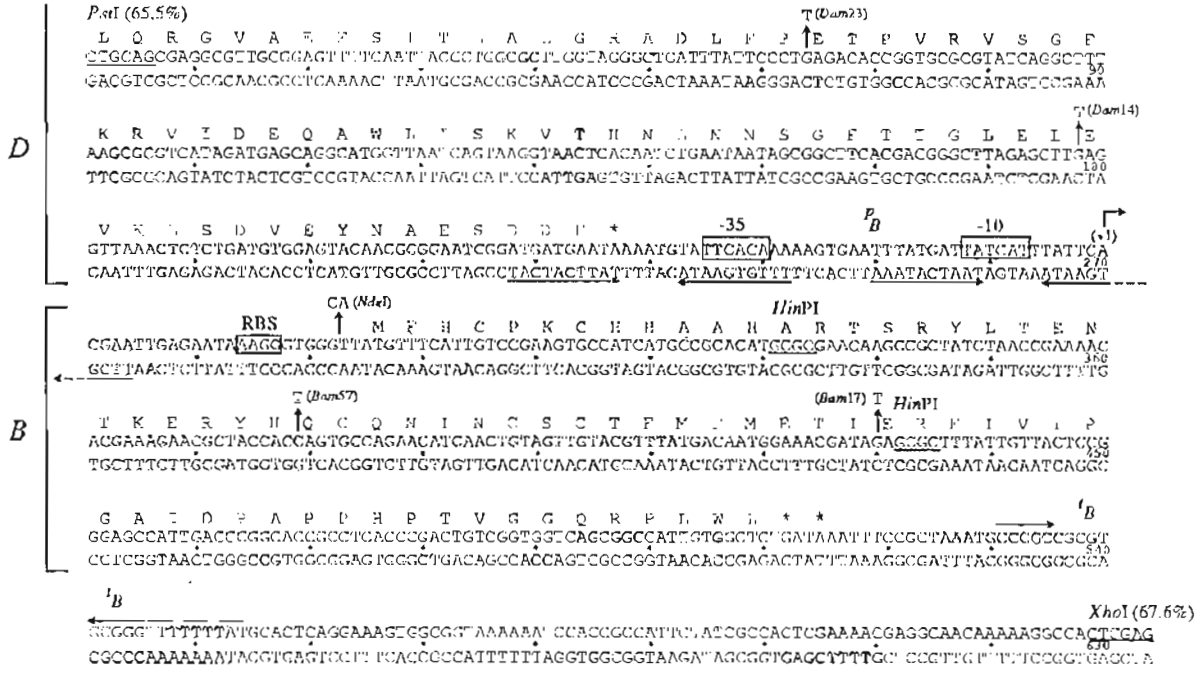

PstI (65.5%)
T (Dam23)
D
-35
-10
(+1)
B
RBS
CA (NdeI)
HinPI
T (Dam57)
(Dam17) T
HinPI
XhoI (67.6%)

Figure 7.4 (cont.)

6. Primer extension of RNA isolated from a culture of the single copy *Bam* plasmid pEC448 and the *c*I expression plasmid pEC453.

7. Identical to track 5, but the culture also carried the p_{V}*galK* plasmid pEC433.

8. Identical to track 6, but the culture also carried the p_{V}*galK* plasmid pEC433.

9. Primer extension of RNA isolated from a 186 *cIts* culture carrying the single copy *B* plasmid pEC447.

10. Identical to track 9, but the culture also carried the p_{V}*galK* plasmid pEC433.

11. Primer extension of RNA isolated 35 min after heat-induction of a 186 *cIts* lysogen (0.5 ug total RNA extended). Used as a positive control to demonstrate that primer was in excess to RNA.

12. Primer extension of RNA isolated from a culture carrying the single copy plasmid pOU61, the *c*I expression plasmid pEC453 and the p_{V}*galK* plasmid pEC433.

13. Primer extended in the absence of RNA.

14. Primer alone.

A sequencing ladder was included to provide size markers. The sequencing ladder was generated from M13 clone mEC648 (a M13mp18 clone of the *Hin*dIII-*Xho*I (61.3% to 67.6%) fragment cloned into the *Hin*dIII and *Sal*I sites; Section 2.4.2), and corresponds as read to the *r*-strand of the sequence in Figure 7.3. As shown on the left of the Figure, the 101 b primer runs at a position corresponding to sequence position 581 (Figure 7.3). The extension product runs at sequence position 517 (Figure 7.3), giving an extension product of 64 bases. A minor extension product sized at 135 bases is also observed. The size of this extension product indicates that the 5'-end of the transcript is located at sequence coordinate 199 (in the *D* gene) from the *Pst*I site at 65.5%.

Figure 7.4 Investigation of repression of the *B* promoter by primer extension analysis on *in vivo* RNA.

RNA was prepared from non-induced cultures of 186 *cIts* and 186 *cIts Bam*17 lysogens of the non-permissive strain E864, and 35 minutes after heat-induction of a culture of a 186 *cIts* lysogen of strain E864 as described in Sections 2.32.1 and 2.32.2. RNA was also prepared from cultures of E864 grown at 30°C carrying (i) the single copy *B* plasmid pEC447 and pACYC184 (ii) the single copy *Bam* plasmid pEC448 and pACYC184 (iii) the single copy *B* plasmid pEC447 and the *c*I plasmid pEC453 (iv) the single copy *Bam* plasmid pEC448 and the *c*I plasmid pEC453 (v) the single copy *B* plasmid pEC447, the *c*I plasmid pEC453 and the p_V*galK* plasmid pEC433 (vi) the single copy *Bam* plasmid pEC448, the *c*I plasmid pEC453 and the p_V*galK* plasmid pEC433 (vii) the parent plasmid pOU61, the *c*I plasmid pEC453 and the p_V*galK* plasmid pEC433. Cultures of a 186 *cIts* lysogen of strain E864 carrying the single copy *B* plasmid pEC447, and the same lysogen carrying the single copy *B* plasmid pEC447 and the p_V*galK* plasmid pEC433, were also used to prepare RNA.

The extension reaction was based on the method of McKnight *et al.* (1981) and is described in Section 2.33. The 101 base *Hin*PI-*Hin*PI (sequence coordinates 434 to 334) primer (Figure 7.3) specific for rightward message was constructed by extension (Section 2.30.2) on the *l*-strand M13 clone mEC648 (Section 2.4.2), followed by restriction with *Hin*PI and fractionation on a 5% non-denaturing polyacrylamide gel (Section 2.26.2(a)). This primer (2.0×10^5 cpm per reaction) was denatured, annealed to *in vivo* RNA (5 ug, unless otherwise stated) and extended with AMV reverse transcriptase in the presence of all four dNTPs as described in Section 2.33. Electrophoresis was as described in Section 2.26.2(b). The autoradiograph was exposed at room temperature for 14 hours.

Gel Track

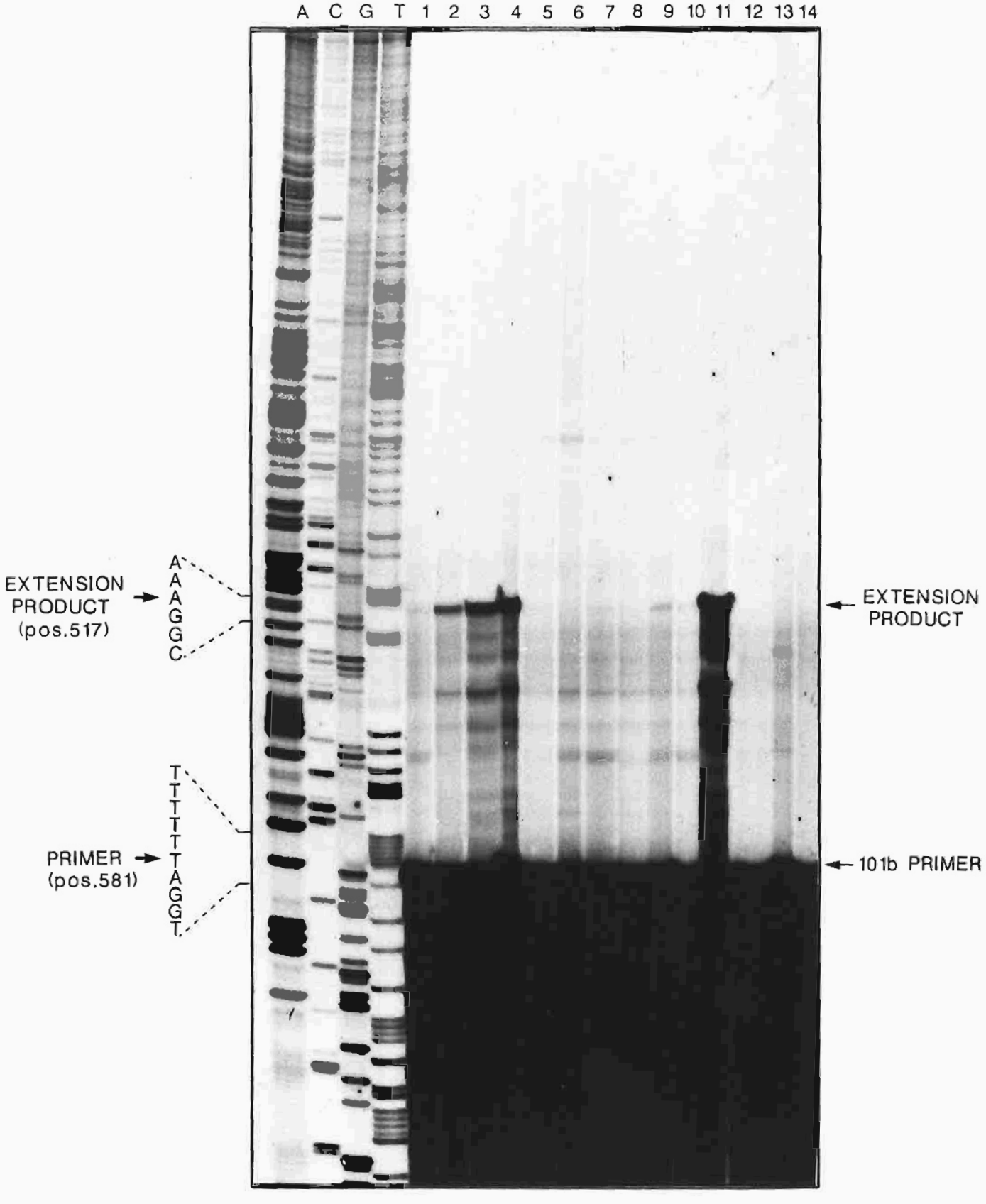
1. Primer extension of RNA isolated from a non-induced culture of a 186 *cIts* lysogen.

2. Primer extension of RNA isolated from a non-induced culture of a 186 *cIts Bam*17 lysogen.

3. Primer extension of RNA isolated from a culture of the single copy *B* plasmid pEC447 and pACYC184.

4. Primer extension of RNA isolated from a culture of the single copy *Bam* plasmid pEC448 and pACYC184.

5. Primer extension of RNA isolated from a culture of the single copy *B* plasmid pEC447 and the *c*I expression plasmid pEC453.


A C G T 1 2 3 4 5 6 7 8 9 10 11 12 13 14
EXTENSION PRODUCT (pos.517)
A A A G G C
PRIMER (pos.581)
T T T T T T A G G T
EXTENSION PRODUCT
101b PRIMER

An extension product of the same size was observed on extension on RNA from the single copy *B* and *Bam* plasmid clones pEC447 and pEC448 in the presence of plasmid pACYC184 (Figure 7.4, tracks 3 and 4 respectively). Transcription from the *B* promoter on the *Bam*17 plasmid pEC448 appeared to be greater than that from the *B* promoter on the *B* plasmid pEC447, but this was probably a product of track distortion, and demonstrated the absence of direct repression of the *B* promoter by B protein alone. The level of transcription from both plasmids was also increased over the level of transcription detected from a 186 *cIts Bam*17 lysogen (Figure 7.4, track 2), which was consistent with the results of the slot blot analysis of transcription from the same plasmids (Section 7.2.2). This suggested that the increased level of *B* gene transcription from plasmids pEC447 and pEC448 (with repect to lysogenic transcription) shown by slot blot analysis (Section 7.2.2) was likely to be a result of increased transcription from the *B* promoter.

The intensity of the 64 base extension product from the single copy *B* plasmid pEC447 was markedly decreased when CI protein was supplied from pEC453 (Figure 7.4, track 5) and demonstrated that the control of *B* gene transcription was in fact indirect, with the CI protein having a role in the repression of transcription from the *B* promoter. Unexpectedly, the intensity of the 64 base extension product was also markedly decreased from the single copy *Bam*17 plasmid pEC448 (Figure 7.4, track 6) and superficially indicated that CI protein was able to repress transcription from p_B in the absence of B protein. This reduction in transcription from the *B* promoter of the *Bam*17 gene on plasmid pEC448 in the non-suppressing host was unexpected, since the difference in the levels of transcription of the prophage *B* and *Bam* genes (Kalionis *et al.*, 1986b; Figure 7.4, track 1 versus 2) demonstrated that the *B* gene product was involved in controlling its own transcription.

One explanation to accomodate the contradiction of track 2 (*Bam*17 in the lysogenic state) and track 6 (*Bam*17 on pEC448 and CI supplied from pEC453) is that the prematurely terminated protein of 45 amino acids produced with the *Bam*17 product has little repressor activity when associated with a limited amount of CI protein in a lysogen (track 2), but is activated to repress at the elevated amounts of CI protein present with the cloned *c*I gene (track 6). Indeed, in this regard, it is noted that the level of *Bam*17 transcription detected from the single copy *Bam* plasmid pEC448 in the absence of CI

protein was increased over that detected from a *Bam*17 lysogen (Figure 7.4, track 4 versus track 2). This difference in transcription levels could be a result of the fact that transcription of the promoter of the *Bam*17 gene in a lysogen is not fully derepressed, due to the residual activity of the *Bam* gene product and the presence of CI protein. If this were the case, supplying CI protein from plasmid pEC453 to the *Bam*17 gene of a lysogen should result in a further reduction in transcription. Furthermore, it would be important to repeat these experiments with the *Bam*57 lysogen (terminating at 28 amino acids; Kalionis *et al.*, 1986b) or preferably with an in-frame deletion mutant. While such experiments remain to be performed, the present indications remain strong that the B and CI proteins in some manner cooperate to impose repression on *B* gene transcription, since the *B* gene does not appear to control CI expression. This cooperative interaction is drawn from the fact that the autogenous control of *B* gene transcription indicates the involvement of the *B* gene product. However, the *B* gene does not appear to control *c*I expression as demonstrated by the fact that immunity to superinfection is conferred by a 186 *cIts Bam*17 lysogen in a non-permissive host (Hocking and Egan, 1982a), and transcription of the *c*I gene has been shown to be not dependent upon the *B* gene product by Northern analysis Kalionis *et al.* (1986b), which therefore indicates some form of cooperative interaction between the proteins.

It was also of interest to determine if the presence of the p_V*galK* plasmid had any effect on the repression of *B* gene transcription. The presence of the p_V*galK* plasmid pEC433 did not appear to have any significant effect on the repression by the CI protein of the *B* promoter on plasmids pEC447 and pEC448 (Figure 7.4, tracks 7 and 8), as transcription from these plasmids remained repressed in the presence of the p_V*galK* plasmid pEC433. The presence of the p_V*galK* plasmid pEC433 also appeared to have no effect on the extent of repression of the *B* promoter on the single copy *B* plasmid in a 186 *cIts* lysogen carrying the single copy *B* plasmid pEC447 and the p_V*galK* plasmid pEC433 (Figure 7.4, track 10), as compared to the same lysogen wihout the p_V*galK* plasmid (Figure 7.4, track 9).

Several minor extension products of a smaller size than the 64 base product were observed (Figure 7.4). Several of these extension products were observed in the control reaction of primer extended in the absence of RNA (Figure 7.4, track 13) and indicated

that they were likely to be due to self-priming reactions. Other extension products of a smaller size than 64 bases observed may have been due to prematurely-terminated extension products or extension on degradation products. However, a minor extension product sized at 135 bases was also produced, which was specific to the *B* or *Bam*17 genes as it was not observed on extension on RNA isolated from a culture carrying the parent plasmid pOU61 (Figure 7.4, track 12). This extension product has also been observed by Kalionis (1985) on extension on lysogenic RNA with a *B* specific primer. The significance and origin of this transcript is unknown, but the size of the extension product demonstrated that the 5'-end of this transcript was located in the *D* gene at sequence coordinate 199 (Figure 7.3).

It is important to note that the studies on *B* gene transcription by primer extension analysis were performed on accumulated transcripts. The possibility does exist that the control of *B* gene transcription is at the level of message stability, rather than at the level of transcription initiation. If this were the case, the CI protein would appear to destabilize the *B* gene message. Pulse-labelling studies would resolve this point. Nevertheless, the *c*I product was demonstrated to decrease transcription initiated from the *B* promoter and regardless of whether the control of *B* gene transcription is at the level of initiation of transcription or message stability, CI protein would appear to be still involved in controlling the level of transcription of the *B* gene present at some level. However, since the *c*I gene product is involved in repression of transcription from the early lytic promoter p_R (Dodd *et al.*, 1990), it was considered likely that it would also have a similar role at the *B* promoter.

7.2.5 CI protein does not act to antagonize *B*-dependent activation of late transcription.

In Section 7.3.2(b) it was concluded that the ability of the lysogenic state to prevent the single copy *B* plasmid pEC447 activating p_V transcription was due to the CI protein blocking p_B transcription or *B* function itself. In Section 7.2.4 it was shown that indeed CI does act to repress p_B transcription and it remained to be confirmed that CI does not also act at the level of *B* function.

To investigate whether B protein was still competent to activate p_V transcription in the presence of CI protein, the ability of the B protein from the *B* expression plasmid

pEC434 to activate p_V transcription on the $p_V galK$ plasmid pEC433 was tested when CI protein was expressed from plasmid pEC453. In this case, CI protein presumably cannot affect *B* transcription by repression, since the *B* gene on plasmid pEC434 is transcribed from the *trc* promoter (Section 5.2.1).

B protein expressed from the *B* expression plasmid pEC434 has previously been shown to activate p_V transcription on the $p_V galK$ plasmid pEC433 (Section 5.2.4). Transformation of strain E864, carrying the *B* expression plasmid pEC434 and the $p_V galK$ plasmid pEC433, with the *cI* expression plasmid pEC453 did not result in any inhibition of the ability of B protein to activate p_V transcription on the $p_V galK$ plasmid (117 units; Table 7.5), as compared with the galactokinase activity from the $p_V galK$ plasmid on transformation with the parent plasmid pACYC184, which gave 88.1 units (Table 7.5). The CI protein from plasmid pEC453 was active in preventing activation of late transcription from the late promoter p_V, when the source of B was under p_B control (pEC447, Table 7.5). Therefore, the ability of B protein to activate late transcription as efficiently in the presence of CI protein as it did in the absence of CI protein, demonstrated that the CI protein did not act to prevent B protein activating late transcription *per se*. This also indicated that the CI protein probably had no effect in destabilizing *B* mRNA (as *B*-dependent activation still occurred), and it was concluded that the inability of B protein to activate late transcription in the presence of the CI protein was not a result of direct interference, but a consequence of repression of *B* gene transcription.

7.3 SUMMARY.

The aim of this chapter was to investigate whether the inability of a prophage copy of the *B* gene to activate late transcription in the lysogenic state was a direct consequence of a low B protein concentration due to autogenous repressive control of *B* gene transcription or due to the expression of a lysogenic function which directly antagonized the ability of B protein to activate late transcription, or a combination of these possibilities.

No evidence of direct repression of *B* gene transcription by B protein was observed in the work presented and indicated that the autogenous control of *B* gene transcription was indirect. It was found that the *B* gene cloned onto a single copy plasmid under its own control was able to activate p_V transcription, but unable to activate p_V transcription when

TABLE 7.5

B-dependent activation of p_V transcription in the presence of CI protein.

Culture[a]	Galactokinase activity[b]
p_V*galK* pOU61B^+ pACYC184	40.1±4.0
p_V*galK* pOU61B^+ p*cI*	1.2±0.4
p_V*galK* p$p_{trc}B^+$ pACYC184	88.1±8.3
p_V*galK* p$p_{trc}B^+$ p*cI*	117 ±10

Notes to Table 7.6

a. Cultures of the non-permissive strain E864 carrying the p_V*galK* plasmid (pEC433), the single copy *B* plasmid pOU61B^+ (pEC447) and either plasmid pACYC184 or the CI expressing plasmid p*cI* (pEC453) were grown at 30°C in M63 supplemented media with the appropriate antibiotics to A_{650}=0.3-0.6, samples taken and the galactokinase activity determined. Cultures of strain E864 carrying the p_V*galK* plasmid (pEC433), the *B* plasmid expression plasmid p$p_{trc}B^+$ (pEC434) and either plasmid pACYC184 or the *cI* expressing plasmid p*cI* (pEC453) were grown at 30°C in M63 supplemented media with the appropriate anibiotics to A_{650}=0.3-0.6, samples taken and the galactokinase activity determined. The p_V*galK* plasmid (pEC433), the single copy *B* plasmid pOU61B^+ (pEC447), the plasmid pACYC184, the 1.3 kb *Hae*III-*Hae*III (71.2% to 75.5%) plasmid clone p*cI*(pEC453), and the *B* expression plasmid p$p_{trc}B^+$ (pEC434) are described in detail in Section 2.3.

b. The galactokinase activity was determined as described in Section 2.18, and are the average of two separate experiments.

introduced into a 186 lysogen. The inability of this single copy plasmid copy of the *B* gene to activate p_V transcription in a 186 lysogen was concluded to be due to the expression of the CI protein, since a plasmid clone containing only the *c*I gene was able to prevent the single copy plasmid clone of the *B* gene activating p_V transcription. This demonstrated that the inability of B protein to activate late transcription in the lysogen was either due to the *c*I gene product repressing *B* gene transcription or that the *c*I gene product acted to directly inhibit B protein activating late transcription.

Using primer extension analysis to detect transcription from the *B* promoter, it was found that the level of transcription from p_B on the single copy *B* plasmid was markedly decreased in the presence of CI protein, consistent with the CI protein being involved in repression of *B* gene transcription. Thus the control of *B* gene transcription was concluded to be indirect and to involve the *c*I gene product. Therefore, the CI protein which is normally involved in maintainence of the lysogenic state by repression of the early lytic promoter p_R (Dodd *et al.*, 1990), appeared to have a second role in being involved in the repression of *B* gene transcription in the lysogenic state.

The CI protein was also found to repress transcription of the *Bam*17 gene on the single copy plasmid and this superficially indicated that CI was able to repress transcription from p_B in the absence of *B* gene product. This was contrary to the autogenous control of *B* gene transcription postulated from the difference in transcript levels of the *B* and *Bam*17 genes in a lysogen (Kalionis *et al.*, 1986b), which indicated that the *B* gene product had some role in repressing its own transcription. Since B protein does not appear to control CI expression from immunity and *in vivo* transcription studies (Hocking and Egan, 1982a; Kalionis *et al.*, 1986b), it appeared that the autogenous control of *B* gene transcription in the lysogenic state was likely to be due to B protein modifying the CI protein to alter its specifity. Therefore the ability of the CI protein expressed from a plasmid to repress transcription of the *Bam*17 gene on a single copy plasmid may have been due to the *Bam* gene product having some residual activity enabling it to modify CI activity.

B protein expressed from the *B* expression plasmid pEC434 was able to activate p_V transcription in the presence of a plasmid clone expressing the *c*I gene and demonstrated that CI did not directly antagonize *B* gene function. Therefore the inability of the *B* gene

to activate late transcription when transcription was repressed was concluded to be a direct consequence of repression, which presumably resulted in insufficient expression of B protein for activation of transcription.

CHAPTER 8

CONCLUDING DISCUSSION

CHAPTER 8. CONCLUDING DISCUSSION.

Finnegan and Egan (1981) first reported that the phage genes *A* and *B* were involved in transcription of the late genes of phage 186. Kalionis *et al.* (1986b) argued that the role of the replication gene *A* was to provide a replicating template for *B* to function and also proposed that another phage function, whose expression was under *c*I repressor control, was also required for late transcription.

In this work it was demonstrated that *B* is the only phage function necessary for late transcription and neither a second function under *c*I control nor a replicating template is required for B to activate a 186 late promoter. In this discussion I will first review the argu ments of Finnegan and Egan (1981) and Kalionis *et al.* (1986b) in light of the present results, then discuss activation of transcription from the late promoters of 186, the control of *B* gene transcription, the inability of P4 to transactivate a repressed 186 prophage, and finally present the current model for the sequence of events that leads to late gene expression after induction of a prophage.

8.1 B IS THE ONLY PHAGE FUNCTION NECESSARY FOR ACTIVATION OF 186 LATE PROMOTERS.

In this work it was demonstrated that the *B* gene product is the only phage function required directly for the activation of a 186 late promoter, since the *B* gene product alone expressed from a plasmid was capable of efficiently activating late transcription in the absence of other phage functions (Section 5.2.4).

The need of an additional function other than B for activation of late transcription, the expression of which was under *c*I control, was based by Kalionis *et al.* (1986b) on the fact that satellite phage P4 could rescue late functions from a derepressed 186 prophage, but not from a repressed 186 prophage (Sauer *et al.*, 1982). The possibility that the single copy nature of the repressed prophage represented inadequate gene dosage was dismissed by Kalionis *et al.* (1986b), as P4 could efficiently rescue late functions from the derepressed prophage of the replication mutant 186 *cIts Aam* (Sauer *et al.*, 1982), which would therefore remain in single copy. Kalionis *et al.* (1986b) proposed that a 186 function under *c*I repressor control was necessary for P4 δ to transactivate late transcription of 186 and, as B and δ are very closely related (Kalionis *et al.*, 1986b),

activation by B must also need this same function. However, since no other function apart from B is required for activation of late transcription (Section 5.2.4), the same is expected to be true for δ. In this regard, preliminary experiments have demonstrated that the δ gene product alone is capable of activating p_V transcription on the p_V*galK* plasmid, since the δ gene product expressed from the plasmid pCH13 (Halling and Calendar, 1990) is capable of activating the p_V*galK* plasmid pEC433 *in vivo* in the absence of other phage functions (R. Tiwari, personal communication). The possible reasons for the inability of P4 to transactivate a repressed 186 prophage are discussed in Section 8.5.

8.2 A REPLICATING TEMPLATE IS NOT NECESSARY FOR B TO ACTIVATE LATE TRANSCRIPTION.

Kalionis *et al.* (1986b) entertained a requirement of a replicating template for B to activate transcription because of the apparent enigma that a superinfecting immunity-insensitive phage, mutant in a late gene (e.g. 186 *vir Dam*), could not rescue by complementation the appropriate late function from a prophage in the marker rescue studies of Hocking and Egan (1982a), in spite of the fact that the superinfecting phage could replicate and presumably provide excess B function needed for activation. The same enigma exists with the fact that 186 late mutants are very poorly complemented by a coinfecting 186 *Aam* phage (Hocking and Egan, 1982c). Kalionis *et al.* (1986b) reasoned that this inability of the prophage function to complement the superinfecting phage was not a consequence of the single gene dosage of the prophage function, since P4 could rescue sufficient late functions from a single copy 186 template for a burst size 40% of maximum (Sauer *et al.*, 1982), but rather that the inability of B protein to activate late transcription must have been related to the template itself. Since late transcription was known to be dependent upon the activity of gene *A* (Finnegan and Egan, 1981), which is *cis*-acting and under *c*I repressor control, Kalionis *et al.* (1986b) concluded that phage replication provided a specific template topology for B to activate late transcription.

However, the ability of excess B function to efficiently activate a late promoter present on a prophage in the absence of expression of the *A* gene (Section 6.2.1(b)), demonstrated that there was no requirement for a specific template topology to be provided by phage replication in order for B to activate late transcription. The inability to rescue

late functions from the prophage in the marker rescue experiments of Hocking and Egan (1982a), could be explained by the combination of the single copy status of the function to be rescued and the fact that the late promoters on the prophage must compete for available B protein with the vast excess of late promoters present on the superinfecting phage, presumably resulting in an insufficient level of the function to be rescued for phage morphogenesis during lytic infection. The lack of complementation with 186 *Aam* phage can be similarly explained. The rescue of late functions from a derepressed 186 *Aam* prophage by P4 superinfection in fact is not comparable as the large excess of competing late promoters is absent in such a case.

In the case of the closely related phage P2, Halling and Calendar (1990) have concluded that the *ogr* gene requires a replicating template in order to activate late transcription with the following reasoning. Starting with the conclusion that a P2 phage unable to replicate cannot transcribe its late genes, they determined that a non-replicating template can express its *ogr* gene and therefore concluded that it must be at the level of Ogr functioning that non-replication blocks late transcription. The data used to arrive at this conclusion, the ability of the replication defective mutant P2 ogr^{+} *Aam*127 to complement the deletion mutant P2 *delogr*17 A^{+}, does indicate the ability of the *ogr* gene to be expressed from a non-replicating template, but indicates nothing about the ability of Ogr to function on a non replicating template. In light of the result that B protein is able to activate the late promoters of a prophage (Section 6.2.1(b)), my prediction is that the role of replication in late gene transcription in the normal vegetative cycle of P2 (Funnell and Inman, 1982; Geisselsoder *et al.*, 1973; Lengyel and Calendar, 1974) will be to provide the elevated levels of Ogr required for activation, and that P2 Ogr will join its close relatives P4 δ (Sauer *et al.*, 1982) and 186 B (Section 6.2.1(b)) in not requiring for its function the topology of a replicating template.

8.3 ACTIVATION OF TRANSCRIPTION FROM THE LATE PROMOTERS.

8.3.1 Role of the late control proteins in activation of late transcription.

The 186 *B* gene product is the only phage function required to directly activate the 186 late promoter p_V (Section 5.2.4), and it seems likely that this will extend to the

remaining 186 late promoters. The 186 B, P2 Ogr and P4 δ proteins show extensive homology in their amino acid sequences, with the δ protein resembling a head to tail dimer of B and Ogr (C. Halling, personal communication). These proteins appear to be functionally interchangeable in their activation of late transcription. This is drawn from the following facts. The P2 late promoters are normally under the control of the P2 *ogr* gene (Sunshine and Sauer, 1975; Christie and Calendar, 1983; 1985), but can also be activated by the 186 *B* gene product in P2-186 hybrid phages (Hocking and Egan, 1982b). The δ gene product of P4 can substitute for the 186 *B* function (Sauer *et al.*, 1982) and the *ogr* function of P2 (Sauer *et al.*, 1982; Halling and Calendar, 1990), and the sites of initiation of transcription from the P2 late promoters activated by P4 δ are the same as those when activated by P2 Ogr (Christie and Calendar, 1983; 1985). The P4 late promoters p_{sid} and p_{LL} are also positively regulated by the δ gene product (Dale *et al.*, 1986; Deho *et al.*, 1988) and the P4 p_{sid} promoter can be activated *in vitro* by the δ and *ogr* gene products in the absence of other phage functions (Keener *et al.*, 1988). Since the 186 *B* gene product was shown to be the only phage function required to directly activate late transcription (Section 5.2.4), the late control proteins P2 Ogr and P4 δ are also expected to be the only phage functions required for activation of late transcription.

The inability of the prophage copy of the *B* gene to activate transcription was concluded to be due to an insufficient concentration of B protein, a direct consequence of the repression of *B* gene transcription (Section 7.2.5). In the lysogenic state there appeared to be a sufficient concentration of B protein for repression (as the *B* gene product was involved in repressing its own transcription; Kalionis *et al.*, 1986b), but an insufficient concentration for the activation of late transcription (Section 4.2.4(a)). It was therefore concluded that in order to activate late transcription, B protein was required at high concentration, at least with respect to the concentration in the lysogen. This need of a high concentration of B protein for activation at the late promoters could be to satisfy a low affinity binding constant for B at the late promoters. Alternatively it could reflect a possible requirement for the need of a multimeric form of B for activation (as discussed in Section 8.3.4), or a combination of a low affinity binding constant and the need of a multimeric form.

The late control proteins of phages 186, P2 and P4 appear to be positive

transcription factors which act at the late promoters of these phages. This raises the question as to whether the late control proteins act as alternative σ factors or accessory factors in activating transcription from the late promoters. The late control proteins of these phages appear to interact with one or both of the α subunits of RNA polymerase to effect late transcription (Section 1.5.5; Sunshine and Sauer, 1975; Halling *et al.*, 1990; Fujicki *et al.*, 1976). However, the late control proteins are unlikely to function as novel σ subunits of RNA polymerase, as there is no detectable homology between the known σ factors of bacteria and their phages (Helmann and Chamberlin, 1988) and the late control proteins of phages 186, P2 and P4 (Christie *et al.*, 1986; Halling *et al.*, 1990). Consistent with the fact that these proteins are unlikely to function as new σ factors, the *in vitro* activation of the P4 late promoter p_{sid} by the P4 *delta* gene product has been shown to be dependent upon the σ^{70} subunit (Keener *et al.*, 1988). Therefore, the *B*, *ogr* and δ gene products are unlikely to replace the vegetative σ subunit in the process of activation, and presumably act as accessory factors in conjunction with RNA polymerase holoenzyme

The examiners note that it would be appropriate to add the reference to Lee T-C. and Christie G.E. (1990). *J. Biol. Chem.*, **265**: 7472-7477, in which the Ogr protein is shown to be a zinc binding protein.

to contain a λ Cro-like DNA-binding domain as determined by the method of Dodd and Egan (1987) (I. Dodd, personal communication). The "zinc-finger" DNA-binding motif, originally identified in *Xenopus laevis* transcription factor TFIIIA (Miller *et al.*, 1985) and identified in several other proteins (Evans and Hollenberg, 1988), comprises a pair of cysteine residues and a pair of histidine residues that form a tetrahedral co-ordination complex with a zinc(II) ion, the intervening residues (12 or 13) creating a peptide loop which interacts with the DNA (Miller *et al.*, 1985). Halling *et al.* (1990) noted that the 186 B, P2 Ogr and P4 δ proteins have pairs of cysteine residues which could possibly act to chelate zinc and have proposed that the late control proteins could possibly be a novel class of zinc finger DNA-binding proteins using the pairs of cystein residues in the proteins to chelate a zinc(II) ion. However, these cysteine residues are located at positions 4 and 7, and 30 and 35 (Figure 1.5) and would yield an intervening loop of 22 amino acids. Obviously if the late control proteins were DNA-binding proteins of this type, they would

represent a novel class of zinc-dependent DNA-binding proteins.

There are a number of bacterial activator proteins which are known or likely to bind to DNA, but which contain DNA-binding motifs of unknown structure (Raibaud and Schwartz, 1984; Busby, 1986). The late control proteins of 186, P2 and P4 may represent such a class of bacterial activators. Alternatively, it is also possible that these proteins may not bind DNA independently of RNA polymerase holoenzyme, but may perform the function of an auxilliary σ subunit.

8.3.2 Host functions involved in activation of late transcription.

8.3.2(a) Role of the bacterial RNA polymerase.

The involvement of the host RNA polymerase in the activation of late transcription was demonstrated by the fact that P2 late gene transcription is blocked by the *E.coli* C *rpoA*109 mutation, which is a structural mutation in the α subunit of RNA polymerase (Sunshine and Sauer, 1975; Fujicki *et al.*, 1976). Dominant mutations of P2 were isolated that suppressed the *rpoA*109 mutation by restoring P2 late transcription, which defined the P2 *ogr* gene (Sunshine and Sauer, 1975; Birkeland and Lindqvist, 1986; Christie *et al.*, 1986), and indicated that the *ogr* gene may interact directly with one or both of the α subunits of RNA polymerase. Keener *et al.* (1988) have also shown that *in vitro*, activation of the P4 *sid* promoter by the δ gene product utilizes the σ^{70} subunit. Thus it would appear that the entire RNA polymerase holoenzyme will be involved in activation of transcription from the late promoters.

The *B* gene of 186 is functionally interchangeable with the P2 *ogr* gene (Hocking and Egan, 1982b) and the two proteins display considerable homology in their amino acid sequences (Kalionis *et al.*, 1986b). Since the mechanisms of action are likely to be similar (Hocking and Egan, 1982b; Kalionis *et al.*, 1986b), the *B* gene product is also expected to interact with the α subunit of RNA polymerase to effect activation of late transcription. However, the function of the 186 *B* gene is unaffected by the *rpoA*109 mutations (Sauer, 1979). Nevertheless, one class of mutants isolated by the lethality of pEC434 to a 186 lysogen (Appendix I) is expected to be in the *rpoA* gene, coding for the α subunit of RNA polymerase. Isolation of 186 mutants able to overcome such mutants should map in the *B*

gene and would confirm the existence of an interaction between the α subunit of RNA polymerase and the B protein.

It has become apparent that the α subunit of RNA polymerase does play a selective role in the initiation of transcription, since mutations in the α subunit have been identified which appear to affect transcription of bacterial genes. A deficiency in the utilization of arabinose, melibiose and glutamate and a concurrent reduction in transport of sulphate is conferred by the *phs* mutation in the *rpoA* gene (Rowland *et al.*, 1985; Giffard *et al.*, 1985). However, the *phs* mutation does not block P2 development (Rowland *et al.*, 1985). Second-site suppressors of the transcriptional defect conferred by the *envZ* mutation, the porin regulatory gene, also appear to map in the *rpoA* gene (Garrett and Silhavy, 1987) and another mutation in the *rpoA* gene interferes with suppression of an *envZ* mutation by a compensatory mutation in the *ompR* gene, the positive regulator of the *ompF* and *ompC* genes (Matsuyama and Mizushima, 1987).

8.3.2(b) Other host functions.

It may be that the host RNA polymerase and the late control proteins of 186, P2 and P4 are the only functions required to activate late transciption. If this were the case, then the only class of mutants that may be identified by the prophage lethality protocol (Appendix I) are those in the α subunit of RNA polymerase (*rpoA*), as was demonstrated for the block to P2 late gene transcription imposed by the *rpoA*109 mutation (Sunshine and Sauer, 1975; Fujicki *et al.*, 1976), since a similar interaction is envisaged for the 186 B protein. If other host functions are involved in the activation process, then these may be exposed by the prophage lethality protocol, discussed in Appendix I.

8.3.3 DNA sequence elements involved in activation of transcription from the late promoters.

8.3.3(a) Upstream sequences.

The 186 late promoters are positively regulated by the *B* gene product (Finnegan and Egan, 1981; Section 5.2.4). The P2 late promoters are normally under the control of the P2 *ogr* gene (Sunshine and Sauer, 1975; Christie and Calendar, 1983; 1985), but can

also be activated by the 186 *B* gene product in P2-186 hybrid phages (Hocking and Egan, 1982b). The δ gene product of P4 can substitute for the 186 *B* function (Sauer *et al.*, 1982) and the *ogr* function of P2 (Sauer *et al.*, 1982; Halling and Calendar, 1990). The P4 late promoters p_{sid} and p_{LL} are positively regulated by the δ gene product (Dale *et al.*, 1986; Deho *et al.*, 1988) and the P4 p_{sid} promoter can be activated *in vitro* by the δ and *ogr* gene products in the absence of other phage functions (Keener *et al.*, 1988). Therefore the B, Ogr and δ proteins appear to be functionally interchangeable and it was not surprising that the DNA sequence of the late promoters of these phage revealed areas of homology, consistent with a common mechanism of activation (Section 3.2.6).

The DNA sequences preceding the transcription initiation sites of the 186 late promoters p_V and p_{12} (Section 3.2.5(b)), the P2 late promoters (Christie and Calendar 1983; 1985) and the P4 late promoters p_{sid} (Dale *et al.*, 1986) and p_{LL} (Deho *et al.*, 1988) are shown in Figure 8.1. All showed very little homology to the *E. coli* promoter consensus sequence in the -10 and -35 regions (Figure 8.1), consistent with the requirement for phage encoded proteins in the activation of transcription. McClure (1985) concluded that positively regulated promters in *E. coli* recognized by RNA polymerase utilizing the σ^{70} subunit have the following features: (i) in the -10 region, at least two out of the three most highly conserved bases TA---T are present and (ii) at least one of the most highly conserved bases of the TTG sequence in the -35 region are present. McClure (1985) also concluded that in positively regulated promoters recognized by RNA polymerase utilizing the σ^{70} subunit, that the positive factor binds at or upstream of the -35 region, but the discovery that the MerR transcription factor binds between the -35 and -10 regions apparently contradicts this conclusion (O'Halloran *et al.*, 1989).

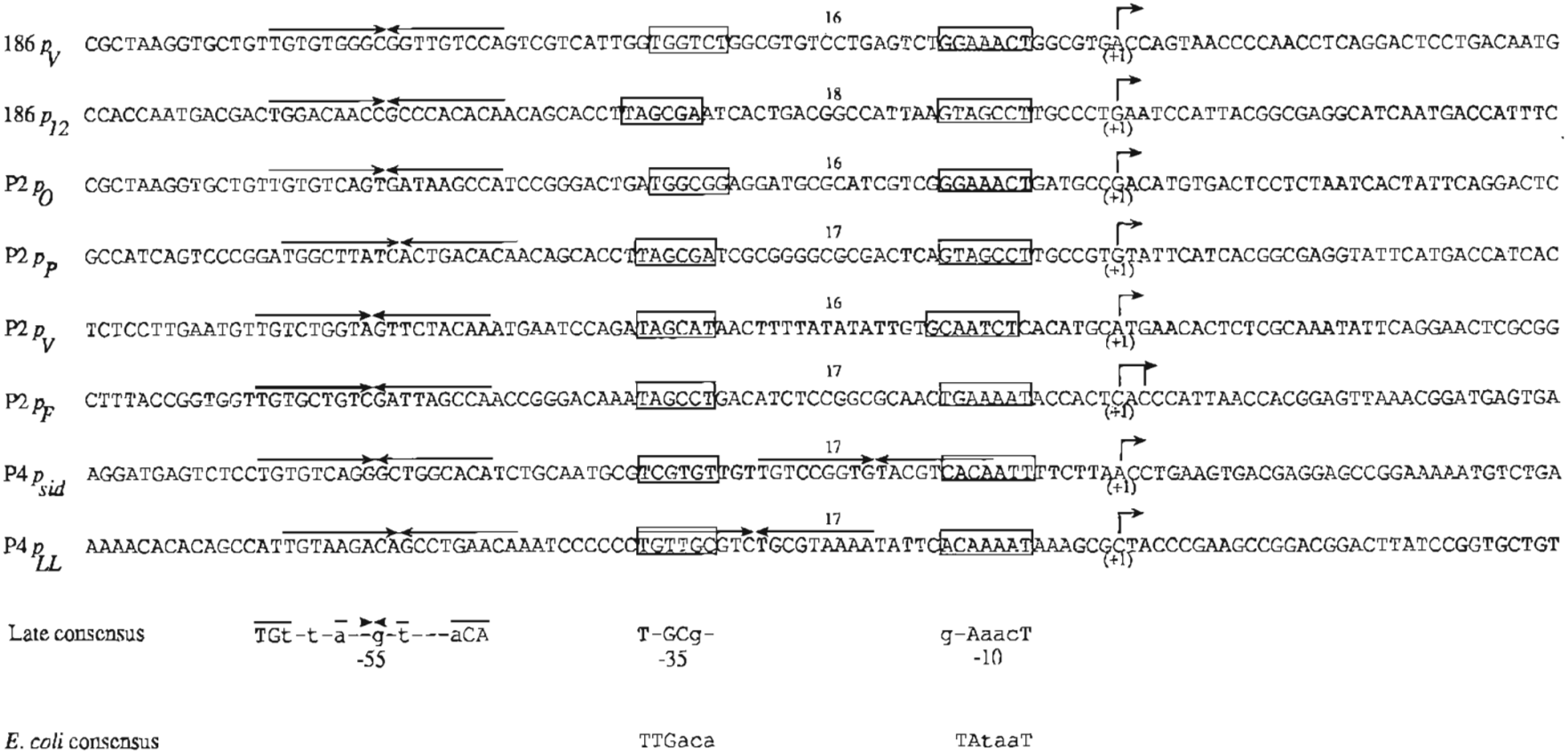
The late promoters of 186, P2 and P4 have the characteristics of late promoters recognized by RNA polymerase utilizing the σ^{70} subunit as described by McClure (1985). The -10 regions of the late promoters of 186, P2 and P4 showed considerable sequence homology and share with the *E. coli* consensus the conserved T-residue at the -7 position and the conserved A residue at the -11 position, as shown in Figure 8.1. Considerable sequence conservation was also noted about the -35 regions of the late promoters, and this conserved sequence (5'-TAGCGT-3') was found to be positioned 16-18 base pairs upstream of the conserved -10 region (Figure 8.1). This conserved sequence in the -35

Figure 8.1. Comparison of the late promoter sequences of 186, P2 and P4.

Nucleotide sequences of the predicted 186 late promoters p_V and p_{12} (Section 3.2.5(b)), the P2 late promoters (Christie and Calendar, 1985) and the P4 late promoters p_{sid} (Dale *et al.*, 1986) and p_{LL} (Deho *et al.*, 1988), aligned at the transcription initiation sites. The transcription initiation sites for the P2 late promoters (Christie and Calendar, 1985), the P4 late promoters (Dale *et al.*, 1986; Deho *et al.*, 1988) and the 186 late promoters p_V and p_{12} (Sections 4.2.1 and 4.2.2) are shown (+1). Boxed regions indicate sequences in the late promoters showing homology and are positioned about the -10 and -35 regions, as originally identified by Christie and Calendar (1985). The number of intervening bases between the conserved sequences at the -10 and -35 positions are shown.

Dale *et al.* (1986) identified an inverted repeat structure centered at the -56 position in the P4 p_{sid} promoter and noted homology with the P2 late promoters in the corresponding region (-55 position). These inverted repeat structures are indicated by converging arrows. Inverted repeat structures are also present in the 186 late promoters p_V and p_{12} (Section 3.2.5(b)). Further copies of this inverted repeat were also noted in this work in the P4 p_{sid} and p_{LL} promoters.

The conserved sequences about the -10, -35 and -55 positions have been incorporated into a late promoter consensus sequence as shown below the promoter sequences. Upper case letters represent nucleotides present in 8/8 or 7/8 of each sequence position. Lower case letters represent nucleotides present in 6/8 or 5/8 of each sequence position. The *E. coli* promoter consensus sequence (Hawley and McClure, 1983) is shown for comparison.


186 p_V CGCTAAGGTGCTGTTGTGTGGGCGGTTGTCCAGTCGTCATTGGTGGTCTGGCGTGTCCTGAGTCTGGAAACTGGCGTGACCAGTAACCCCAACCTCAGGACTCCTGACAATG
16
(+1)
186 p_{12} CCACCAATGACGACTGGACAACCGCCCACACAACAGCACCTTAGCGAATCACTGACGGCCATTAAGTAGCCTTGCCCTGAATCCATTACGGCGAGGCATCAATGACCATTTC
18
(+1)
P2 p_O CGCTAAGGTGCTGTTGTGTCAGTGATAAGCCATCCGGGACTGATGGCGGAGGATGCGCATCGTCGGGAAACTGATGCCGACATGTGACTCCTCTAATCACTATTCAGGACTC
16
(+1)
P2 p_P GCCATCAGTCCCGGATGGCTTATCACTGACACAACAGCACCTTAGCGATCGCGGGGCGCGACTCAGTAGCCTTGCCGTGTATTCATCACGGCGAGGTATTCATGACCATCAC
17
(+1)
P2 p_V TCTCCTTGAATGTTGTCTGGTAGTTCTACAAATGAATCCAGATAGCATAACTTTTATATATTGTGCAATCTCACATGCATGAACACTCTCGCAAATATTCAGGAACTCGCGG
16
(+1)
P2 p_F CTTTACCGGTGGTTGTGCTGTCGATTAGCCAACCGGGACAAATAGCCTGACATCTCCGGCGCAACTGAAAATACCACTCACCCATTAACCACGGAGTTAAACGGATGAGTGA
17
(+1)
P4 p_{sid} AGGATGAGTCTCCTGTGTCAGGGCTGGCACATCTGCAATGCGTCGTGTTGTTGTCCGGTGTACGTCACAATTTTCTTAACCTGAAGTGACGAGGAGCCGGAAAAATGTCTGA
17
(+1)
P4 p_{LL} AAAACACACAGCCATTGTAAGACAGCCTGAACAAATCCCCCCTGTTGCGTCTGCGTAAAATATTCACAAAATAAAGCGCTACCCGAAGCCGGACGGACTTATCCGGTGCTGT
17
(+1)
Late consensus TGt-t-a--g-t---aCA -55 T-GCg- -35 g-AaacT -10
E. coli consensus TTGaca TAtaaT

region possessed two of the highly conserved nucleotides in the -35 TTG sequence of the *E. coli* consensus sequence (McClure, 1985), but deviated significantly from the *E. coli* consensus in the remainder of this conserved sequence.

Dale *et al.* (1986) identified an inverted repeat structure at the -56 position in the P4 p_{sid} promoter and noted elements of this inverted repeat structure in the corresponding region of the P2 late promoters (Figure 8.1). Deho *et al.* (1988) also noted elements of this inverted repeat at the -54 position of the P4 p_{LL} promoter (Figure 8.1). Elements of this inverted repeat were also identified at the -55 position in both the 186 late promoters p_V and p_{12} (Figure 8.1; Section 3.2.6). The inverted repeat elements at the -55 position in the 186 and P2 late promoters and the P4 p_{LL} promoter showed less dyad symmetry than the same element in the P4 *sid* promoter, but some degree of dyad symmetry was still apparent (Figure 8.1). Deletion and mutation analysis have revealed that certain nucleotides in the inverted repeat at the -56 position in the p_{sid} promoter are essential for activation of this promoter by the P4 δ gene product (van Bokkelen *et al.*, 1990), and this is expected to extend to the other late promoters, considering the extensive sequence conservation of this element.

It is interesting to note that the arrangement of the divergent 186 late promoters p_V and p_{12} was such that they shared the same inverted repeat element, positioned at the -55 position with respect to the transcription initiation sites for both promoters (Figure 3.10). This arrangment has also been noted for the corresponding p_Q and p_P promoters of P2 (Halling, 1989) and indicates that these divergent promoters in 186 and P2 utilize the same element in the control of their transcription. As suggested by Halling (1989), the arrangement of the inverted repeat elements in the divergent promoters p_Q and p_P of P2 also implies that the inverted repeat element at position -55 is functional in either orientation.

Finally, further copies of the inverted repeat element were also detected upon inspection of the DNA sequence of the P4 p_{sid} promoter (at position -18) and the P4 p_{LL} promoter (at position -27), which were not noted by the respective authors (Figure 8.1), and it may be that these second copies of this inverted repeat overlapping the -10 and -35 regions of the p_{sid} and p_{LL} promoters have roles in regulation of transcription from these

promoters. The possible role of these additional inverted repeat elements in the regulation of the P4 late promoters is discussed Section 8.3.4.

8.3.3(b) Downstream sequences.

In the activation of transcription from the 186 late promoter p_V there appeared to be a possible requirement of a downstream element (between +18 and +229), as activation of p_V transcription in the *galK* promoter analysis system was significantly decreased for such a deletion late after prophage induction (Section 4.2.5(a)). There is no detectable sequence homology downstream of the late promoters of 186 and P2, apart from the ten nucleotide element referred to as the *B box* (Kalionis *et al.*, 1986b). In three of the four P2 late promoters (p_O, p_P and p_V) this element is positioned approximately 25 nucleotides downstream of the transcription initiation site (Figure 8.2). However, the *B box* sequences identified in the 186 late promoter p_V were considerably further downstream and present in both orientations (Figure 8.2). No *B box* sequences were detectable downstream of the 186 late promoter p_{12} (Section 3.2.7). Two of the *B box* sequences downstream of p_V, which were contained on the 490 bp *Alu*I (11.0% to 12.6%) restriction fragment used to construct the p_V*galK* plasmid, appeared to have no effect on activation of transcription, as derivatives of the p_V promoter with either or both of the *B box* sequences deleted had no effect on the ability of 186 p_V to be activated late after prophage induction (Section 4.2.5). The absence of any effect on activation upon deletion of the *B boxes* also demonstrated that these sequence elements were not responsible for the decrease in activation observed upon deletion of DNA sequences downstream of the +18 position in the promoter p_V (Section 4.2.5(c)). Since the *B box* sequences downstream of p_V apparently had no role in controlling transcription in the *galK* promoter analysis system, this raised the question as to what sequence element(s) downstream of the +18 position was responsible for the reduction in activation observed on deletion of DNA downstream of the +18 position, and also raised the question as to the possible role of the *B box* sequence elements.

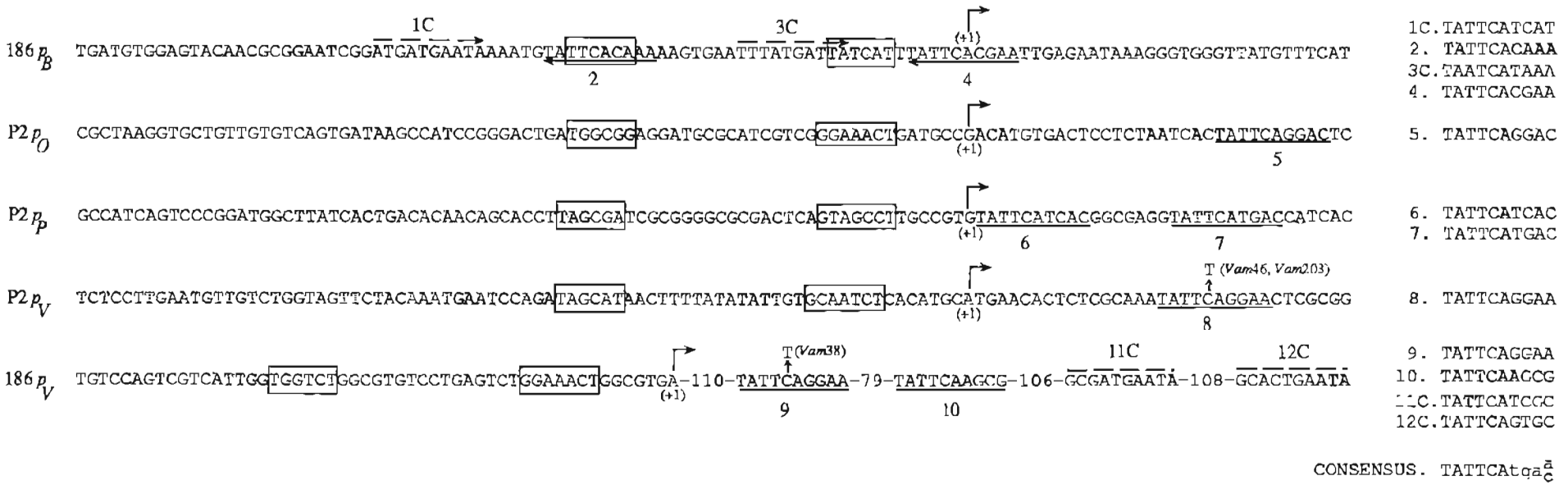
The existence of the *B box* in three of the four late promoters of P2 (Christie and Calendar, 1985) and downstream of the 186 *V* promoter does imply a functional role for this element. The absence of this conserved sequence downstream of the 186 late promoter p_{12} (Section 3.4.2) and the P2 late promoter for the *F* gene (Christie and

Figure 8.2 The *B box*.

Promoter of the 186 *B* gene (Kalionis *et al.*, 1986b) and the late promoters of the P2 genes *O*, *P* and *V* (Christie and Calendar, 1985) presented aligned at the proposed mRNA start-point of each gene (+1). The -10 and -35 regions of the 186 *B* promoter (Kalionis *et al.*, 1986b) and the P2 late promoters (Christie and Calendar, 1985) are boxed. The inverted repeat structures overlapping the *B* promoter are indicated by converging arrows (Kalionis *et al.*, 1986b). One arm of these inverted repeats has been found downstream of three of the four P2 late promoters (underlined regions 5-8), and this sequence has been designated the *B box* (Kalionis *et al.*, 1986b). The 186 late promoter p_V is presented with the *B box* sequences identified downstream of the predicted transcription initiation site (regions 9, 10, 11C and 12C; Section 3.2.7). Intervening nucleotides are indicated.

B box sequences found in the left strand of the sequence are indicated by solid lines (regions 2,4, 5-10) and sequences on the complementary strand are indicated by dashed lines (regions 1C, 3C, 11C and 12C). Base pair changes associated with the amber mutation in the 186 *V* gene (*Vam*38; Dibbens, 1984) and amber muations in the P2 *V* gene (*Vam*46, 203; Christie and Calendar, 1985) are shown.

A consensus sequence for the *B box* sequences identified is presented to the right of the Figure. Upper case letters indicate nucleotides present in 12/12 or 11/12 of the sequence postions. Lower case letters indicate sequences present in 5 or more of the sequences.


1C
3C
(+1)
186 p_B TGATGTGGAGTACAACGCGGAATCGGATGATGAATAAAATGTATTCACAAAAAGTGAATTTATGATTATCATTTATTCACGAATTGAGAATAAAGGGTGGGTTATGTTTCAT
2
4
1C.TATTCATCAT
2. TATTCACAAA
3C.TAATCATAAA
4. TATTCACGAA
P2 p_O CGCTAAGGTGCTGTTGTGTCAGTGATAAGCCATCCGGGACTGATGGCGGAGGATGCGCATCGTCGGGAAACTGATGCCGACATGTGACTCCTCTAATCACTATTCAGGACTC
(+1)
5
5. TATTCAGGAC
P2 p_P GCCATCAGTCCCGGATGGCTTATCACTGACACAACAGCACCTTAGCGATCGCGGGGCGCGACTCAGTAGCCTTGCCGTGTATTCATCACGGCGAGGTATTCATGACCATCAC
(+1)
6
7
6. TATTCATCAC
7. TATTCATGAC
T (*Vam*46, *Vam*203)
P2 p_V TCTCCTTGAATGTTGTCTGGTAGTTCTACAAATGAATCCAGATAGCATAACTTTTATATATTGTGCAATCTCACATGCATGAACACTCTCGCAAATATTCAGGAACTCGCGG
(+1)
8
8. TATTCAGGAA
T(*Vam*38)
11C
12C
9. TATTCAGGAA
186 p_V TGTCCAGTCGTCATTGGTGGTCTGGCGTGTCCTGAGTCTGGAAACTGGCGTGA-110-TATTCAGGAA-79-TATTCAAGCG-106-GCGATGAATA-108-GCACTGAATA
(+1)
9
10
10. TATTCAACCG
11C.TATTCATCGC
12C.TATTCAGTGC
CONSENSUS. TATTCAtgaa$^{a}_{c}$

Calendar, 1985) suggests that the *B box* may not be a DNA sequence motif essential for activation of late transcription. Consistent with this is the fact that the 186 *Vam*38 mutation (a C to T transition in the *l*-strand at position 1024; Dibbens, 1984), maps in the first *B box* identified downstream of the *V* promoter, altering the invariant C residue at position 5 of the first *B box* (TATTCAGAA; Figure 8.2). Since the *Vam*38 mutation is suppressed by a *supE* strain (Hocking and Egan, 1982c), this indicated that the *Vam*38 mutation does not abolish transcription from p_V. Similarly, the P2 *Vam*46 and *Vam*203 mutations, which change the same invariant C residue in the *B box* sequence located downstream of the P2 *V* promoter (Figure 8.2; Christie and Calendar, 1985), are also suppressible.

The inverted repeat sequences overlapping the *B* promoter are candidates for the operator sites for the repressor of *B* gene transcription (Kalionis *et al.*, 1986b). One arm of these inverted repeats corresponds to the *B box* sequences found downstream of the 186 late promoter p_V and the P2 late promoters p_V, p_P and p_O (Christie and Calendar, 1985). Therefore, in the lysogenic state it may be that the repressor of *B* gene transcription is also bound at the *B box* sequences downstream of the late promoter p_V and could have some role in preventing B protein activating late transcription. However, deletion derivatives of the p_V promoter with either or both of the *B box* sequences deleted showed no significant activation in the lysogenic state (Section 4.2.5(c)), and indicated that these elements were not involved in inhibiting activation of transcription in the lysogenic state.

If the *B boxes* are not involved in the process of initiation of transcription from the late promoters, then it is conceivable that the *B box* may have a role in the termination/antitermination of transcription of each of the operons in which it is represented, exerting its influence on transcription distal to the 490 bp *Alu*I restriction fragment used to construct the p_V*galK* plasmid. In this case, the *B box* would act as a signal to modify the transcribing RNA polymerase. Such mechanisms are observed in the lambdoid phages (Herskowitz and Hagen, 1980; Friedman and Gottesman, 1983; Rybchin, 1984; Franklin 1985a,b; Tanaka and Matsushiro, 1985), in *E. coli* rRNA operons (reviewed by Morgan, 1986) and in satellite phage P4, which encodes a function Psu, which antiterminates transcription in P4, P2 and *E. coli* genes by a mechanism which is expected to be unlike λ antitermination (Sunshine *et al.*, 1971; Sauer *et al.*, 1981; Lagos *et*

al., 1986). Northern analysis of transcription initiating from promoter p_V and derivatives with in-frame deletions of the *B boxes* introduced into the phage, may indicate if the *B box* has a role in transcription termination/antitermination, if a change in the transcription pattern is observed.

The *B box* was considered unlikely to have a role in the control of translation, as there appeared to be no positional correlation of the *B box* sequence elements in the P2 late promoters and the 186 *V* promoter, with respect to the predicted ribosome-binding sites or reading-frames of the genes with which they are associated. However, gene fusions, using the *V* gene and derivatives with the *B boxes* deleted in-frame, so as to direct translation of a reporter gene, may expose a role of the *B box* at the translational level.

8.3.4 Proposed mechanism of activation of transcription from the late promoters.

It is likely that the late control proteins B, Ogr and δ will be the only phage functions required for activation of late transcription from the late promoters (Section 8.2.1). The late control proteins are also likely to interact with the α subunit of RNA polymerase (Sunshine and Sauer, 1975; Fjuicki *et al.*, 1976; Halling and Calendar, 1990). In addition, the bacterial RNA polymerase appears to utilize the σ^{70} subunit during activation of late transcription (Keener *et al.*, 1988).

The late control proteins could activate the late promoters by binding to promoter DNA and providing favourable protein-protein interactions that aid promoter recognition or utilization by RNA polymerase (Halling *et al.*, 1990). Alternatively they could function as auxiliary σ factors, interacting with RNA polymerase but not binding to the DNA independently. However, the inverted repeat element at the -55 position in the late promoters is suggestive of a DNA-binding site and it may be that the late control proteins are accessory factors which act by binding to the DNA to improve promoter recognition or utilization, as has been suggested previously by Halling (1989). There is little doubt that the inverted repeat element is essential to activation of the late promoters, since mutations in this element completely abolish activation of the P4 p_{sid} promoter by the δ gene product (van Bokkelen *et al.*, 1990). The lower extent of dyad symmetry in the late promoters of 186 and P2 with respect to the same element at the -56 position of p_{sid}, may be a reflection

of a lower affinity of the late control proteins binding to these sites in the late promoters of 186 and P2.

If the late control proteins do interact with the DNA at the -55 position, then the dyad symmetry of this element would imply either that a dimeric protein binds or a single protein with rotational symmetry of two domains binds to this site. It is tempting to speculate that a dimeric complex of B protein (or Ogr protein) binds to the inverted repeat to activate late transcription, and this could be the reason for the need of B protein at high concentration to activate late transcription. The δ protein resembles a head to tail dimer of the B or Ogr proteins (C. Halling, personal communication) and could therefore bind to the inverted repeat without the need for dimerization.

As noted in Section 8.3.3(a), copies of the inverted repeat elements were also found in positions overlapping the -10 and -35 regions of the P4 p_{sid} promoter and the p_{LL} promoter (Figure 8.1). The positions of these inverted repeats would imply that DNA-binding of the late control proteins could compete for binding by RNA polymerase at the promoter and therefore result in negative regulation of these promoters. The p_{sid} promoter transcribes the δ gene (Dale *et al.*, 1986) and since the δ gene product activates the p_{sid} promoter (Dale *et al.*, 1986; Keener *et al.*, 1988), this indicates that the δ protein positively regulates its own transcription. The inverted repeat element overlapping the -10 region may therefore be involved in negative autoregulation by virtue of the competition between binding of the δ protein and RNA polymerase. However it is important to note that the MerR transcription factor binds between the -35 and -10 regions to positively regulate transcription (O'Halloran *et al.*, 1989) and the same could be true for binding of the late control proteins to the inverted repeat elements overlapping the promoters of p_{sid} and p_{LL}.

Finally, there may be the involvement of a downstream sequence element in the activation of late transcription, as demonstrated by the reduction in activation of the 186 late promoter p_V as result of a downstream deletion of the promoter (Section 4.2.5(a)). The identity of this element and how such a downstream element would influence activation of transcription remains to be determined.

8.3.5 Future studies.

8.3.5(a) Identification of the functions involved in activating late transcription.

One of the aims in studying activation of late transcription in 186 is the identification of the phage and host functions absolutely required to activate late transcription. As demonstrated in this work, B protein is the only phage function required to activate late transcription from the p_V promoter (Section 5.2.4) and it is likely that the other late promoters will also show a similar requirement. The entire bacterial RNA polymerase is also likely to be involved in the activation of late transcription, since the late control proteins Ogr and δ interact with the α subunit of RNA polymerase (Sunshine and Sauer, 1975; Fujicki *et al*, 1976; Halling *et al*, 1990) and activation of the P4 p_{sid} promoter *in vitro* by the δ gene product is dependent upon the σ^{70} subunit of RNA polymerase (Keener *et al*., 1988). Therefore it may be that the B protein and the host RNA polymerase utilizing the σ^{70} subunit are the only functions required to directly activate late transcription. This could be tested by determining the ability of purified B protein and RNA polymerase holoenzyme to activate transcription from a late promoter *in vitro* in the absence of other host and phage functions.

Should other host functions be involved in activation of late transcription, these should be exposed by the protocol for the isolation of survivors resistant to increased *B* expression in a 186 lysogen, as described in Appendix I. Phage mutants able to overcome the block to late transcription imposed by the host mutants isolated should map in the *B* gene, and would demonstrate an interaction between the host functions identified and the B protein. The lethality of increased *B* expression to a prophage also allows the isolation of mutants in the *B* gene on the expression plasmid unable to activate late transcription, since these mutants should be capable of transforming a 186 lysogen. The isolation of missense mutations in the *B* gene by such a procedure may expose regions of the B protein involved in activation.

8.3.5(b) Determination of the DNA sequence elements involved in activation of transcription from 186 late promoter p_V.

Deletion and mutation analysis of the P4 *sid* promoter have provided evidence that the DNA sequence elements at positions -55, -35 and -10 are all important in activation of transcription (van Bokkelen *et al.*, 1990). Considering the extensive homology between the late promoters of 186, P2 and P4, there appears little doubt that the same regions will be important in activation of transcription from the 186 late promoter p_V.

Transcription from the late promoter p_V appears to require the novel situation of an element(s) downstream of position +18 (Section 4.2.4(a)). To determine if there is any requirement for a downstream element, the activation of a series of constructs with sequential 3'-deletion derivatives of p_V, created for instance by *Bal*31 exonuclease (Maniatis *et al.*, 1982), fused to the *galK* gene could be investigated. In these experiments, B protein should be supplied from the *B* expression plasmid pEC434 (Section 5.2.1) or an equivalent expression plasmid, to investigate the effects of the deletions on *B*-dependent activation in the absence of other phage functions. *Bal*31 analysis could also be used in a similar manner to define the upstream elements involved in activation of late promoter p_V.

The most direct approach to determining whether the *B box* sequences have any role in phage development in 186, would be the construction of *B box* deletion derivatives of the late promoter p_V (as described in Section 4.2.5(b)) and introduction of these derivatives into the phage chromosome. *In vivo* transcription studies could then be performed to determine if there was any effect on transcription of the genes encoded on the *VUTSRQ* transcript. Introduction of the p_V promoter with the *B box* deletions into the 186 chromosome could be achieved by homologous recombination *in vivo* or *in vitro* recombination, and expression of genes *VUTSRQ* from a plasmid clone should allow propagation of the mutant phage, if deletion of the *B box* sequences did prove detrimental to phage development. Since B protein is also likely to activate the P2 late promoters directly, the role of the *B boxes* downstream of the P2 late promoters p_Q , p_P and p_V (Christie and Calendar, 1983; 1985) could be also investigated in a similar manner to that used for investigation of the role of the *B box* sequences downstream of the late promoter p_V (Section 4.2.5(c)). Activation of the P2 late promoters by B protein alone is likely, and

the *B*-dependent activation of derivatives of the P2 late promoters with the *B boxes* deleted could be determined.

8.3.5(c) DNA-binding activity of B protein and identification of the presumptive DNA-binding site(s).

To determine if B protein has any capacity to bind at the late promoter p_V, gel retardation analysis (Fried and Crothers, 1981) using crude or purified extracts of B protein can be used to determine if B protein has any capacity to specifically retard a restriction fragment carrying the late promoter. If the B protein does appear to have DNA-binding capacity by this assay, the location of the binding site can be investigated using DNAse protection studies (Galas and Schmitz, 1978) or methylation interference (Siebenlist and Gilbert, 1980).

Obviously such studies would be considerably aided by using purified, active B protein. Gel retardation experiments have been performed using crude extracts of B protein expressed from plasmid pEC434 (Section 5.2.1) and a DNA restriction fragment spanning the late promoter p_V. In these preliminary experiments no retardation of the late promoter p_V was observed specific for the B protein, but it appeared that the absence of retardation was most likely a result of the crude extracts not containing active B protein, since these same extracts were unable to activate galactokinase expression from the p_V*galK* plasmid pEC433 (Section 5.2.3) in a coupled *in vitro* transcription-translation system (R.Tiwari, personal communication). This *in vitro* activation of the p_V promoter is currently being employed as an assay system for purification of active B protein.

Genetic studies may also allow identification of the site on the B protein involved in activation, and identification of the presumptive DNA-binding site(s) in the late promoters. The *galK* gene provides a genetic marker that can be used to positively or negatively select for *galK* gene expression (McKenney *et al.*, 1981). Activation of transcription from the late promoter p_V can be negatively selected against by fusing the *V* promoter to the *galK* gene and selecting for *galK*$^-$ survivors in a *galE*$^-$*T*$^+$*K*$^-$ strain in the presence of galactose (*galK* expression is lethal in a *galE*$^-$*T*$^+$*K*$^-$ strain in the presence of galactose; Nikaido, 1961). Using this negative selection procedure, isolation of missense mutants in the *B* gene unable to activate late transcription (survivors in a *galE*$^-$*T*$^+$*K*$^-$ strain

in the presence of galactose) may expose regions of the B protein involved in activation of transcription or sites involved in interaction with the host RNA polymerase. Such a procedure could also be used to expose the site of interaction on the B protein involved in multimerization, if a multimeric complex of B protein proves to be involved in activation at the late promoter. In a similar manner, mutagenesis of the late promoter p_V and isolation of survivors unable to activate late transcription may expose the sites in the promoter DNA involved in the activation of transcription.

Alternatively, the ability to select for *galK* expression may allow identification of the presumptive DNA-binding site of the B protein in the late promoter, by isolating promoter mutants able to restore activation of late transcription in the presence of the mutant B proteins isolated above. Such mutants can be positively selected for, since they would confer the ability to grow in a $galE^+T^+K^-$ strain on minimal media that contain galactose as the sole carbon source.

8.4 CONTROL OF *B* GENE TRANSCRIPTION.

8.4.1 Negative autoregulation of *B* gene transcription.

Kalionis *et al.* (1986b) concluded that the *B* gene product autogenously repressed its own transcription in the lysogenic state, either directly or indirectly. In this work, the control of *B* gene transcription was shown not to be direct, but to be indirect and concluded to require the involvement of both the *B* gene product and the *c*I gene product for repression of *B* gene transcription (Section 7.2.4). Therefore, the CI protein which is normally involved in maintainence of the lysogenic state by repression of the early lytic promoter p_R (Dodd *et al.*, 1990), appeared to have a second role in being involved in the repression of *B* gene transcription in the lysogenic state.

Since the autogenous control of *B* gene transcription in the lysogenic state was indirect, this meant that either the expression of the *c*I gene was under the control of B protein or that the B protein modified the activity of the CI protein, so that it repressed transcription from the *B* promoter. However, expression of the *c*I gene does not appear to be under *B* gene control (Hocking and Egan, 1982a; Kalionis *et al.*, 1986b) and this suggested that the autogenous control of *B* gene transcription in the lysogenic state was

likely to be the result of the B protein altering the activity of the CI protein so that it would repress *B* gene transcription.

The fact that the CI protein was able to repress transcription from the promoter of a single plasmid copy of the *Bam*17 gene, when CI was supplied from a plasmid clone (Section 7.2.4), superficially indicated that CI protein alone was sufficient to repress the *B* promoter. However, the autogenous control of *B* gene transcription in the lysogenic state implied that B protein had some role in repression, and it was therefore concluded that the *Bam*17 gene product was likely to possess sufficient activity to allow modification of the activity of CI protein at the elevated levels expressed from the plasmid clone, and thereby cause repression of *B* gene transcription (Section 7.2.4). Consistent with partial activity of the *Bam*17 gene product, evidence suggested that the level of transcription from the *Bam*17 gene in a non-permissive 186 lysogen did not reflect complete derepression of the *B* promoter (Section 7.2.4).

The fact that B protein is likely to interact with the CI protein to alter its specificity also suggested that increased B expression could possibly lead to prophage induction, by altering the specificity of the CI protein from that of the primary operator site overlapping the early lytic promoter p_R (Lamont *et al.*, 1988), to that of the operator overlapping the *B* promoter. However it was found that increased *B* expression from a plasmid, although lethal to a prophage, did not cause prophage induction (Section 6.2.1(c)). This fact indicated that either sufficient CI protein with specificity for the primary operator was still present even at high B concentration or that the CI protein retained dual specificity in the presence of B protein.

8.4.2 Mechanism of repression.

In its simplest form, the mechanism of repression at the *B* promoter would be by DNA binding of either an altered form of the CI protein or a B/CI protein complex to some operator site overlapping or in the vicinity of the *B* promoter. Kalionis *et al.* (1986b) identified two inverted repeat structures overlapping the *B* promoter, and suggested that these were likely candidates for the binding sites of the repressor of *B* gene transcription. The ability of a restriction fragment spanning -164 to +21 from the transcription initiation site of the *B* promoter (a *Pst*I-*Nde*I (65.5% to 66.5%) restriction fragment), to induce

expression from the prophage copy of the *B* gene (Section 6.2.3), demonstrated that this restriction fragment was likely to contain operator sites for the repressor of *B* gene transcription.

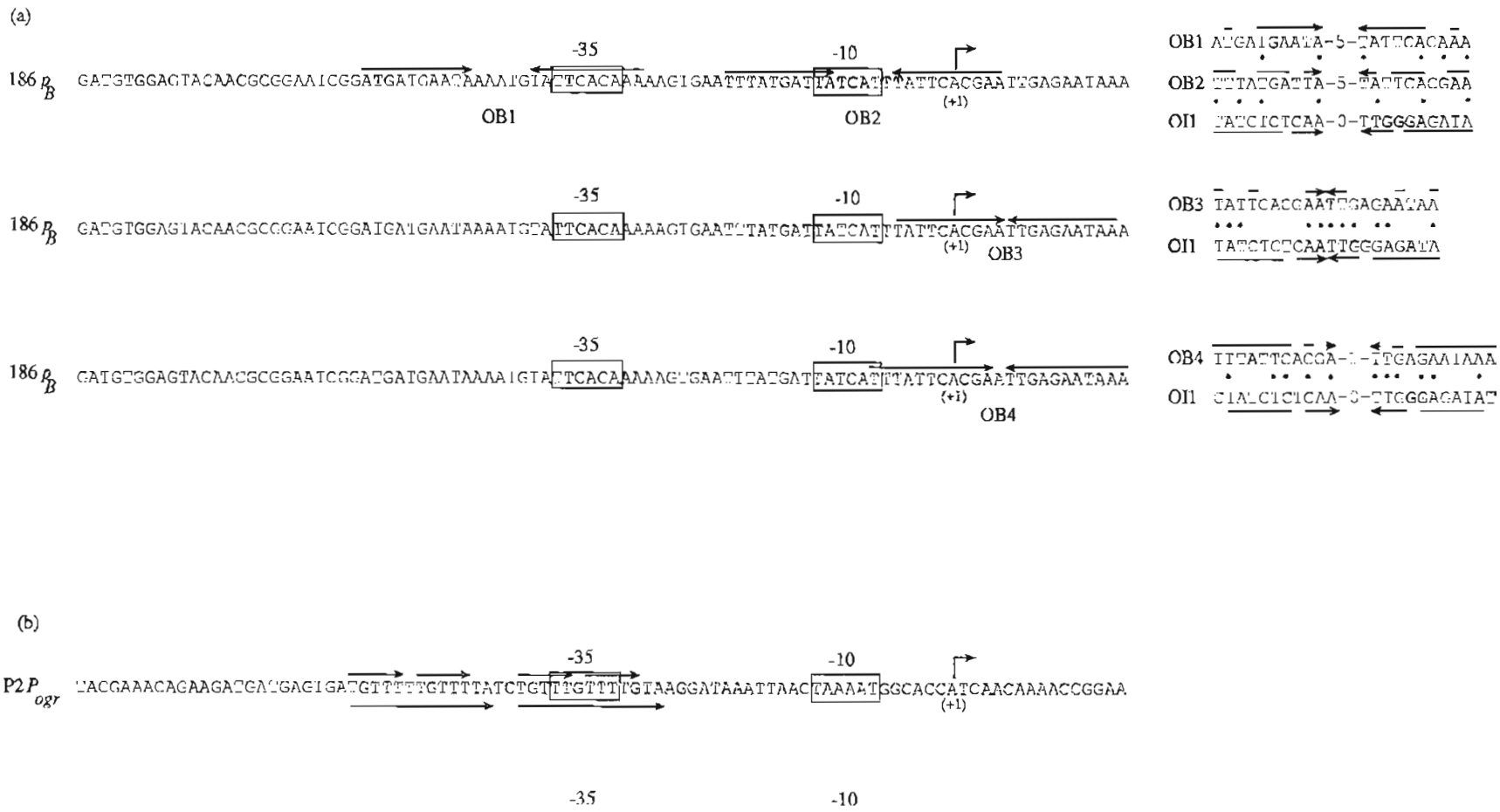
The primary binding site of the CI repressor has been located by sequence analysis of 186 virulent mutants able to form plaques in the presence of CI repressor. These mutants all had mutations in an inverted repeat (OI1) overlapping the early lytic promoter p_R (Lamont *et al.*, 1988) and therefore this inverted repeat is the primary candidate for the CI repressor binding site. The inverted repeats overlapping the *B* promoter (Kalionis *et al.*, 1986b) are the obvious candidates for the operator sites involved in repression of *B* gene transcription. However, these inverted repeat sequences overlapping the *B* promoter (OB1 and OB2, Figure 8.3(a)) do not show significant homology with the primary CI binding site overlapping the p_R promoter. A sequence with limited homology to OI1 and some limited dyad symmetry was found near the transcription initiation site of p_B (OB3; Figure 8.3(a)) and CI protein could conceivably bind at this site to block transcription. A further inverted repeat structure was also identified overlapping the *B* promoter (OB4; Figure 8.3(a)). However, all these inverted repeat structures were sufficiently different from the primary CI binding site OI1 to suggest the need for some change in the DNA-binding specificity of the CI protein.

Nevertheless, the inverted repeats identified by Kalionis *et al.* (1986b) remain the primary candidates for the operator site of the repressor of *B* gene transcription. If these inverted repeats overlapping the *B* promoter do prove to be the operator sites for repression, this would indicate that in order for CI protein to repress transcription, it must acquire a new DNA-binding specificity to recognise the operator site involved in repression of *B* gene transcription, presumably by the B protein altering the specificity in some way. This alteration in the DNA-binding specificity of the CI protein could be achieved by the activation of a DNA binding domain on the protein separate to that used to presumably repress transcription at p_R. In this regard, the λ Int protein has two automonous DNA-binding domains with different recognition specificities (Moitoso de Vargas *et al.*, 1988). Alternatively, the recognition specificity of the presumptive DNA-binding domain of the CI protein could be altered by B protein, in a similar way to that observed for alteration of the DNA-binding specificity of the *S. cerevisiae* α2 repressor by

Figure 8.3 Potential operator sites in the 186 *B* and P2 *ogr* promoters.

a. Potential operator sites in the vicinity of the 186 *B* promoter. The nucleotide sequence, transcription initiation site (+1) and predicted -10 and -35 regions (boxed) of the 186 *B* gene are as determined by Kalionis *et al.* (1986b). Potential operator sites are indicated by converging arrows. Potential operator sites OB1 and OB2 were identified by Kalionis *et al.* (1986b), while sites OB3 and OB4 were identified in this work. Homology of the potential operator sites to the primary CI operator site overlapping the 186 p_R promoter (OI1; Lamont *et al.*, 1988) is shown to the right of the Figure. Homology between the operator sites is indicated and the number of intervening bases between each arm of the repeats also shown.

b. Potential operator sites in the P2 *ogr* promoter. The nucleotide sequence, transcription initiation site (+1) and predicted -10 and -35 regions (boxed) are as first reported in Pritchard and Egan (1985), and by Birkeland and Lindqvist (1986) and Christie *et al.* (1986). Direct repeat sequences identified in this work are shown by repeated arrows.


(a)
186 P_B
-35
-10
(+1)
OB1
OB2
OB3
OB4
OB1
OB2
OI1
OB3
OI1
OB4
OI1
(b)
P2 P_{ogr}
-35
-10
(+1)
-35
-10

a1 protein (Goutte and Johnson, 1988). Determination of whether CI protein requires B protein for DNA binding at the *B* promoter and localization of the operator sites will provide evidence to how CI controls *B* gene transcription.

Finally, Birkeland and Lindqvist (1988) have suggested that the P2 *ogr* gene product may also negatively regulate its own transcription. This came from the fact that transcription from the promoter of an *ogr* deletion mutant (P2 *delogr*17) was reduced upon infection of a cell containing constitutive Ogr protein expressed from a plasmid. No inverted repeats can be detected in the *ogr* promoter (Kalionis *et al.*, 1986b), but direct repeat structures are evident (Figure 8.3(b)). It will be of interest to see what mechanism of control of *ogr* gene transcription is employed, if any, and whether this parallels the control of *B* gene transcription in 186. Recent evidence has in fact indicated that the *ogr* promoter on an infecting, non-replicating P2 phage is repressed by the P2 repressor gene *C* expressed on a plasmid (N. Birkeland, personal communication).

8.4.3 Possible site of interaction of the CI protein with B protein.

The *rpoA*109 mutation in the α subunit of RNA polymerase blocks P2 late transcription (Sunshine and Sauer, 1975; Fujicki *et al.*, 1976) and mutations in P2 that overcome this block map in the *ogr* gene (Sunshine and Sauer, 1975; Birkeland and Lindqvist, 1986; Christie *et al.*, 1986). The extensive amino acid homology between the 186 B protein and the P2 Ogr protein (Kalionis *et al.*, 1986b) and their ability to both activate the P2 late promoters (Hocking and Egan 1982b; Sunshine and Sauer, 1975), indicates that their functional mechanism is likely to be very similar and for these reasons it is anticipated that the B protein will have a corresponding interaction with the α subunit of RNA polymerase. The *rpoA*109 mutation is a structural mutation in the α subunit and the likely region of interaction on the α subunit has been identified, since the amino acid change that resulted in the *rpoA*109 mutation in *E. coli* C has been determined (Fujicki *et al.*, 1976). It was therefore of interest to examine the amino acid sequence of the CI protein for any similarity to the site of interaction on the α subunit of RNA polymerase, as B protein may interact with the CI protein to control its own transcription (Section 7.2.4) and could conceivably use a common protein-protein interaction site.

Interestingly, a region of five amino acids at position 97 of the CI protein showed some homology to the site of interaction on the α subunit of RNA polymerase (Figure 8.4). Of these five amino acids, three residues were identical and the remaining two represented conservative changes. However, whether the B and CI proteins do interact, and if this does represent the site of that interaction, remains to be determined.

8.4.4 Future studies.

8.4.4(a) Autogenous control of *B* gene transcription.

Future studies of the repressive control of *B* gene transcription should be to confirm that CI protein in concert with B protein is involved in repression at p_B and should also concentrate on identification of the operator sites, as this will give an indication as to how repression is manifested at the *B* promoter.

The involvement of the *c*I gene product in repression of *B* gene transcription came from the fact that the *c*I gene was the only gene completely contained on the 1270 bp *Hae*III-*Hae*III (71.2% to 75.5%) restriction fragment, and a function expressed from this region was involved in repression from p_B (Section 7.2.4). Cloning of the *c*I gene and a *c*I$^-$ derivative (away from the 186 p_L promoter) onto an expression plasmid in the absence of other phage genes and investigation of the ability of these constructs to repress *B* gene transcription will demonstrate the absolute involvement of the CI protein.

The *c*I gene product was found to repress transcription of the *Bam*17 gene on a single copy plasmid clone, under conditions of increased expression of the CI protein from a multicopy plasmid clone (Section 7.2.4). This absence of involvement of the *B* gene product may have been a consequence of the high levels of CI protein expressed from the multicopy plasmid clone activating some partial activity of the *Bam*17 gene product. To investigate the involvement of the *B* gene product in repression of its own transcription as observed in the lysogenic state, cloning the *c*I gene under its own control onto the single copy *B* and *Bam*17 plasmid clones (pEC447 and pEC448; Section 7.2.2), to reproduce the single copy status of the *c*I gene in the lysogen, may allow demonstration of the involvement of the *B* gene product in repression at lysogenic levels of CI protein. Alternatively, increasing the CI concentration in a *Bam*17 lysogen, by transformation with

Figure 8.4 Possible site of interaction on the 186 CI protein with the 186 B protein.

(a) The deduced amino acid sequence of the *rpoA*109 mutation in the alpha subunit of *E. coli* C compared to the wild-type (*rpoA*$^+$) is shown (Fujicki *et al.*, 1976). The amino acid sequence of the *rpoA*109 mutation was determined by quantitative amino acid analysis, while the wild-type sequence was determined by Edman degradation. The assignment of the position of the altered residue was not able to be determined (Fujicki *et al.*, 1976).

(b) The amino acid sequence of the wild-type *rpoA* gene of *E. coli* K12, deduced from the DNA sequence (Meek and Hayward, 1984) is shown. Homology of the 186 CI protein at residue 97 of the protein to the amino acid sequence of the *E. coli* K12 *rpoA*$^+$ gene is shown. Identical amino acids are represented by three circles, while amino acid showing conservative homomlogy are indicated by two circles. The amino acid sequence of the 186 CI protein is as determined by translation of the DNA sequence (Kalionis *et al.*, 1986a).

(a)

E. coli C *rpoA*109	thr	glu	val	glu	(his ,	leu)	lys	thr	pro
E. coli C *rpoA* $^{+}$	thr	glu	val	glu	leu	leu	lys	thr	pro

(b)

E. coli K12 *rpoA* $^{+}$	thr	glu	val	glu	leu	leu	lys	thr	pro
				•••	••	•••	•••	••	
186 CI			glu	glu	ile	leu	lys	ser	asp

the CI expression clone, may demonstrate repression of the promoter of the *Bam*17 gene in the lysogenic state to a level comparable with that observed for the wild-type protein and would further confirm the activity of the *Bam*17 gene product in repressing transcription.

To determine if the *B* gene product is involved in repressing its own transcription, single copy plasmid clones carrying the *Bam*57 gene could be constructed, identical to the single copy plasmid clone of the *Bam*17 derivative described in this work. The *Bam*57 mutation is predicted to produce a truncated protein in a non-permissive host of 28 amino acids (Kalionis *et al.*, 1986b) and may therefore be deficient in the residual activity predicted to be associated with the *Bam*17 gene product. Should the *Bam*57 gene also show repression of transcription at elevated CI concentration, then construction of an extensive in-frame deletion of the *B* gene by oligonucleotide mutagenesis to completely remove the *B* gene product and determination of whether CI protein can now repress transcription in the complete absence of *B* gene product will provide an answer as to whether CI protein alone is capable of repression at p_B *in vivo*.

In vitro studies may also provide an insight into the phage functions involved in repression and the identification of the site of repression. Gel retardation analysis (Fried and Crothers, 1981) could be used to determine if CI protein alone binds to the *B* promoter. Crude extracts of CI protein expressed from a plasmid are functional in gel retardation of DNA fragments carrying the primary CI operator (I.Dodd, personal communication). The ability of CI protein to bind *in vitro* to the *B* promoter in such experiments would show the lack of an absolute requirement for B protein in the binding of CI protein to the *B* promoter. If no retardation is observed, the ability of CI protein in conjunction with B protein (either expressed in the same strain, or by addition of B protein expressed from a different source) to bind can be investigated. If mobility shift is observed, either by CI alone or in conjunction with B protein, the operator site can be located by DNAse protection studies (Galas and Schmitz, 1978) or methylation interference (Siebenlist and Gilbert, 1980).

Genetic studies may also allow identification of the operator sites and the presumptive site of interaction between the DNA and the repressor on the repressor molecule itself, using the *galK* gene to positively or negatively select for *galK* gene expression (McKenney *et al.*, 1981). Transcription from p_B can be positively selected for

by fusing the *B* promoter to the *galK* gene (in single copy) and selecting for $galK^+$ colonies in a $galE^+T^+K^-$ strain on minimal media that contain galactose as the sole carbon source. Repression of the *B* promoter in the presence of CI and B proteins should confer a $galK^-$ phenotype and the inability to grow. Mutagenesis of the p_B promoter fused to the *galK* gene and transformation into a $galE^+T^+K^-$ strain carrying sources of B and CI proteins should allow the identification of potential operator site mutants by their ability to grow on minimal media plus galactose (the promoter will still be functional in such a case). Such a procedure could also be used to isolate mutants in the repressor (CI, or B and CI in combination) unable to repress transcription. Further, the ability to select against *galK* expression (*galK* expression is lethal in a $galE^-T^+K^-$ strain in the presence of galactose; Nikaido, 1961) could be used to select for revertants in the operator sites or in the other interacting protein components of the repressor, which are able to re-exert repression. Such a procedure could therefore be used to isolate mutants at the presumptive sites of interaction between the B and CI proteins, if both proteinsare found to be required for repression. This could be achieved with the use of a mutation in one of the proteins which is at the presumptive site of interaction between the two proteins, and the is olation of a suppressing mutation on the other protein component able to re-exert repression

8.4.4(b) Control of late transcription at the level of *B* gene expression.

The initial event in activating late transcription appears to be the derepression of *B* gene transcription due to loss or inactivation of the CI protein. To specifically examine if the loss of CI protein does lead to induction of *B* gene transcription and thereby activate late transcription, the wild-type *c*I gene and the *cIts* derivative may be cloned onto a single copy plasmid carrying the *B* gene. At the permissive temperature (30°C), *B* gene transcription should be repressed and there should be no evidence of activation of p_V transcription on the $p_V galK$ plasmid. At the non-permissive temperature (39°C), the construct carrying the wild-type *c*I gene is expected to show no derepression of *B* gene transcription and no evidence of activation of p_V transcription, while the *cIts* construct should show derepression of *B* gene transcription and activation of p_V transcription. However, the temperature dependent increase in copy number of plasmid pOU61 (Larsen

et al., 1984), used to construct the single copy *B* and *Bam* plasmids (Section 7.2.2), makes it unsuitable as a vector system for this purpose.

There remains the question as to what role increasing gene dosage would have on the repression of *B* gene transcription and therefore the activation of late transcription, if CI protein was still functional and expressed with increasing dosage. In single copy, the *B* and *c*I gene products should repress *B* gene transcription and as a consequence there should be insufficient free B protein to activate late transcription. However, an increase in template number could result in activation of late transcription, either by a consequence of the increase in free B concentration expected as a result of increasing gene dosage, or alternatively an increase in gene dosage could result in derepression of *B* gene transcription due to some change in the state of B protein, and thereby activate late transcription.

To examine if increasing *B* gene dosage does activate late transcription under conditions of expression of CI protein, the wild-type *B* and *c*I genes could be cloned onto the temperature dependent replicon pOU61. At 30°C, *B* gene transcription is expected to be repressed and activation of late transcription absent. Increasing the copy number of the plasmid (by temperature shift) will provide an increased dosage of both the *B* and *c*I genes, although the relative gene dosage will remain unchanged. Under such conditions of increased copy number, the repression of *B* gene transcription and activation of late transcription can be determined. The absence of activation of late transcription at increased gene dosage would demonstrate that sufficient repression is still manifested to prevent an increase in the concentration of B protein for activation of late transcription. However, it is anticipated that the increase in copy number will result in an increase in free B concentration (as discussed in Appendix II) and that activation of late transcription will result as a consequence of this increase in concentration.

8.5 CONCERNING P4 TRANSACTIVATION OF P2 AND 186 PROPHAGE.

Upon infection of a P2 lysogen, P4 derepresses the P2 prophage helper and allows replication of the phage *in situ* (as prophage excision is inefficient in P2) and presumably allows activation of the late genes under their normal mode of control by the P2 Ogr protein (Sauer *et al.*, 1982; Six and Lindqvist, 1978). The ε gene of P4 is responsible for

this derepression of the P2 prophage, as a P4 ε⁻ phage is unable to transactivate a P2 prophage (Geisselsoder *et al.*, 1981). The ε gene product does not appear to effect P4 development (Six and Lindqvist, 1978) and may act as an antirepressor (Bertani and Six, 1988). This inability of P4 to transactivate in the absence of the P4 ε function demonstrates that if derepression of the P2 prophage does not occur, then sufficient transactivation by the δ protein cannot take place to rescue late functions from the P2 prophage, despite the fact that P4 can rescue sufficient late functions from a single copy P2 prophage, as demonstrated by rescue of late functions from a P2 prophage with a mutation in the replication gene *A* (*Aam*127) by wild-type P4 (Halling and Calendar, 1990).

It is an interesting fact that despite the very close relationship of 186 with P2, P4 is unable to transactivate a repressed 186 prophage (Sauer *et al.*, 1982), although P4 can rescue sufficient late functions from a single copy 186 template for a viable infection (Sauer *et al.*, 1982). This raises the question as to why P4 is not competent to plate on a 186 lysogen, and similarly why a P4 ε⁻ phage cannot plate on a P2 lysogen.

The δ gene product is likely to be the only phage function required to activate late transcription (as discussed in Section 8.1), and therefore it could be that the repressed states of P2 and 186 both block δ function by the expression of some function which directly or indirectly blocks δ function or expression. However, the ability of P4 to rescue late functions from a P2 prophage may be due to the ability of the P4 ε function to derepress the P2 prophage and allow expression of the P2 late genes by their normal mode of control, while the inability of P4 to rescue late functions from a 186 prophage may be due to the fact that the prophage is unable to be derepressed and that P4 δ is also blocked in its ability to activate late transcription.

This inability of λP4 to rescue late functions from the prophage is not due to some dependence on the expression of the helper prophage late control gene (Ogr or B), as demonstrated by the fact that wild-type P4 efficiently plates on a P2 *delogr* lysogen (Halling and Calendar, 1990). This dependence of the helper late control gene may have been expected, since the P4 δ gene is transcribed from the late promoter p_{sid} (Dale *et al.*, 1986) and this promoter is activated by Ogr protein (Dale *et al.*, 1986; Keener *et al.*, 1988). Thus, when P4 infects a P2 lysogen, the p_{sid} promoter is expected to be activated by P2 Ogr protein (Dale *et al.*, 1986). Similarly one might expect 186 B to activate the P4 p_{sid}

promoter on P4 infection of a 186 lysogen. However the result with the *delogr* P2 lysogen described above suggests that sufficient δ expression occurs in the absence of this reciprocal activation to allow efficient activation of the prophage late functions by the δ protein.

Since the 186 CI and B proteins are likely to interact (Section 7.2.4), a similar interaction could be envisaged between the 186 CI protein and the P4 δ protein, and it may be that both the 186 CI repressor and the P2 C repressor both directly block P4 δ function or P4 development. Whether 186 CI protein does block P4 δ function can be investigated, by determining if the 186 CI expression clone constructed in this work, blocks P4 δ function in activating late transcription on the p_V*galK* plasmid.

8.6 MODEL FOR THE INDUCTION OF LATE GENE TRANSCRIPTION IN PHAGE 186.

A working model for the induction of late gene expression in 186 is that in the lysogenic state the 186 *B* gene is transcribed and that the protein product of the *B* gene represses its own transcription indirectly by modification of the DNA-binding specificity of the CI protein, resulting in a concentration of B protein insufficient for activation of late transcription. Upon prophage induction, the pool of CI repressor is depleted (due to either inactivation or degradation of the CI protein), allowing derepression of *B* gene transcription. This derepression of transcription provides the increase in concentration of B protein necessary for activation of the late promoters, probably by the conversion of the B protein to a multimeric form.

Finally, the activation of late transcription due to the increase in B concentration, as a result of derepression of *B* gene transcription, should lead to the activation of the late promoter for the transcription unit encoding genes *GFED*, the transcription from which continues into the *B* gene (Section 1.5.3(b); Kalionis *et al.*, 1986b), resulting in the *B* gene product positively controlling its own transcription upon activation of the late promoters. This positive autogenous control of *B* gene transcription, coupled with the increase in *B* gene transcription due to the increase in gene dosage as a result of replication, should provide the large amounts of B protein presumably required for efficient activation of the

numerous late promoters late in lytic infection (Kalionis *et al.*, 1986b). Further, this positive autogenous control also finally decouples *B* gene transcription from any possible repressive effects on transcription from the *B* promoter (Kalionis *et al.*, 1986b).

APPENDIX I

ON THE USE OF *B* EXPRESSION PLASMID pEC434 FOR THE ISOLATION OF HOST MUTANTS BLOCKED IN ACTIVATION OF LATE TRANSCRIPTION

APPENDIX I. ON THE USE OF *B* EXPRESSION PLASMID pEC434 FOR THE ISOLATION OF HOST MUTANTS BLOCKED IN ACTIVATION OF LATE TRANSCRIPTION.

I.1 INTRODUCTION.

The results presented in Section 6.2.1(a) demonstrated that the *B* expression plasmid pEC434 was unable to transform a 186 lysogen. This inability to transform a 186 lysogen was found to be dependent upon functional B protein expressed from plasmid pEC434, due to the presence of the prophage and not a lethal consequence of *B* expression itself (Section 6.2.1(a)). Transformation of a 186 lysogen with pEC434 was achieved by increasing the concentration of *lacI* repressor, with the use of a multicopy $lacI^q$ clone (Section 6.2.1(a)). It was surmised that inadequate control of the *trc* promoter of plasmid pEC434 was responsible for a level of *B* expression, which for an unknown reason, was lethal to the prophage. The inability of a 186 prophage to be stably maintained with the *B* expression plasmid pEC434 in the absence of $lacI^q$ control was shown not to be a result of B protein causing prophage induction (Section 6.2.1(c)) and it appeared that *B* expression from plasmid pEC434 was activating transcription of the prophage late genes with lethal consequences.

The lethality of the *B* expression plasmid pEC434 to a 186 prophage suggested a protocol for the isolation of host mutants specifically blocked in activation of late transcription. Isolation of survivors resistant to derepressed *B* expression from plasmid pEC434 in a 186 lysogen should be defective in the lethal function itself or in *B*-dependent activation of transcription at some level. One class of mutants expected to be isolated by such a procedure, should be mutants in host genes which block *B*-dependent activation of transcription of the lethal prophage function and such mutants should therefore be blocked in activation of 186 late transcription.

This appendix investigates whether the lethality of derepressed *B* expression from pEC434 in a 186 lysogen was a suitable protocol for the isolation of host mutants blocked in activating late gene transcription. The isolation of such host mutants will expose host genes involved in activating late gene transcription and contribute to the understanding of the mechanism by which late gene transcription is controlled.

I.2 THE USE OF *B* EXPRESSION PLASMID pEC434 FOR THE ISOLATION OF HOST MUTANTS BLOCKED IN ACTIVATION OF LATE TRANSCRIPTION.

I.2.1 The effect of induction of *B* expression from pEC434 on the viability of a 186 lysogen.

To demonstrate that induction of *B* expression from the *B* expression plasmid pEC434 was lethal to a 186 lysogen, the viability of 186 lysogenic cells carrying the *B* expression plasmid pEC434 and the *lacI*^{q} plasmid pMC7, was tested in the presence of the gratuitous inducer IPTG. This was performed by determining the cell viability of wild-type 186 lysogens of the non-permissive host E251 (E635), carrying the *B* expression plasmid pEC434 and the *lacI*^{q} clone pMC7, on M13 minimal plates (Section 2.9.2) supplemented with IPTG (1 mM final concentration). The results are presented in Table I.1.

In the presence of IPTG, the cell viability of a 186 lysogen carrying plasmids pEC434 and pMC7 was dramatically reduced, plating with an efficiency of less than 2.2 x 10^{-6} compared with the plating efficiency of the same lysogen in the absence of inducer (Table I.1). As anticipated, this marked reduction in cell viability was shown to be dependent upon the *B* gene product, since the non-permissive lysogen carrying the *Bam* expression plasmid pEC435 and pMC7 showed no significant reduction in cell viability in the presence of IPTG (efficiency of plating 0.87; Table I.1). Induction of *B* expression itself from plasmid pEC434 was not responsible for the extensive cell death, as the corresponding non-lysogen showed no significant reduction in cell viablity on induction with IPTG (efficiency of plating 0.74; Table I.1). It was therefore concluded that induction of *B* expression from pEC434 was lethal to a 186 lysogen.

The nature of the prophage function conferring lethality is unknown. Since increased *B* expression in a 186 lysogen is lethal, this function is presumably under *B* gene control, either directly or indirectly. Since the late genes of 186 are under positive control by the *B* gene product (Finnegan and Egan, 1981; Section 5.2.4), the 186 function conferring lethality on increased *B* expression was expected to be located in the late region. Initially it was considered that activation of transcription of the lysis gene *P* may be responsible for the observed cell death, since considerable cell lysis was observed upon

TABLE I.1

Viability of 186 lysogens on induction of *B* expression from plasmid pEC434.

Culture[a]	Conditions	Viability[b]
p*lacI*^{q} p$p_{trc}B^+$	30°C/1 mM IPTG	0.74
(186$^+$) p*lacI*^{q} p$p_{trc}B^+$	30°C/1 mM IPTG	$<2.2\times10^{-6}$
(186$^+$) p*lacI*^{q} p$p_{trc}Bam$	30°C/1 mM IPTG	0.87

Notes to Table I.1

a. Cultures of the non-permissive strain E251 and the corresponding lysogen E251(186$^+$) (E635), carrying the p*lacI*^{q} plasid pMC7 (Calos, 1978; Section 2.3.1) and either the p$p_{trc}B^+$ plasmid (pEC434) or the p$p_{trc}Bam^+$ plasmid (pEC435), were grown at 30°C in M13 minimal media containing the appropriate antibiotics to A_{600}=0.8. Dilutions of the cultures were plated onto M13 minimal plates containing the appropriate antibiotics supplemented with 1 mM IPTG (final concentration) or without IPTG. Plates were incubated at 30°C for 36 hours and colony forming units scored. Plasmids pEC434 and pEC435 are described in Section 2.3.2.

b. The viability is expressed as the ratio of colony forming units obtained in the presence of 1 mM IPTG to the colony forming units obtained in the absence of IPTG.

IPTG induction in liquid with a 186 lysogen carrying plasmids pEC434 and pMC7. However, the inability of the *B* expression plasmid pEC434 to transform a 186 *cIts Pam*16 lysogen of the non-permissive host E251 (E296) suggested that the lysis gene was not entirely responsible for the lethal effect, presuming that the *Pam* mutation in the non-permissive host did not have residual activity. Experiments to map the B activated phage function lethal to the host are currently being performed in this laboratory. Preliminary studies have shown that the lethal function is located in the late region of 186 (R. Tiwari, personal communication).

I.2.2 Use of plasmid pEC434 to isolate mutants defective in activation of late transcription.

Since extensive cell death was apparent when *B* expression from plasmid pEC434 was induced with IPTG in a 186 lysogen (Section I.2.1), this lethality was investigated as a method to isolate mutants defective in activation of late gene transcription. In order to isolate mutants blocked in *B*-dependent activation of transcription at some level, the protocol chosen was to isolate survivors of a 186 lysogen, carrying plasmids pEC434 and pMC7, which were resistent to treatment with IPTG. The classes of mutants predicted to survive induction of *B* expression in a 186 lysogen were:

1. Plasmid mutants which eliminated *B* gene expression or which eliminated expression of functional B protein from plasmid pEC434. These mutations could be defective in the cloned *B* gene or the plasmid signals controlling *B* gene expression. Plasmid DNA isolated from such surviving colonies should transform a 186 lysogen with high efficiency in the absence of $lacI^q$ control and should not be capable of activating p_V transcription from the $p_V galK$ plasmid pEC433 (Section 5.2.4).
2. Prophage mutants which eliminated the lethal prophage function activated by increased *B* expression. Such mutants were anticipated to be still capable of activating 186 late transcription.
3. Host mutants which blocked the ability of B protein expressed from plasmid pEC434 to activate transcription of the lethal prophage function.
4. Host mutants deficient in the ability to derepress *lacI* controlled transcription with the inducer IPTG.

Of the mutants described above, host mutations were considered most likely to confer resistance to 186 *vir* infection, since activation of transcription of the essential late genes (Hocking and Egan, 1982c) should be absent or significantly reduced in such mutants. Resistance to 186 *vir* infection was therefore chosen as a suitable initial screening step in the identification of host mutants blocked in activation of late transcription. [However, it was also conceivable that a mutation in the *B* gene of plasmid pEC434 could result in a B protein defective in activation, but able to block 186 *vir* infection by a *trans*-dominant effect. A similar argument could also be applied to potential prophage mutants in the presumptive lethal function. Nonetheless both these classes of 186 resistant mutants, if identified, would be of considerable interest in their own right.]

In order to isolate mutants blocked in late activation, the protocol chosen was to isolate survivors by plating a 186 lysogenic culture carrying the *B* expression plasmid pEC434 and the *lacI*^{q} clone pMC7 onto plates supplementedwith IPTG, using the same protocol described in Table I.1. However, no surviving colonies were isolated by such a procedure. In order to enrich the culture for cells resistant to induction with 1 mM IPTG, a culture of the 186 lysogen carrying plasmids pEC434 and pMC7 was induced with IPTG in liquid culture and the cells then allowed to propagate after induction. The liquid induction procedure is described in Table I.2, and essentially involved growing a culture of host E251 lysogenic for wild-type 186 (E635) and carrying plasmids pEC434 and pMC7, at 30^{o}C to late log phase (A_{600} = 0.8) in M13 minimal medium containing ampicillin (25 ug/ml) and tetracycline (15 ug/ml). Induction in liquid was peformed by the addition of IPTG to a final concentration of 1 mM. Considerable cell lysis was observed 2-3 hours post IPTG induction. The culture was grown for 16 hours at 30^{o}C and subsequently subcultured into the same medium supplemented with 1 mM IPTG and grown to late log phase (A_{600}=0.8). Various dilutions of the culture were plated onto M13 minimal plates (containing the appropriate antibiotics) with and without IPTG added to 1 mM final concentration, and survivors scored after growth for 36 hours at 30^{o}C.

Survival on plates supplemented with IPTG, following the above induction procedure, was found to be at a rate of survival of approximately 3.5 x 10^{-3}, with respect to the survival of the same culture plated in the absence of IPTG (Table I.2). This demonstrated that a significant level of survival of non-IPTG resistant cells was occurring

TABLE I.2

Production of IPTG resistant survivors of 186 lysogens carrying *B* expression plasmid pEC434.

Culture[a]	Ratio of IPTG resistant survivors[b]
(186$^+$) p*lacI*^{q} p$p_{trc}B^+$	3.5×10^{-3}
(186$^+$) p*lacI*^{q} p$p_{trc}Bam$	0.61

Notes to Table I.2

a. Cultures of E251(186$^+$) (E635) carrying the p*lacI*^{q} plasmid pMC7 (Calos, 1978; Section 2.3.1) and the p$p_{trc}B^+$ plasmid (pEC434) or the p$p_{trc}Bam$ plasmid (pEC435) were grown at 30°C in M13 minimal medium containing the appropriate antibiotics to A_{600}=0.4, and IPTG added to 1 mM final concentration. Cultures were grown for 16 hours at 30°C, subcultured into M13 minimal media containing the appropriate antibiotics supplemented with 1 mM IPTG (final concentration) and grown at 30°C to A_{600}=0.8. Dilutions of the cultures were plated onto M13 minimal plates containing the appropriate antibiotics supplemented with 1 mM IPTG (final concentration) or without IPTG. Plates were grown at 30°C for 36 hours and colony forming units scored. Plasmids pEC434 and pEC435 are described in Section 2.3.2.

b. Ratio of IPTG resistant survivors is the ratio of colony forming units obtained on M13 minimal plates supplemented with IPTG to the colony forming units obtained on M13 minimal plates without IPTG.

on IPTG induction in liquid culture. Nonetheless, the survivors isolated on minimal plates supplemented with IPTG remained resistant to this treatment, as shown by their ability to be replated on IPTG plates with almost 100% efficiency. Treatment of a 186 lysogen, carrying the *Bam* expression plasmid pEC435 and plasmid pMC7 with IPTG in an identical manner as described above, had little effect on cell viability, plating with a rate of survival in the presence of IPTG of 0.61, with respect to plating in the absence of IPTG (Table I.2).

Of 150 survivors resistent to IPTG that were tested, all were found to be resistent to 186 cI^- infection as determined by cross-streaking (Section 2.15.1). This indicated that loss of the prophage was not the reason for their acquired resistance to IPTG. However, only eleven of the 150 survivors were found to be resistant to 186 *vir*1, as determined by cross-streaking. The vast majority of survivors isolated which were resistant to IPTG (but 186 *vir* sensitive), were presumably mutants in the *B* expression plasmid pEC434 not expressing functional B protein, mutants in the lethal prophage function activated by B protein from plasmid pEC434, or mutants resistant to induction with IPTG. However, only the 186 *vir* resistant survivors were chosen for further characterization of their ability to activate late transcription. It was also recognized that that these survivors may not have been of an independent origin, but their ability to activate late transcription was characterized to assess the procedure as a method for the isolation of host mutants blocked in late activation.

I.2.3 Functionality of *B* expression plasmid pEC434 in the 186 resistant mutants.

Before investigating whether the 186 resistant survivors of IPTG induction were defective in activation of late gene transcription, it was appropriate to determine whether the *B* expression plasmid pEC434 from these mutants was still capable of activating late gene transcription. This was done for two reasons. Firstly, in order to eventually assess if the mutants were blocked in late activation, the mutants were to be transformed with the $p_{Y}galK$ plasmid (pEC433) and the ability of the *B* expression plasmid to activate p_{Y} transcription tested. It was therefore necessary to show that inability to activate was not due to an inability of the *B* expression plasmid to activate late transcription, due to some plasmid mutation. Secondly, it was also possible that mutants in plasmid pEC434 were

responsible for conferring 186 *vir* resistance, and therefore that the 186 resistance observed was not due to a host mediated block to activation of late transcription.

Plasmid DNA was prepared (Section 2.21.1) from the eleven 186 resistant colonies isolated in Section I.2.2. Plasmid DNA from seven of the 186 resistant colonies was found to be unable to transform a 186 lysogen, demonstrating that plasmid pEC434 from these survivors was expressing functional B protein capable of killing a lysogen. In addition, plasmid DNA preparations from these same seven 186 resistant colonies was capable of activating p_V transcription on the p_V*galK* plasmid pEC433, as shown by the red colony phenotype on MacConkey indicator plates when transformed into a $galE^+T^+K^-$ strain (E863) carrying the p_V*galK* plasmid pEC433. Therefore, plasmid pEC434 in these 186 resistant mutants was concluded to be active in expressing functional B protein. Further, plasmid pEC434 DNA isolated from these 186 resistant mutants did not confer 186 *vir*1 resistance when transformed into strain E251, as determined by cross-streaking (Section 2.15.1), showing that the resistance to 186 in these survivors was likely to be host or prophage mediated.

Plasmid DNA isolated from the remaining four 186 *vir* resistant colonies was capable of transforming a 186 lysogen at an equivalent efficiency to the *Bam* expression plasmid pEC435, and was also unable to activate the p_V*galK* plasmid pEC433, as demonstrated by the white colony phenotype on MacConkey indicator plates when transformed into E863 carrying the p_V*galK* plasmid pEC433. These results demonstrated that the *B* expression plasmid pEC434 from these mutants was not expressing functional B protein capable of activating late gene transcription. Nor did the plasmid DNA from these 186 resistant mutants confer 186 *vir*1 resistance when transformed into strain E251, as determined by cross-streaking (Section 2.15.1). Despite the fact that plasmid pEC434 in these mutants was not capable of activating late transcription, presumably due to mutations in the *B* gene, the nature of the resistance to 186 *vir* infection in this class of mutants remains unknown, although it was possible that the resistance to 186 *vir* could be conferred from these plasmids on transformation into a 186 lysogen.

I.2.4 Curing the 186 resistant mutants of plasmid pMC7.

The seven 186 resistant mutants isolated that carried functional plasmid pEC434 were finally screened for their ability to activate p_V transcription on the p_V*galK* plasmid pEC433. An inability of the *B* expression plasmid pEC434 to activate p_V transcription in these 186 resistant strains would indicate that host mutants had been identified which blocked activation of late gene transcription. To test whether the *B* expression plasmid pEC434 was capable of activating p_V transcription, it was first necessary to introduce the p_V*galK* plasmid pEC433 into these 186 resistant mutants.

The seven 186 resistant mutants isolated by the IPTG induction protocol (Section I.2.2) carried plasmid pMC7 in addition to the *B* expression plasmid pEC434. Plasmid pEC434 is a derivative of the R1-derived vector pEC612 (Section 2.3.2). Plasmid pMC7 is a *lacI*^{q} clone of the tetracycline resistant, ColE1 derivative pMB9 (Calos, 1978). The p_V*galK* plasmid pEC433 carries the ColE1 related replicon from pMB1 (de Boer, 1984) and therefore was expected to be unable to be maintained in the same host as plasmid pMC7 due to plasmid incompatability. In order to introduce the p_V*galK* plasmid pEC433 into the 186 resistant mutants carrying functional *B* expression plasmid pEC434 and thereby determine if late gene transcription was blocked, it was first necessary to cure the 186 resistant mutants of plasmid pMC7.

Initial attempts at curing the 186 resistant mutants of plasmid pMC7 (which confers resistance to tetracycline) by several passages of growth in M13 minimal media and ampicillin (to maintain selection for pEC434) were unsuccessful (Section 2.15.3). No loss of tetracycline resistance was observed after three passages of growth to stationary phase in M13 minimal media supplemented with ampicillin, as determined by plating the cultures onto M13 plates with and without the addition of tetracycline (15 ug/ml), and assessing the loss of tetracycline resistance (Section 2.15.3). Since no loss of tetracycline resistance was identified by such a procedure, it was concluded that propagating the cells in minimal media in the absence of tetracycline selection was not sufficient to result in loss of plasmid pMC7. However, the p_V*galK* plasmid pEC433 was successfully introduced into the seven 186 resistant mutants by direct transformation (Section 2.15.2(b)). In this method, the 186 resistant cells were transformed (Section 2.15.2(b)) with the kanamycin resistant p_V*galK* plasmid pEC433, allowed to express the antibiotic resistance for two hours in

minimal media, subcultured into minimal media supplemented with ampicillin (25 ug/ml) and kanamycin (25 ug/ml), and allowed to reach stationary phase. Cells were plated onto M13 minimal plates supplemented with ampicillin (25 ug/ml) and kanamycin (25 ug/ml) and individual colonies picked and tested for their resistance to tetracycline (15 ug/ml). Three of the seven 186 resistant mutants transformed were found to give tetracycline sensitive colonies in 40%-80% of the colonies tested, and agarose gel electrophoresis of plasmid DNA preparations (Section 2.21.1) from these tetracycline sensitive mutants confirmed the loss of plasmid pMC7 (data not shown).

Using this method of direct transformation and kanamycin selection, four of the seven 186 resistant mutants showed no loss of tetracycline resistance in over 100 colonies tested for each mutant. The colonies isolated on kanamycin selection were all found to be resistent to tetracycline (15 ug/ml), in addition to ampicillin (25 ug/ml) and kanamycin (25 ug/ml). This inability to be cured of plasmid pMC7 in these mutants was surmised to be due to the need to maintain *lacI* repression of *B* expression from plasmid pEC434, as it appeared that loss of $lacI^q$ control in these survivors would result in derepression of *B* gene transcription and killing of the lysogen.

I.2.5 Activation of p_V transcription in 186 resistant mutants.

To determine whether the 186 resistant mutants cured of plasmid pMC7 and transformed with the p_V*galK* plasmid pEC433 were capable of activating late transcription, the level of galactokinase activity from the p_V*galK* plasmid pEC433 in the presence of the *B* expression plasmid pEC434 in some of these mutants was determined.

The results are presented in Table I.3. The level of galactokinase activity was determined in the presence of 1 mM IPTG, in order that the level of p_V transcription determined was under conditions of certain derepression of *B* gene transcription from plasmid pEC434. The p_V*galK* plasmid pEC433 showed extensive activation in the presence of the original *B* expression plasmid pEC434 in the non-lysogenic strain E251 (74.1 units; Table I.3). However, each of the three 186 resistant mutants cured of tetracycline resistance showed a reduction of galactokinase activity to a level of approximately 10% of the non-lysogenic control (186 resistant mutants 18, 28 and 37; Table I.3). This was not a result of decreased template number of pEC433 in these

TABLE I.3

Activation of p_V transcription in 186 resistant mutants.

Culture[a]	Galactokinase activity	Relative[b] activity
$pp_{trc}B^+$ p_VgalK	74.1	100%
(186$^+$) $pp_{trc}B^+$ p_VgalK 186 resistant No. 18	9.6	13.0%
(186$^+$) $pp_{trc}B^+$ p_VgalK 186 resistant No. 28	7.2	9.7%
(186$^+$) $pp_{trc}B^+$ p_VgalK 186 resistant No. 37	5.1	6.9%
$pp_{trc}B^+$(No. 37) p_VgalK	69.3	93.5%
(186$^+$) $pp_{trc}B^+$ p_VgalK 186 resistant No. 26	65.3	88.1%

Notes to Table I.3

a. Cultures of the non-permissive strain E251 and the corresponding 186 resistant mutants (derived from E251(186$^+$)) carrying the $pp_{trc}B^+$ plasmid (pEC434) and the p_VgalK plasmid (pEC433) were grown overnight in M13 minimal medium containing the appropriate antibiotics at 30°C. The cultures were subcultured into the same media and grown at 30°C to A_{600}=0.2, IPTG (1 mM final concentration) added and incubated at 30°C for 2 hours, and the galactokinase activty determined (Section 2.18). 186 resistant mutants 18, 28 and 37 were able to be cured of plamid pMC7, while 186 resistant mutant 26 was not (Section I.2.4). $pp_{trc}B^+$ p_VgalK designates the corresponding non-lysogen E251 carrying the $pp_{trc}B^+$ plasmid pEC434 and the p_VgalK plasmid pEC433. $pp_{trc}B^+$(No. 37) designates plasmid pEC434 isolated from mutant 37 tramsformed into E251 carrying the p_VgalK plasmid pEC433. Plasmids pEC434 and pEC433 are described in Section 2.3.2.

b. Percentage activity relative to E251 carrying the $pp_{trc}B^+$ plasmid pEC434 and the p_VgalK plamid pEC433 (74.1 units).

mutants, as the copy number of plasmid pEC433 in these 186 resistant derivatives was equivalent to that of the non-lysogenic control, as determined by total plasmid copy number estimation of whole cell lysates (Section 2.19.3; data not shown). In addition, plasmid pEC434 DNA isolated from one of the 186 resistant mutants (186 resistant mutant 37) was not reduced in its ability to activate p_V transcription, since transformation of this plasmid DNA into strain E251 carrying the p_V*galK* plasmid pEC433 gave equivalent levels of galactokinase to the original non-lysogenic control containing the original *B* expression plasmid pEC434 (Table I.3; 69.3 units). Therefore the 186 resistant mutants showing decreased p_V activation were concluded to be severely reduced in their ability to activate late transcription.

The level of galactokinase activity from a 186 resistant mutant of the class retaining the tetracycline resistant plasmid pMC7 (186 resistant mutant 26; Section I.2.4), but containing the *B* expression plasmid pEC434 and the p_V*galK* plasmid pEC433, was also determined. Consistent with the fact that this class of mutants was unable to be cured of the *lacI*^{q} plasmid pMC7, this mutant showed equivalent levels of galactokinase activity to the non-lysogenic control (Table I.3; 65.3 units). These mutants were concluded to be still functional in activation of late transcription, consistent with the inability of these mutants to be cured of the *lacI*^{q} plasmid pMC7, since this would result in cell death (Section I.2.1). The reason for the resistance to IPTG and 186 *vir* in this class of mutants remains to be determined.

I.3 SUMMARY.

The results presented established that the induction of *B* expression from pEC434 in a 186 lysogen resulted in the isolation of several mutants severely reduced in their ability to activate late transcription. These mutants are likely to be host mutants blocked in activation of late transcription, although it is conceivable that these mutants could be prophage mutants blocking late gene activation.

Studies are currently underway by R. Tiwari in this laboratory to determine the precise chromosomal locations of mutants isolated by a similar procedure which are blocked in activation of late transcription. These mutants were obtained by isolation of surviving colonies on transformation of a 186 lysogen with plasmid pEC434, in the

absence of *lacI*^{q} control. In this protocol, nitrosoguanidine mutagenesis of the host lysogen was used to increase the rate of survival upon transformation. The map positions of the loci of these mutants may enable the identification of the host functions involved in activation of late transcription, exposed by this protocol. Since the *B* gene is expected to interact with the α subunit of RNA polymerase, by analogy with the interaction of the P2 *ogr* gene product with the α subunit of RNA polymerase (Section 1.5.5), one class of host mutants isolated is expected to map in the host *rpoA* gene. A critical verification that the host mutants isolated were indeed blocked in activation of late transcription, would be the isolation of mutants in the *B* gene of 186 that were able to overcome this block to 186 infection.

APPENDIX II

THE EFFECT OF REPLICATION ON THE EXPRESSION OF AN AUTOGENOUSLY REPRESSED GENE

APPENDIX II. THE EFFECT OF REPLICATION ON THE EXPRESSION OF AN AUTOGENOUSLY REPRESSED GENE.

II.1 INTRODUCTION.

Phage replication is a dynamic process that provides a large increase in the number of phage chromosomes in a relatively short time. Earlier in my thesis research it was believed probably that transcription of the 186 gene *B* was directly autogenously controlled and the question arose as to what effect replication would have on the expression of a gene which represses its own transcription. The process of replication not only provides an increase in the dosage of the negatively autoregulated gene, but also provides a rapid increase in the concentration of operator sites. This appendix investigates the role of an increase in gene dosage and an increase in operator concentration on the expression of an autogenously controlled gene. Since the product of the autogenously controlled gene may also be involved in controlling expression of other genes, replication has the potential to indirectly control gene expression by controlling expression of the autoregulated gene.

II.2 TRANSIENT DEREPRESSION OF TRANSCRIPTION OF AN AUTOGENOUSLY REPRESSED GENE BY REPLICATION.

Consider the situation in which the repression of a gene is by the binding of the product of that gene to block initiation of transcription. The process of replication will provide a rapid increase in the template number and therefore a corresponding increase in the concentration of operator sites for the repressor of the autoregulated gene. The dynamic nature of this rapid increase in free operator concentration should result in a reduction in the pool of free repressor molecules and therefore should lead, at least transiently, to a large derepression of transcription. This derepression would lessen in extent as the system moved to its new steady state. However, the large transient derepression of transcription may be all that is needed to initiate a cascade of events, involving the expression of other genes under the control of the repressor function. Therefore replication would appear to have the capacity to control gene expression by derepression of an autogenously controlled gene.

II.3 RELATIONSHIP OF THE DOSAGE OF AN AUTOGENOUSLY REPRESSED GENE TO THE CONCENTRATION OF FREE REPRESSOR.

The second effect of replication will be to provide an increase in the dosage of an autogenously repressed gene. This raises the question as to what effect an increase in gene dosage would have on the concentration of free repressor at the new steady state. Under conditions of increased gene dosage, the concentration of free repressor is expected to increase by consideration of the free repressor and free operator concentrations at equilibrium, even though it may seem an enigma that an increase in the dosage of an autogenously repressed gene will lead to an increase in free repressor concentration.

Consider the simple system in which the autogenous repression of a gene is by the binding of a single repressor molecule (the product of the autoregulated gene) to a single operator site and that this binding prevents the initiation of transcription from the promoter for the repressor gene. In such a system, the dissociation constant K_D is defined

$$RO \rightleftharpoons R + O;$$

$$K_D = \frac{[R][O]}{[RO]}$$

where [R] = free repressor concentration, [O] = free operator concentration and [RO] = bound repressor-operator concentration.

In this system, an operator site not occupied by repressor is free to bind RNA polymerase and therefore available to initiate transcription of the repressor gene. Consider the situation where the dosage of an autogenously controlled gene increases. An increase in gene dosage necessarily means an increase in the total operator concentration $[O_T]$, which is given by the sum of the free operator concentration [O] and the bound repressor-operator concentration [RO] (i.e. $[O_T] = [O] + [RO]$). If the free repressor concentration were to remain unchanged under such conditions, the term {[O]/[RO]} must remain unchanged to maintain the constant K_D. Since the total operator concentration has increased, the terms [O] and [RO] must increase proportionately to maintain constant {[O]/[RO]}. However, an increase in free operator concentration should result in an increase in free repressor concentration due to increased transcription of the repressor gene, inconsistent with the original maintainence of an unchanged repressor concentration. If the free repressor concentration were to increase this would necessitate a decrease in the

term {[O]/[RO]} to maintain the constant K_D. An absolute increase in the free operator concentration [O] is still compatible with a decrease in the term {[O]/[RO]} upon increasing the total operator concentration, as [O] and [RO] may both increase, but [RO] may increase over [O] relatively, in order to decrease the term {[O]/[RO]}. Thus an increase in free repressor concentration upon increasing gene dosage is the only situation consistent with maintainence of the equilibrium constant and the relationship between the free operator concentration and free repressor concentration. Thus it would appear that an increase in the dosage of a negatively autoregulated gene will lead to an increase in free repressor concentration, and therefore that replication has the potential to control gene expression in this manner.

II.4 SUMMARY.

It is concluded that replication could have two effects on the expression of an autogenously repressed gene. In the first stages of the disturbed equilibrium there would be a large but transient derepression of transcription, which may be all that is needed to initiate a cascade of events. Secondly, in the steady state (if reached), the concentration of free repressor of the autogenously controlled gene is expected to increase with increasing gene dosage, and this increase in concentration could also have a role in controlling gene expression. Thus it would appear that as a general principle, replication could be used to control expression of an autogenously repressed gene, and thereby indirectly control the expression of other genes under the control of the autoregulated repressor.

BIBLIOGRAPHY

BIBLIOGRAPHY

Adhya, S. and Gottesman, M. (1978). *Ann. Rev. Biochem.*, **47**: 967-996.

Adhya, S. and Miller, W. (1979). *Nature*, **279**: 492-494.

Adhya, S., Sarkar, P., Valenzuela, D. and Maitra, V. (1979). *Proc. Nat. Acad. Sci., U.S.A.*, **76**: 1613-1617.

Albright, L.M., Kassavetis, G.A. and Geiduschek, E.P. (1988). *J. Bacteriol.*, **170**: 1279-1289.

Anderson, J.E., Ptashne, M. and Harrison, S. (1985). *Nature*, **316**: 596-601.

Aoyama, T., Takanami, M., Ohtsuka, E., Taniyama, Y., Marumoto, R., Sato, H. and Ikehara, M. (1983). *Nucl. Acids Res.*, **11**: 5855-5864.

Appleyard, R.K. (1954). *Genetics*, **39**: 440-452.

Auerswald, E.-A., Ludwig, G., Schaller, H. (1981). *Cold Spring Harbor Symp. Quant. Biol.*, **45**: 107-113.

Baldwin, R.L., Barrand, P., Fritsch, A., Goldthwait, D.A. and Jacob, F. (1966). *J. Mol. Biol.*, **17**: 343-357.

Barrett, K.J., Gibbs, W. and Calendar, R. (1972). *Proc. Nat. Acad. Sci., U.S.A.*, **69**: 2986-2990.

Barrett, K.J., Marsh, M.L. and Calendar, R. (1976). *J. Mol. Biol.*, **106**: 683-707.

Beck, C.F. and Warren, R.A.J. (1988). *Microbiol. Rev.*, **52**: 318-326.

Bertani, G. (1975). *Mol. Gen. Genet.*, **136**: 107-137.

Bertani, L.E. (1968). *Virology*, **36**: 87-103.

Bertani, L.E. and Bertani, G. (1971). *Adv. Genet.*, **16**: 199-237.

Bertani, L.E. and Six, E.W. (1988). In "The Bacteriophages" (R. Calendar, ed.). Vol. 2, pp. 72-143 (Plenum Publishing).

Birkeland, N.K. and Lindqvist, B.H. (1986). *J. Mol. Biol.*, **188**: 487-490.

Birkeland, N.K., Christie, G.E. and Lindqvist, B.H. (1988). *Gene.*, **73**: 327-335.

Birnboim, H.C. and Doly, J. (1979). *Nucl. Acids Res.*, **7**: 1513-1523.

Bonekamp, F., Anderson, H.D., Christensen, T. and Jensen, K.F. (1985). *Nucl. Acids* Res., **13**: 4113-4123.

Bowden, D.W. and Calendar, R. (1979). *J. Mol. Biol.*, **129**: 1-18.

Bradley, C., Ong, P.L. and Egan, J.B. (1975). *Mol. Gen. Genet.*, **140**: 123-135.

Brody, E., Rabussay, D. and Hall, D.H. (1983). In "Bacteriophage T4" (Matthews, C.K., Kutter, E.M., Mosig, G and Berget, P.B., eds.) pp. 174-183. (Amer. Soc. for microbiol., Washington D.C.).

Brosius, J., Dull, T.J., Sleeter, D.D. and Noller, H.F. (1981). *J. Mol. Biol.*, **148**: 107-121.

Brosius, J., Erfle, M. and Storella, J. (1985). *J. Biol. Chem.*, **260**: 3539-3541.

Bruening, G., Gould, A.R., Murphy, P.J. and Symons, R.H. (1982). *F.E.B.S. letters*, **148**: 71-78.

Brumby, A. (1986). Homours Thesis, University of Adelaide.

Busby, S.J. (1986). In "Regulation of Gene expression - 25 years on. Soc. Gen. Microbiol. Symp. Vol. 39" (Booth, I.R. and Higgins, C.F., eds.) pp. 51-77. (Cambridge University Press).

Bushman, F.D. and Ptashne, M. (1988). *Cell*, **54**: 191-197.

Calendar, R., Lindahl, G., Marsh, M. and Sunshine, M. (1972). *Virology*, **47**: 68-75.

Calos, M.P. (1978). *Nature*, **274**: 762-765.

Campbell, A. (1965). *Virology*, **27**: 329-339.

Capaldi, R.A. and Vanderkooi, G. (1972). *Proc. Nat. Acad. Sci., U.S.A.*, **69**: 930-932.

Carter, D. (1985). Honours Thesis, University of Adelaide.

Ceruzzi, M.A.F., Bektesh, S.L. and Richardson, J.P. (1985). *J. Biol. Chem.*, **260**: 9412-9418.

Chamberlin, M.J. and Ryan, T. (1982). In "The Enzymes Vol. 15." (Boyer, P., ed.) pp. 87-108. (Acad. Press, N.Y.).

Chang, A.C.Y. and Cohen, S.N. (1978). *J. Bacteriol.*, **134**: 1141-1166.

Chattoraj, D.K. (1977). In "DNA Insertion Elements, Plasmids, and Episomes" (A.I. Bukhari, J.A. Shapiro, and S.L. Adhya, eds.), pp. 733-736 (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.).

Chattoraj, D.K. and Bertani, G. (1980). *Mol. Gen. Genet.*, **178**: 85-90.

Chattoraj, D.K. and Inman, R.B. (1972). *J. Mol. Biol.*, **66**: 423-434.

Chattoraj, D.K. and Inman, R.B. (1973). *Proc. Nat. Acad, Sci., U.S.A.* **70**: 1768-1771.

Chattoraj, D.K. and Inman, R.B. (1974). *Proc. Nat. Acad. Sci., U.S.A.*, **71**: 311-314.

Chattoraj, D.K., Schnos, M. and Inman, R.B. (1973). *Virology*, **55**: 439-444.

Chattoraj, D.K., Younghusband, H.B. and Inman, R.B. (1975). *Mol. Gen. Genet.*, **136**: 139-149.

Chen, C.-Y.A., Galluppi, G.R. and Richardson, J.P. (1986). *Cell*, **46**: 1023-1028.

Christie, G.E. and Calendar, R. (1983). *J. Mol. Biol.*, **167**: 773-790.

Christie, G.E. and Calendar, R. (1985). *J. Mol. Biol.*, **181**: 373-382.

Christie, G.E., Haggard-Ljungquist, E., Feiwell, R. and Calendar, R. (1986). *Proc. Nat. Acad. Sci., U.S.A.*, **83**: 3238-3242.

Court, D., Brady, C., Rosenberg, M., Wulff, D.L., Behr, M., Mahoney, M. and Izumi, S. (1980). *J. Mol. Biol.*, **138**: 231-254.

Crosby, B. (1985). Honours Thesis, University of Adelaide.

Dale, E.C., Christie, G.E. and Calendar, R. (1986). *J. Mol. Biol.*, **192**: 793-803.

Debarbouille, M., Shuman, A.H., Silhavy, T.J. and Schwartz, M. (1978). *J. Mol. Biol.*, **124**: 359-371.

de Boer, H.A. (1984). *Gene*, **30**: 251-255.

de Crombrugghe, B., Busby, S. and Buc, H. (1984). *Science*, **224**: 831-837.

Deho, G., Zangrossi, S., Ghisotti and Sironi, G. (1984). *J. Virol.*, **62**: 1697-1704.

Devereux, J., Haberli, P. and Smithies, O. (1984). *Nuc. Acids Res.*, **12**: 387-395.

Dibbens, J. (1984). Honours Thesis, University of Adelaide.

Dodd, I.B. and Egan J.B. (1987). *J. Mol. Biol.*, **194**: 557-564.

Dodd, I.B., Kalionis, B. and Egan J.B. (1990). *J. Mol. Biol.*, in press.

Dodd, I.B. and Egan J.B. (1990). *J. Mol. Biol.*, submitted.

Doi, R.H. and Wang, L.-F. (1986). *Microbiol. Rev.*, **50**: 227-243.

Doolittle, R.F. (1981). *Science*, **214**: 149-159.

Dunn, J.J. and Studier, F.W. (1981). *J. Mol. Biol.*, **148**, 303-330.

Dunn, J.J. and Studier, F.W. (1983). *J. Mol. Biol.*, **166**: 477-535.

Englesberg, E. and Wilcox, G. (1974). *Ann. Rev. Genet.*, **8**: 219-242.

Evans, R.M. and Hollenberg, S.M. (1988). *Cell*, **52**: 1-3.

Fickett, J.W. (1982). *Nucl. Acids Res.*, **10**: 5303-5318.

Finnegan, E.J. (1979). Ph.D. Thesis, University of Adelaide.

Finnegan, E.J. and Egan, J.B. (1979). *Molec. Gen. Genet.*, **172**: 287-293.

Finnegan, E.J. and Egan, J.B. (1981). *J. Virol.*, **38**: 987-995.

Franklin, N. (1985a). *J. Mol. Biol.*, **181**: 75-84.

Franklin, N. (1985b). *J. Mol. Biol.*, **181**: 85-91.

Fried, M. and Crothers, D.M. (1981). *Nucl. Acids Res.*, **9**: 6505-6525.

Friedman, D.I. and Gottesman, M. (1983). In "Lambda II" (Hendrix, R.W., Roberts, J.W., Stahl, F.W. and Weisberg, R.A., eds.). pp. 21-51. (Cold Spring Harbor Lab., Cold Spring Harbor N.Y.).

Fujicki, H., Palm, P., Zillig, W. and Calendar, R. (1976). *Mol. Gen. Genet.*, **145**: 19-22.

Funnell, B.E. and Inman, R.B. (1982). *J. Mol. Biol.*, **154**: 85-105.

Galas, D.J. and Schmitz, A. (1978). *Nucl. Acids Res.*, **5**: 3157-3170.

Galas, D.J., Eggert, M. and Waterman, M.S. (1985). *J. Mol. Biol.*, **186**: 117-128.

Galloway, J.L. and Platt, T. (1985). In "Regulation of gene expression - 25 years on. Soc. Gen. Microbiol. Symp. Vol. 39" (Booth, I.R. and Higgins, C.F., eds.) pp. 135-178. (Cambridge University Press).

Garen, A., Garen, S. and Wilhelm, R.C. (1965). *J. Mol. Biol.*, **14**: 167-178.

Garrett, S. and Silhavy, R.J. (1987). *J. Bacteriol.*, **169**: 1379-1385.

Gartenberg, M.R. and Crothers, D.M. (1988). *Nature*, **333**: 824-829.

Geiduschek, E.P. and Kassavetis, G.A. (1988). In "The Bacteriophages",.(R. Calendar, ed.). Vol. 1, pp. 93-115 (Plenum Publishing).

Geisselsoder, J., Mandel, M. and Calendar, R. (1973). *J. Mol. Biol.*, **77**: 405-415.

Geisselsoder, J., Youderian, P., Deho, G., Chidambaram, M., Goldstein, R.N. and Ljungquist, E. (1981). *J. Mol. Biol.*, **148**: 1-19.

Gibbs, W., Goldstein, R.N., Weiner, R., Lindqvist, B. and Calendar, R. (1973). *Virology*, **53**: 24-39.

Gibbs, W., Eisen, H. and Calendar, R. (1983). *J. Virol.*, **47**: 392-398.

Giffard, P.M., Rowland, G.C., Kroll,R.G., Stewart, L.M.D., Bakker, E.P. and Booth, I.R. (1985). *J. Bacteriol.*, **164**: 904-910.

Glisin, V., Crkvenjakov, R. and Byus, C. (1974). *Biochem.*, **13**: 2633-2637.

Goldfarb, D.S., Wong, S.L., Kudo, T. and Doi, R.H. (1983). *Mol. Gen. Genet.*, **191**: 319-325.

Gonzalez, N., Wiggs, J. and Chamberlin, M. (1977). *Arch. Biochem. Biophys.*, **182**: 404-408.

Goutte, C. and Johnson, A.D. (1988). *Cell*, **52**: 875-882.

Gouy, M. and Gautier, C. (1982). *Nucl. Acids Res.*, **10**: 7055-7074.

Grantham, R., Gautier, C., Gouy, M., Jacobzone, M. and Mercier, R. (1981). *Nuc. Acids Res.*, **9**: 43-74.

Gribskov, M., Devereux, J. and Burgess, R.R. (1984). *Nuc. Acids Res.*, **12**: 539-549.

Grosjean, H. and Fiers, W. (1982). *Gene*, **18**: 199-209.

Grossman, A.D. and Losick, R. (1986) In "Regulation of gene expression - 25 years on. Soc. Gen. Microbiol. Symp. Vol 39" (Booth, I.R. and Higgins, C.F., eds.) pp. 127-138. (Cambridge University Press).

Grossman, A.D., Erichson, J.W. and Gross, C.A. (1984). *Cell*, **38**: 383-390.

Grunstein, M. and Hogness, D.S. (1975). *Proc. Nat. Acad. Sci., U.S.A.*, **72**: 3961-3965.

Guarente, L., Nye, J.S., Hochschild, A. and Ptashne, M. (1982). *Proc. Nat. Acad. Sci., U.S.A.*, **79**: 2236-2239.

Guidolin, A., Zingg, J.-M., Lehnherr, H. and Arber, W. (1989). *J. Mol. Biol.*, **208**: 615-622.

Gussin, G.N., Johnson, A.D., Pabo, C.O. and Sauer, R.T. (1983). In "Lambda II" (Hendrix, R.W., Roberts, J.W., Stahl, F.W., and Weisberg, R.A., eds.) pp. 93-121. (Cold Spring Harbor Lab., Cold Spring Harbor N.Y.).

Haggard-Ljungquist, E., Kockum, K. and Bertani, L.E. (1987). *Mol. Gen. Genet.*, **208**: 52-56.

Hall, M.N. and Silhavy, T.J. (1979). *J. Bacteriol.*, **140**: 843-847.

Hall, M.N. and Silhavy, T.J. (1981a). *J. Mol. Biol.*, **146**: 23-47.

Hall, M.N. and Silhavy, T.J. (1981b). *J. Mol. Biol.*, **151**: 1-15.

Halling, C. (1989). Ph.D Thesis, University of California (Berkeley).

Halling, C. and Calendar, R. (1990). *J. Bacteriol.*, submitted.

Halling, C., Sunshine, M.G., Lane, K.B., Six, E.W. and Calendar, R. (1990). *J. Bacteriol.*, submitted.

Harris, J.D. and Calendar, R. (1978). *Virology*, **85**: 343-358.

Hawley, D.K. and McClure, W.R. (1982). *J. Mol. Biol.*, **157**: 493-525.

Hawley, D.K. and McClure, W.R. (1983). *Nucl. Acids Res.*, **11**: 2237-2255.

Haynes, L.L. and Rothman-Denes, L.B. (1985). *Cell*, **41**: 597-605.

Helmann, J.D. and Chamberlin, M.J. (1988). *Ann. Rev. Biochem.*, **57**: 839-872.

Herskowitz, I. and Hagen, D. (1980). *Ann. Rev. Genet.*, **14**: 399-445.

Ho, Y.-S. and Rosenberg, M. (1988). In "The Bacteriophages",.(R. Calendar, ed.). Vol. 2, pp. 725-756. (Plenum Publishing).

Ho, Y.-S., Wulff, D. and Rosenberg, M. (1983). *Nature*, **304**: 703-708.

Ho, Y.-S., Wulff, D. and Rosenberg, M. (1986). in "Regulation of Gene expression - 25 years on. Soc. Gen. Microbiol. symp. Vol. 39" (Booth, I.R. and Higgins, C.F., eds.) pp. 79-103. (Cambridge University Press).

Hochschild, A., Irwin, N. and Ptashne, M. (1983). *Cell*, **32**: 319-325.

Hocking, S.M. (1977). Ph.D. Thesis, University of Adelaide.

Hocking, S.M. and Egan, J.B. (1982a). *Mol. Gen. Genet.*, **187**: 87-95.

Hocking, S.M. and Egan, J.B. (1982b). *Mol. Gen. Genet.*, **187**: 174-176.

Hocking, S.M. and Egan, J.B. (1982c). *J. Virol.*, **44**: 1068-1071.

Hocking, S.M. and Egan, J.B. (1982d). *J. Virol.*, **44**: 1056-1067.

Holmes, W.M., Platt, T. and Rosenberg, M. (1983). *Cell*, **32**: 1029-1032.

Hooper, I. (1979). Ph.D. Thesis, University of Adelaide.

Hooper, I. and Egan, J.B. (1981). *J. Virol.*, **40**: 559-601.

Hoopes, B.C. and McClure, W.R. (1987). In "Escherichia coli and Salmonella typhimuriam - Cellular and Molecular Biology" (Neidhardt, F.C. chief ed., Ingraham, J.L., Brooks Low, K., Magasanik, B., Schaechter, M. and Umbarger, H.E. eds.) Vol. 2. pp. 1231-1240. (American Society for Microbiology, Washington D.C.).

Huddleston, V., (1970). Honours Thesis, University of Adelaide.

Hyde, J.M. and Bertani, G. (1975). *J. Gen. Virol.*, **28**: 415-419.

Ikemura, T. (1981a). *J. Mol. Biol.*, **146**: 1-21.

Ikemura, T. (1981b). *J. Mol. Biol.*, **151**: 389-409.

Inman, R.B., Schnos, M.,Simon, L.D., Six E.W. and Walker, D.H. Jr. (1971). *Virology*, **44**: 67-72.

Jacobs, K. and Geiduschek, E.P. (1981). *J. Virol.*, **39**: 31-45.

Johnson, W.C., Moran, C.P. Jr. and Losick, R. (1983). *Nature*, **302**: 800-804.

Kalionis, B. (1985). Ph.d. Thesis, University of Adelaide.

Kalionis, B. and Egan, J.B. (1981). *Gene.*, **15**: 95-98.

Kalionis, B., Dodd, I.B. and Egan, J.B. (1986a). *J. Mol. Biol.*, **191**: 199-209.

Kalionis, B., Pritchard, M. and Egan, J.B. (1986b). *J. Mol. Biol.*: **191**: 211-220.

Kammerer, W., Deuschle, U., Gentz, R. and Bujard, H. (1986). *EMBO J.*, **5**: 2995-3000.

Keener, J.E., Dale, E.C., Kustu, S. and Calendar, R. (1988). *J. Bacteriol.*, **170**: 3543-3546.

Kolling, R. and Lother, H. (1985). *J. Bacteriol.*, **164**: 310-315.

Konigsberg, W. and Godson, G.N. (1983). *Proc. Nat. Acad. Sci., U.S.A.*, **80**: 687-697.

Kustu, S., Sei, K. and Keener, J. (1986). in "Regulation of Gene expression - 25 years on. Soc. Gen. Microbiol. symp. Vol. 39" (Booth, I.R. and Higgins, C.F., eds.) pp. 139-154. (Cambridge University Press).

Lagos, R., Jiang, R.Z., Kim, S. and Goldstein, R. (1986). *Proc. Nat. Acad. Sci., U.S.A.*, **83**: 9561-9565.

Lamont, I., Kalionis, B. and Egan J.B. (1988). *J. Mol. Biol.*, **199**: 379-382.

Lamont, I., Brumby, A.M. and Egan, J.B. (1989). *Proc. Nat. Acad. Sci., U.S.A.*, **86**: 5492-5496.

Landick, R., Vaughn, V., Lau, E.T., VanBogelen, R.A., Erickson, J.W. and Neidhardt, F.C. (1984). *Cell*, **38**: 175-182.

Lanzer, M. and Bujard, H. (1988). *Proc. Nat. Acad. Sci., U.S.A.*, **85**: 8973-8977.

Larsen, J.E.L., Gerdes, K., Light, J. and Molin, S. (1984). *Gene*, **28**: 45-54.

Lau, L.F. and Roberts, J.W. (1985). *J. Biol. Chem.*, **260**: 574-584.

Lau, L.F., Roberts, J.W. and Wu, R. (1982). *Proc. Nat. Acad. Sci., U.S.A.*, **79**: 6171-6175.

Lee, N., Gielow, W. and Wallace, R. (1981). *Proc. Nat. Acad. Sci., U.S.A.*, **78**: 752-756.

Lengyel, J.A. and Calendar, R. (1974). *Virology*, **57**: 305-313.

Lengyel, J.A., Goldstein, R.N., Marsh, M., Sunshine, M.G. and Calendar, R. (1973). *Virology*, **53**: 1-23.

Lin, C-S. (1984). *Nucl. Acids. Res.*, **12**: 8667-8684.

Lindahl, G. (1969). *Virology*, **39**: 839-860.

Lindahl, G. (1971). *Virology*, **46**: 620-633.

Lindahl, G. (1974). *Mol. Gen. Genet.*, **128**: 249-260.

Lindahl, G., Sironi, G. and Calendar, R. (1970). *Proc. Nat. Acad. Sci., U.S.A.*, **66**: 587-594.

Lindqvist, B.H. (1974). *Proc. Nat. Acad. Sci., U.S.A.*, **71**: 2752-2755.

Lindqvist, B.H. and Bovre, K. (1972). *Virology*, **49**: 690-699.

Ljungquist, E. and Bertani, L.E. (1983). *Mol Gen. Genet.*, **192**: 87-94.

Ljungquist, E., Kockum, K. and Bertani, L.E. (1984). *Proc. Nat. Acad. Sci., U.S.A.*, **81**: 3988-3992.

Lowery, C. and Richardson, J.P. (1977). *J. Biol. Chem.*, **252**: 1381-1385.

Maeda, S., Ozawa, Y., Mizuno, T. and Mizushima, S. (1988). *J. Mol. Biol.*, **202**: 433-441.

Malan, T.P., Kolb, A., Buc, H. and McClure, W.R. (1984). *J. Mol. Biol.*, **180**: 881-909.

Maniatis, T., Fritsch E.F. and Sambrook, J. (1982). "Molecular cloning, a laboratory manual" (Cold Spring Harbor Lab., Cold Spring Harbor N.Y.).

Margolin, W. and Howe, M.M. (1986). *Nucl. Acids Res.*, **14**: 4881-4897.

Margolin, W., Rao, G. and Howe, M.M. (1989). *J. Bacteriol.*, **171**: 2003-2018.

Martin, S.A. and Moss, B. (1975). *J. Biol. Chem.*, **250**: 9322-9329.

Matsuyama, S. and Mizushima, S. (1987). *J. Mol. Biol.*, **195**: 847-853.

Maxam, A.M. and Gilbert, W. (1980). *Meths. Enzymol.*, **65**: 499-560.

Mazodier, P., Cossart, P., Giraud, E. and Gasser, F. (1985). *Nucl. Acids Res.*, **13**: 195-205.

McClure, W.R. (1985). *Ann. Rev. Biochem.*, **54**: 171-204.

McFall, E. (1987). In "Escherichia coli and Salmonella typhimuriam - Cellular and Molecular Biology" (Neidhardt, F.C. chief ed., Ingraham, J.L., Brooks Low, K., Magasanik, B., Schaechter, M. and Umbarger, H.E. eds.) Vol. 2. pp. 1520-1526. (American Society for Microbiology, Washington D.C.).

McKenney, K., Shimatake, H., Court, D., Schmeissner, U., Brady, C. and Rosenberg, M. (1981). In "Gene Amplification and analysis Vol. II" (Chirikjian, J.C. and Papas, T.S., eds.) pp. 384-415. (Elsevier - North Holland).

McKnight, S.L., Gavis, E.R. and Kingsbury, R. (1981). *Cell*, **25**: 385-398.

McLachlan, A.D., Staden, R. and Boswell, D.R. (1984). *Nucl. Acids Res.*, **12**: 9567-9575.

McMaster, G.K. and Carmichael, G.G. (1977). *Proc. Nat. Acad. Sci., U.S.A.*, **74**: 4835-4838.

Meek, D.W. and Hayward, R.S. (1984). *Nucl. Acids Res.*, **14**: 5813-5821.

Melchior, W.B. and von Hippel, P.H. (1973). *Proc. Nat. Acad. Sci., U.S.A.*, **70**: 298-302.

Messing, J. (1979). *Recombinant DNA technical bulletin*, **2**: 43-48.

Miller, J., McLachlan, A.D. and Klug, A. (1985). *EMBO J.*, **4**: 1609-1614.

Moitoso de Vargas, L., Pargellis, C.A., Hasan, H.M., Bushman, E.W. and Landy, A. (1988). *Cell*, **54**: 923-929.

Morgan, E.A. (1986). *J. Bacteriol.*, **168**: 1-5.

Morgan, W.D., Bear, D.G. and von Hippel, P.H. (1983a). *J. Biol. Chem.*, **258**: 9553-9564.

Morgan, W.D., Bear, D.G. and von Hippel, P.H. (1983b). *J. Biol. Chem.*, **258**: 9565-9574.

Morgan, W.D., Bear, D.G., Litchman, B.L. and von Hippel, P.H. (1985). *Nucl. Acids Res.*, **13**: 3739-3754.

Muller-Hill, B., Capo, L. and Gilbert, W. (1968). *Proc. Nat. Acad. Sci., U.S.A.*, **59**: 1259-1264.

Mulligan, M.E., Hawley, D.K., Entriken, R. and McClure, W.R. (1984). *Nucl. Acids Res.*, **12**: 789-800.

Murray, K. and Murray, N.E. (1973). *Nature, New Biol.*, **243**, 134-139.

Ni'Bhriain, N.N., Silvers, S. and Forster, T.J. (1983). *J. Bacteriol.*, *155*: 690-703.

Nicolleti, M. and Bertani, G. (1983). *Mol. Gen. Genet.*, **189**: 343-347.

Nikaido, H. (1961). *Biochim. Biophys. Acta*, **48**: 460-469.

Normark, S., Bergstrom, S., Edlund, T., Grundstrom, T., Jaurin, B., Lindberg, F.P. and Olsson, O. (1983). *Ann. Rev. Genet.*, **17**: 499-525.

Oakley, J.L. and Coleman, J.E. (1977). *Proc. Nat. Acad. Sci., U.S.A.*, **74**, 4266-4270.

O'Halloran,T.V., Frantz, B., Shin, M.K., Ralston, D.M. and Wright, J.G. (1989). *Cell*, **56**: 119-129.

Orosz, J.M. and Wetmur, J.G. (1979). *Biopolymers*, **16**: 1183-1199.

Osborne, A. (1986). Honours Thesis, University of Adelaide.

Pabo, C.O. and Sauer, R.T. (1984). *Ann. Rev. Biochem.*, **53**: 293-321.

Pedersen, S. (1984). *EMBO J.*, **3**: 2895-2898.

Pell, L. (1987). Honours Thesis, University of Adelaide.

Platt, T. (1986). *Ann. Rev. Biochem.*, **55**: 339-372.

Pritchard, M. and Egan, J.B. (1985). *EMBO J.*, **4**: 3599-3604.

Projan, S.J., Carleton, S. and Novick, R.P. (1983). *Plasmid*, **9**: 182-190.

Pruss, G.J. and Drlica, K. (1989). *Cell*, **56**: 521-523.

Pulitzer, J.F., Colombo, M. and Ciaramella, M. (1985). *J. Mol. Biol.*, **182**: 249-263.

Rabussay, D. (1983). In "Bacteriophage T4" (Matthews, C.K., Kutter, E.M., Mosig, G. and Berget, P.B., eds.) pp. 167-173. (Amer. Soc. for microbiol., Washington D.C.).

Raibaud, O, and Schwartz, M. (1984). *Ann. Rev. Genet.*, **18**: 173-206.

Raibaud, O., Gutierrez, C. and Schwartz, M. (1985). *J. Bacteriol.*, **161**: 1201-1208.

Raibaud, O., Vidal-Ingigliardi, D. and Richet, E. (1989). *J. Mol. Biol.*, **205**: 471-485.

Rao, R.N. and Rogers, S.G. (1979). *Gene*, **7**: 79-82.

Rasmussen, P.B., Gerdes, K. and Molin, S. (1987). *Mol. Gen. Genet.*, **209**: 122-128.

Reznikoff, W.S., Siegele, D.A., Cowing, D.W. and Gross, C.A. (1985). *Ann. Rev. Genet.*, **19**: 355-387.

Richardson, H. (1987). Ph.D. Thesis, University of Adelaide.

Richardson, H. and Egan, J.B. (1989). *J. Mol. Biol.*, **206**: 59-68.

Richardson, H., Puspurs, A. and Egan, J.B. (1989). *J. Mol. Biol.*, **206**: 251-255.

Richardson, J.P. (1982). *J. Biol. Chem.*, **257**: 5760-5766.

Richardson, J.P. and Macy, M.R. (1981). *Biochem.*, **20**: 1133-1139.

Richet,E. and Raibaud, O. (1989). *EMBO J.*, **8**: 981-987.

Riva, S., Cascino, A. and Geiduschek, E.P. (1970). *J. Mol. Biol.*, **54**: 85-102.

Robinson, M., Lilley, R., Little, S., Emtage, J.S., Yarronton, G., Stephens, R., Mullican, A., Eaton, M. and Humphreys, G. (1984). *Nuc. Acids Res.*, **12**: 6663-6671.

Rosenberg, M. and Court, D. (1979). *Ann. Rev. Genet.*, **13**: 319-353.

Rosenberg, M., Chepelinsky, A.B. and McKenney, K. (1983). *Science*, **222**: 734-739.

Rowland, G.C., Giffard, P.M. and Booth, I.R. (1985). *J. Bacteriol.*, **164**: 972-975.

Rybchin, V.N. (1984). *Gene*, **27**: 3-11.

Saha, S., Lundqvist, B. and Haggard-Ljungquist, E. (1987a). *EMBO J.*, **6**: 809-814.

Saha, S., Haggard-Ljungquist, E. and Nordstrom, K. (1987b). *EMBO J.*, **6**: 3191-3199.

Saha, S., Haggard-Ljungquist, E. and Nordstrom, K. (1989). *Proc. Nat. Acad. Sci., U.S.A.*, **86**: 3973-3977.

Saint, R.B. (1979). Ph. D. Thesis, University of Adelaide.

Saint, R.B. and Egan, J.B. (1979). *Mol. Gen. Genet.*, **171**: 79-89.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977). *Proc. Nat. Acad. Sci., U.S.A.*, **74**: 5463-5467.

Sauer,B. (1979). Ph.D. Thesis, University of California (Berkeley).

Sauer, B., Ow, D., Ling, L. and Calendar, R. (1981). *J. Mol. Biol.*, **145**: 29-46.

Sauer, B., Calendar, R., Ljungquist, E., Six, E. and Sunshine, M.G. (1982). *Virology*, **116**: 523-534.

Scheritz, R.W., Otwinowski, Z., Joachimiak, A., Lawson, C.L. and Sigler, P.B. (1985). *Nature*, **317**: 782-786.

Schnos, M. and Inman, R.B. (1971). *J. Mol. Biol.*, **55**: 31-38.

Schreier, P.H. and Cortese, R. (1979). *J. Mol. Biol.*, **129**: 169-172.

Sharp, J.A. and Platt, T. (1984). *J. Biol. Chem.*, **259**: 2268-2273.

Shepherd, J.C.W. (1981). *Proc. Nat. Acad. Sci., U.S.A.*, **78**: 1596-1600.

Shih, M.C. and Gussin, G.N. (1984). *J. Mol. Biol.*, **172**: 489-506.

Shine, J. and Dalgarno, L. (1974). *Proc. Nat. Acad. Sci., U.S.A.*, **71**: 1342-1346.

Shulman, M.J., Mizuuchi, K. and Gottesman, M.M. (1976). *Virology*, **72**: 13-22.

Siebenlist, U. and Gilbert, N. (1980). *Proc. Nat. Acad. Sci., U.S.A.*, **77**: 122-126.

Sivaprasad, A.V., Jarvinen, R. and Egan, J.B. (1990). *J. Mol. Biol.*, in press.

Six, E.W. (1975). *Virology*, **67**: 249-263.

Six, E.W. and Lindqvist, B.H. (1978). *Virology*, **87**: 217-230.

Skalka, A. and Hanson, P. (1972) *J. Virol.*, **9**: 583-593.

Souza, L., Calendar, R., Six, E.W. and Lindqvist, B.H. (1977). *Virology*, **81**: 81-90.

Staden, R. (1980). *Nucl. Acids Res.*, **8**: 3673-3694.

Staden, R. (1984a). *Nucl. Acids Res.*, **12**: 505-519.

Staden, R. (1984b). *Nucl. Acids Res.*, **12**: 551-567.

Steger, G., Gross, H., Sanger, H.C., Riesner, D., Hofman, H. and Randles, J.W. (1984). *J. Biomol. Structure and Dyn.*, **2**: 543-571.

Stewart. C. (1988). In "The Bacteriophages",.(R. Calendar, ed.). Vol. 1. pp. 477-515 (Plenum Publishing).

Stormo, G.D., Schneider, T.D. and Gold, L.M. (1982). *Nucl. Acids Res.*, **10**: 2971-2996.

Sunshine, M.G. and Sauer, B. (1975). *Proc. Nat. Acad. Sci., U.S.A.*, **72**: 2770-2774.

Sunshine, M.G., Thorn, M., Gibbs, W. and Calendar, R. (1971). *Virology*, **46**: 691-702.

Surin, B., Jans, D., Fimmel, A., Shaw, D. and Cox, G. (1984). *J. Bacteriol.*, **157**: 772-778.

Szybalski, E.H. and Szybalski, W. (1979). *Gene*, **7**: 217-270.

Szybalski, W., Drahos, D., Luk, K.-C. and Somasekhar, G. (1983). In "Microbiology-1983" (Schlessinger, D., ed.) pp. 35-38. (Amer. Soc. for Microbiol., Washington D.C.).

Tanaka, S. and Matsushiro, A. (1985). *Gene*, **38**: 119-129.

Teo, I., Sedgwick, B., Kilpatrick, M.W., McCarthy, T.V. and Lindahl, T. (1985). *Cell*, **45**: 315-324.

Thomas, P.S. (1980). *Proc. Nat. Acad. Sci., U.S.A.*, **77**: 5201-5205.

Thomas, P.S. (1983). *Meths. Enzymol.*, **100**: 255-266.

Tinoco Jr, I., Borer, P.N., Dengler, B., Levine, M.D., Uhlenbeck, O.C., Crothers, D.M. and Gralla, J. (1973). *Nature*, **246**: 40-41.

Travers, A. and Klug, A. (1987). *Nature*, **327**: 280-281.

Ullrich, A., Gray, A., Wood, W.I., Hayflick, J. and Seeburg, P.H. (1984). *DNA*, **3**: 387-392.

Uzan, M., Leautey, J., d'Aubenton-Carafa, V. and Brody, E. (1983). *EMBO J.*, **2**: 1207-1212.

van Bokkelen, G.B., Dale, E.C., Halling, C. and Calendar, R. (1990). *J. Bacteriol.*, submitted.

von Hippel, P.H., Bear, D.G., Morgan, W.D. and McSwiggen, J.A. (1984). *Ann. Rev. Biochem.*, **53**: 389-446.

Wang, J.C. (1967). *J. Mol. Biol.*, **28**: 403-411.

Wang, J.C. and Schwartz, H. (1967). *Biopolymers*, **5**: 953-966.

Wang, J.C., Martin, K.V. and Calendar, R. (1973). *Biochem.*, **12**: 2119-2123.

Weigle, J. (1966). *Proc. Nat. Acad. Sci., U.S.A.*, **55**: 1462-1466.

Westoo, A. and Ljungquist, E. (1979). *Mol. Gen. Genet.*, **171**: 91-102.

Westoo, A. and Ljungquist, E. (1980). *Mol. Gen. Genet.*, **178**: 101-109.

Wilson, D.B. and Hogness, D.S. (1966). *Meths. Enzymol.*, **8**: 229-240.

Wong, S.L. and Doi, R.H. (1982). *J. Biol. Chem.*, **257**: 11932-11936.

Wood, W.I., Gitschier, J., Lasky, L.A. and Lawn, R.M. (1985). *Proc. Nat. Acad. Sci., U.S.A.*, **82**: 1585-1588.

Woods, W.H. (1972). Ph. D. Thesis, University of Adelaide.

Woods, W.H. and Egan, J.B. (1974). *J. Virol.*, **14**: 1349-1356.

Wulff, D.L. and Rosenberg, M. (1983). In "Lambda II" (Hendrix, R.W., Roberts, J.W., Stahl, F.W. and Weisberg, R.A., eds.) pp. 53-73. (Cold Spring Harbor Lab., Cold Spring Harbor, N.Y.).

Yanisch-Perron, C., Vieira, J. and Messing, J. (1985). *Gene*, **33**: 103-119.

Younghusband, H.B. and Inman, R.B. (1974). *Virology*, **62**: 530-538.

Younghusband, H.B., Egan, J.B. and Inman, R.B. (1975). *Mol. Gen. Genet.*, **140**: 101-110.

Zehring, W.A. and Rothman-Denes, L.B. (1983). *J. Biol. Chem.*, **258**: 8074-8080.

Zehring, W.A., Falco, S.C., Malone, C. and Rothman-Denes, L.B. (1983). *Virology*, **126**: 678-687.

Zoller, M.J. and Smith, M. (1982). *Meth. Enzymol.*, **100**: 468-500.

Zoller, M.J. and Smith, M. (1984). *DNA*, **3**: 479-488.