CHARACTERISATION OF THE *IN VITRO* TRANSCRIPTION PATTERN OF THE TEMPERATE COLIPHAGE 186

A Thesis submitted for

the degree of

Doctor of Philosophy

at the

University of Adelaide

by

Melanie April Pritchard, B.Sc(Hons)

Department of Biochemistry

June, 1984.

FOR MY PARENTS

C O N T E N T S

SUMMARY

The *in vitro* transcription pattern of a prophage can be considered the equivalent of the *in vivo* transcription pattern immediately after infection before phage protein synthesis. That is, the *in vitro* pattern should be indicative of the early transcripts. This thesis describes the establishment of an *in vitro* transcription system for 186 and the characterisation of the transcripts.

186 DNA was transcribed with *E. coli* RNA polymerase and four major transcripts were observed on polyacrylamide gels:

Band 1 - 1540 bases

Band 2 - 1450 bases

Band 3 - 590 bases

Band 4 - 290 bases

The direction of transcription was determined and preliminary mapping of these transcripts was accomplished by hybridizing each transcript to restriction digests of 186 DNA Southern blotted on to nitrocellulose. Then plasmid clones containing appropriate 186 restriction fragments were transcribed *in vitro* which identified promoter containing fragments and supported the hybridization data. Run-off transcripts were analysed, which achieved a more precise knowledge of the location of the 5'-ends of the transcripts. When the DNA sequence became available the 5'-ends of the transcripts were precisely mapped and their accompanying control features identified on the DNA sequence.

Bands 1 and 3 share a promoter, *p*95, which has been mapped to the far right hand end of the 186 chromosome

at 95%. Their function is speculative and is discussed in the context of features found in the DNA near the promoter.

The 5'-end of Band 2 was sequenced and found to represent transcription from a promoter, *p*R, which maps within the first of three computer predicted operators which are presumably involved in the lysis/lysogeny decision by binding repressor.

Band 4 was identified as the B gene transcript. The B gene product is required for late gene expression. A transcript, the same size as the B gene transcript, was identified among the P2 *in vitro* transcripts and proved to be the *ogr* gene transcript, the functional equivalent of the 186 B gene.

The *in vitro* transcription pattern, represented the early transcripts of 186. The transcripts are discussed in relation to the DNA sequence and the biology of the regions from which they originate.

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 no material previously published or written by another person except where due reference is made in the text.

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

M.A. PRITCHARD

ACKNOWLEDGEMENTS

I would like to express my gratitude to the following people for their assistance during the course of my work:

Dr. J.B. Egan for direction, encouragement and for sharing his philosophy.

Professor W.H. Elliott for permission to work in the Department of Biochemistry.

To all members of Lab 19, past and present, not only for creating a friendly, enjoyable atmosphere but also for their critical comments.

Lesley Crocker for the photographic plates that appear in this thesis.

Jeanette Brooker for typing this thesis.

Alison for her help with the preparation of the Introduction.

Sylvia Francis and Mary Milnthorpe for excellent technical assistance.

Special thanks are due to Wolfram Zillig who came to the rescue with RNA polymerase.

ABBREVIATIONS

PEG	-	polyethylene glycol
c.p.m.	-	counts per minute
EDTA	-	ethylenediamine tetraacetate
BSA	-	bovine serum albumin
DTT	-	dithiothreitol
IPTG	-	isopropylthiogalactoside
BCIG	-	5-bromo-4-chloro-3-indoyl-galactoside
SDS	-	sodium dodecyl sulphate
TCA	-	trichloroacetic acid
kb	-	kilobase (1,000 base-pairs)
Tris	-	Tris(hydroxymethyl)aminomethane

CHAPTER 1

GENERAL INTRODUCTION

1. Phage 186

Phage 186 is a temperate bacteriophage of *E. coli* with a double-stranded, non-permuted DNA molecule of molecular weight 19.7 x 10^6 daltons (Wang, 1967). The DNA possesses complementary cohesive ends, nineteen nucleotides long.

A linear genetic map of 186 has been constructed on the basis of two- and three-factor crosses (Figure 1.1) (Hocking and Egan, 1982a). Twenty-two genes essential for lytic growth have been identified, as well as a number of non-essential genes, such as *c*I and *c*II which are involved in the establishment and maintenance of lysogeny, and *int* which is required for lysogenisation of the host. 186 also has a function, defined by a gene *dhr*, which depresses host DNA synthesis upon infection (H. Richardson, pers. commun.).

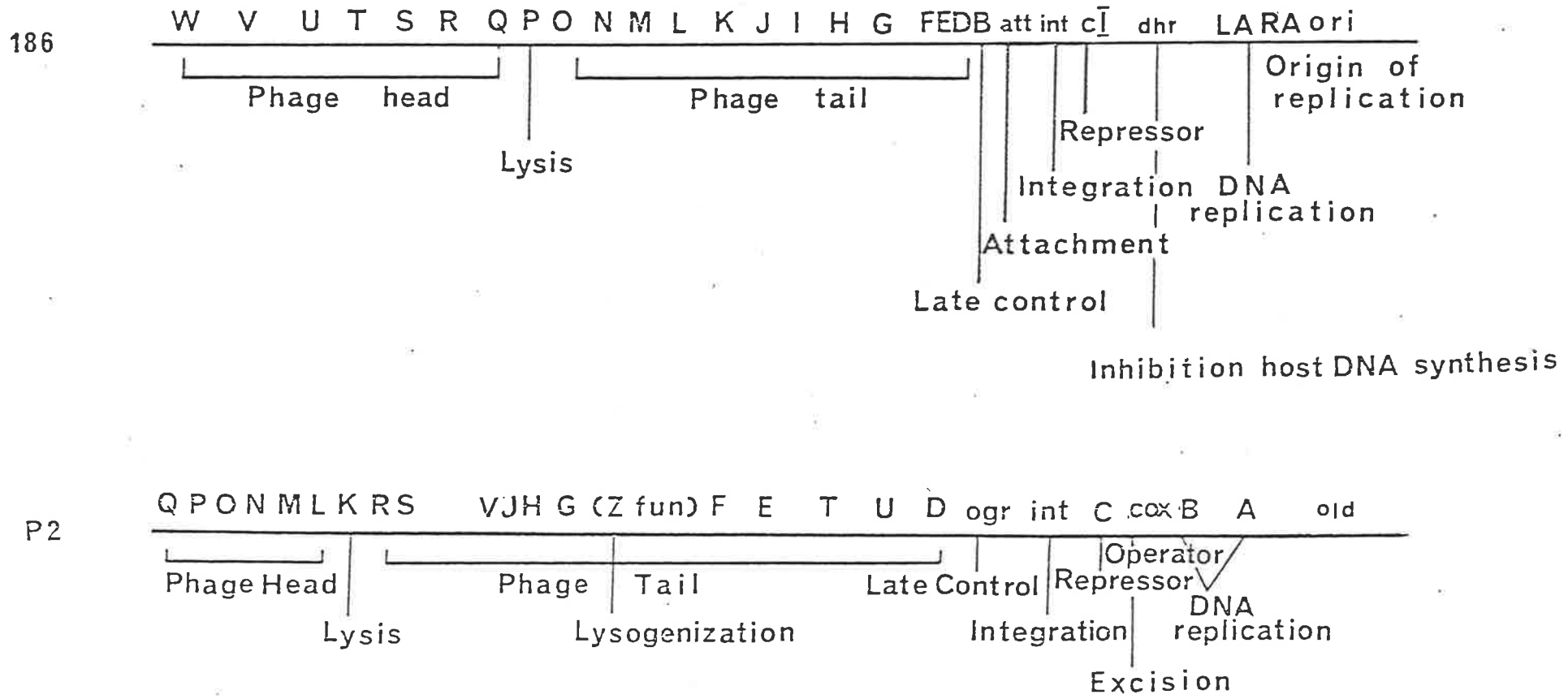
Genes with related functions are clustered on the 186 map. To the left of the *att* site is the late control gene, B. The left most end of the chromosome contains seven genes involved in head morphogenesis, which are separated from a group of eleven genes involved in formation of the phage tail, by genes O and P. Gene P is the lysis gene and the function of O is not clear. Sixteen of these genes are organized into four transcription units based on the polar effects of various amber mutants: groups D to G, H to J, K to M and Q to V. The remaining six essential genes, A, B, N, O, P and W do not belong to polarity groups (Hocking and Egan, 1982b).

Replication of 186 DNA requires the cis-acting product of the A gene (Hocking and Egan, 1982c). A.V. Sivaprasad in our laboratory, has shown by sequence analysis that the A

FIGURE 1.1

GENETIC MAPS OF 186 AND P2

The relative location of the known genes of 186 and P2 are shown in the figure. The gene functions are listed below the map.


186
W V U T S R Q P O N M L K J I H G F E D B att int cI dhr LA RA ori
Phage head
Phage tail
Origin of replication
Repressor
Lysis
Integration
DNA replication
Attachment
Late control
Inhibition host DNA synthesis
P2
Q P O N M L K R S V J H G (Z fun) F E T U D ogr int C cox B A old
Phage Head
Phage
Tail
Late Control
Operator
Repressor
DNA replication
Lysis
Lysogenization
Integration
Excision

gene region codes for two proteins designated LA (left) and RA (right), both defined by amber mutations thought previously to belong to separate alleles of the A gene. Preliminary studies reveal that RA is needed directly for replication whilst LA is required for the efficient translation of RA, i.e., has an indirect role in replication.

A *c*II gene has been defined by an amber mutation which is complemented by mutants in the *c*I repressor gene. The map location of *c*II is not known, although it has been shown to lie to the right of *att* (Hocking and Egan, 1982c).

2. Comparison of P2 and 186

P2 is perhaps the best studied of the non-inducible phage. P2 and 186 show many similarities both in their molecular biology and the morphology of the phage particles.

P2 and 186 have cohesive ends which differ in only two of nineteen residues (Padmanabhan and Wu, 1972; Murray and Murray, 1973). Electron microscope heteroduplex studies have shown that the two phages share homologous sequences, particularly in that part of the genome coding for phage structural components (Skalka and Hanson, 1972; Younghusband and Inman, 1974). The control sequences located on the right third of the genome show no homology under the electron microscope. For comparison of gene function and map location, the genetic maps of 186 and P2 are presented in Figure 1.1.

Replication of P2 DNA requires the products of the early genes A and B (Lindahl, 1971). The A gene of 186 and the A gene of P2 are likely to be site-specific endonucleases which cause a single-strand nick at the origin

of replication (Chattoraj and Inman, 1973; Geisselsoder, 1976). The B gene of P2 is involved in lagging strand synthesis (Funnell and Inman, 1983). There is no evidence in 186 for a gene with analogous function to that of the P2 B gene.

There is no evidence in P2 for a *c*II gene, nor is there a counterpart to *dhr* as P2 does not interfere with host DNA synthesis (Hooper *et al.*, 1981).

λ is unable to plate on a P2 lysogen and this interference is abolished by a mutation in the P2 *old* gene (Lindahl *et al.*, 1970). 186 does not have a gene equivalent to P2 *old* as λ forms plaques on 186 lysogens.

The most striking difference between P2 and 186 is inducibility of the prophage. 186 shares the induction properties of λ, i.e., UV irradiation, mitomycin C and nalidixic acid all cause induction of the 186 prophage (Woods and Egan, 1974). P2 is non-inducible and non-excisable (G. Bertani, 1968) under these conditions.

The non-inducibility of P2 has been attributed to two factors; firstly, P2 repressor is not inactivated by UV irradiation (G. Bertani, 1968) and second, even when repression is lifted by subjecting a prophage with a temperature sensitive repressor to non-permissive temperatures, the prophage fails to excise (L.E. Bertani, 1968).

To explain the non-excisability of P2, the "split-operon" model was proposed by L.E. Bertani (1970). This model relied on P2 *int* transcription being from left to right relative to the genetic map, such that upon integration into the host chromosome, the *int* gene is physically separated from it's promoter which is situated on the opposite side of

the *att* site, hence no *int* protein can be made by the prophage (see Figure 1.2). However, Ljungquist (pers. commun.) sequenced an *int* amber mutant showing that *int* transcription is from right to left which is not compatable with the split-operon theory.

Our laboratory is ultimately interested in the further description of prophage induction and excision. A comparison of 186 with it's non-inducible relative P2 should provide an insight into the reason for this difference in excisability. It should also provide a greater understanding of the control of phage integration and excision, a process that has an analogy in eukaryotic cells with the integration and/or excision of RNA tumor viruses.

3. Interaction of phage P4 with P2 and 186

Satellite phage P4 has a linear, non-permuted, double-stranded DNA molecule (6.8×10^6 daltons), which terminates in single-stranded complementary ends 19 nucleotides long. These complementary ends are identical to those of P2.

Phage P4 is able to replicate it's own DNA and lysogenise a host *E. coli* bacterium in the absence of a helper phage, (Lindquist and Six, 1971); or it may maintain itself as a plasmid (Shore *et al.*, 1978; Goldstein *et al.*, 1982).

For lytic development, P4 requires all the known head, tail and lysis genes of a helper phage such as P2 (Six, 1975). P4 can utilise the genes of a co-infecting helper phage or can stimulate expression of late genes from a prophage helper (Six and Klug, 1973). P4 is able to derepress a helper P2 prophage (Six and Lindquist, 1978) and thereby utilise late genes. This derepression results in *in situ*

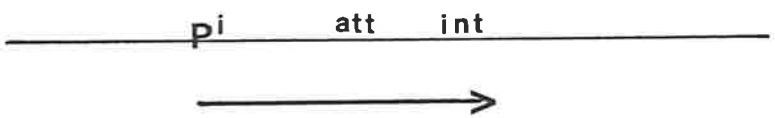
FIGURE 1.2

THE SPLIT-OPERON MODEL

A. The arrangement of the *int* operon in the vegetative phage is shown. *p*i represents the promoter for the *int* gene. The model relies on the *att* site being located between *p*i and *int* and also rightward transcription of *int* indicated by the arrow.

B. The arrangement of the *int* operon upon integration into the host chromosome is shown. The left (L *att*) and right (R *att*) arms of the *att* site are shown. The *int* gene is separated from its promoter.

A



B

unidirectional replication of the integrated helper without excision (Geisselsoder *et al.*, 1981). A P4 mutant exists which is defective in derepression and can not grow on a P2 lysogen; therefore it is likely that P4 requires a repressed P2 function (Geisselsoder *et al.*, 1981).

During lytic co-infection by P4 and P2, P4 interferes with morphogenesis of P2 by directing the assembly of small capsids capable of packaging P4 but unable to package the larger genome of the P2 helper phage (Diana *et al.*, 1978; Geisselsoder *et al.*, 1978; Shore *et al.*, 1978).

As well as causing expression of P2 late genes under their normal mode of control, P4 activates expression of P2 late genes by "*transactivation*" (Souza *et al.*, 1977), which is defined by the ability of P4 to turn on P2 late gene expression in the absence of two early P2 genes, A and B. These genes are normally needed for P2 late gene transcription and for P2 DNA replication (Six, 1975; Harris and Calendar, 1978). The *transactivation* gene of P4 is called δ.

Like P2, 186 can act as a helper phage during a P4/186 mixed infection. P4 is able to *transactivate* 186 late gene expression in the absence of 186 A and B genes. Derepression of a 186 prophage (e.g., by temperature induction) is a prerequisite for a P4 lytic infection (Sauer *et al.*, 1982). It seems that P4 requires a 186 function which is under the negative control of the *c*I repressor protein, but in contrast to P2, P4 cannot derepress a 186 prophage.

4. Comparison between 186-B gene, P2-*ogr* gene and P4-δ gene.

The 186 B, P2 *ogr* and P4 δ genes are functionally similar; they all turn on late gene expression. The *ogr* gene of P2 is also similar to the 186 B gene in it's relative position on the genetic map (Figure 1.1). Hocking and Egan (1982d) have shown that 186/P2 hybrid phages which lack the P2 *ogr* gene rely on the 186 B gene for expression of P2 morphogenic genes. Thus the 186 B gene is able to replace the requirement for *ogr* to turn on P2 late gene expression.

Both 186 B^- and P2 ogr^- phages can support a P4 δ^+ infection, implying that P4 δ gene substitues for B and *ogr* in late turn on (Sauer *et al.*, 1982). However, P4 δ^- mutants are unable to *transactivate* 186 A^- and P2 A^- phage. Sauer *et al.* (1982) explain this phenomenon by considering the following observations. By complementation Sauer *et al.* (1982) concluded that the P2 *ogr* gene can only be expressed when the A gene is functional. Finnegan and Egan (1981) concluded that 186 B gene transcription is dependent upon A gene function. Thus in 186 A^- and P2 A^- mutants, B and *ogr* genes are not expressed, hence, there can be no late gene expression.

5. Transcription Controls

λ and the P2/186 family are the best studied of the temperate coliphages. λ is the paradigm for phage transcriptional control. Some aspects of P2 transcriptional control have been studied and these differ in some respects

from the control of λ.

λ infection results in a sequential expression of phage genes that is determined by the sequential synthesis of regulatory proteins. Immediately after injection of λ DNA into a cell, host RNA polymerase binds at three promoters (*p*R', *p*R, and *p*L) to produce three transcripts (Figure 1.3). The *p*R' initiated transcript is extended by the antiterminator protein Q at *t*R' to enable transcription of late genes (Grayhack and Roberts, 1982). λN protein is encoded by the *p*L transcript and Cro is encoded by *p*R.

N protein, a positive regulator, acts as an antiterminator to allow expression of late genes past the terminators *t*L, *t*R1 and *t*R2 (Friedman *et al.*, 1976a; Adhya *et al.*, 1974; Franklin, 1974). In order for N protein to promote antitermination of transcription, transcription must pass through a region containing a N utilization (*nut*) site (Salstrom and Szybalski, 1978). A number of host functions, *nus*A, *nus*B, *nus*C and *nus*D are required for λN activity (Greenblatt *et al.*, 1980; Friedman and Baron, 1974; Friedman *et al.*, 1976b; Simon *et al.*, 1979; Friedman *et al.*, 1981).

Cro, the protein encoded by the *p*R transcript is a negative regulator which binds to *o*R to block premature synthesis of *c*I. Cro represses early gene transcription from *p*R and *p*L midway through infection (Takeda, 1979; Echols *et al.*, 1973; Reichardt, 1975). λ*cro*$^-$ mutants are unable to grow lytically (Folkmanis *et al.*, 1977). Hence, depression of early protein synthesis is necessary for successful lytic growth.

Negative control of the genes required for the lytic

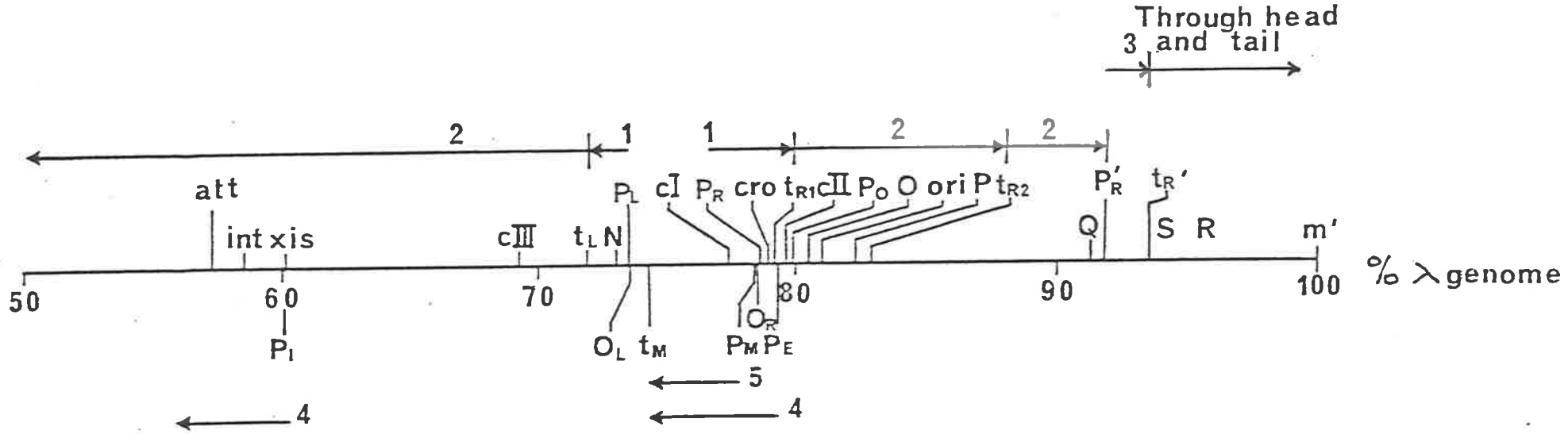
FIGURE 1.3

THE TRANSCRIPTION MAP OF λ

The transcripts essential to lytic development are depicted above the map which shows the control genes and the sites at which they act.

Early transcripts (1) of λ cover the genes N and *cro* and terminate at the *t*L and *t*R1 respectively. The N protein allows elongation of both leftward and rightward early transcripts (2) to cover *c*III, the recombination and integration genes to the left of N, and *c*II, the replication genes and Q which are transcribed rightward. The late gene transcripts (3) are antiterminated by Q at *t*R'.

*c*II and *c*III act together to stimulate transcripts required to establish lysogeny (4) which are depicted below the map. The major transcript made in the prophage state (5) covers the *c*I gene and is maintained by the *c*I protein itself.


Through head
3 and tail
2
1
1
2
2
att
int xis
cIII
tL N
PL
cI PR cro tR1 cII PO O ori P tR2
P'R
tR'
Q
S R
m'
50
60
70
80
90
100
% λ genome
PI
OL tM
OR
PM PE
5
4
4

cycle is mediated by cI, the phage repressor. The repressor binding sites *o*L and *o*R each contain three adjacent repressor binding sites (Figure 1.4), and although these six sites have similar nucleotide sequences, the affinities of cI for these sites differs (Maniatis *et al.*, 1975; Ptashne *et al.*, 1976). The site with the highest affinity for cI is *o*R1, then *o*R2 and *o*R3 (Johnson *et al.*, 1979); likewise at *o*L, the sequence of binding is *o*L1, *o*L2 and *o*L3. The relative affinities of Cro at *o*R are opposite that of cI; *o*R3 > *o*R2 ≃ *o*R1 (Johnson *et al.*, 1978).

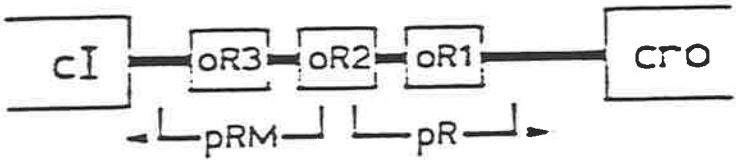
*o*R also controls transcription initiation from another promoter *prm* (the promoter for repressor maintenance). Because of the different binding affinities of Cro and cI to the operator sites their effects on transcription from *p*R and *prm* are quite different (see Herskowitz and Hagen for review, also Gussin *et al.*, 1983). These effects have been studied *in vitro* by monitoring transcription from *p*R and *prm* after addition of varying amounts of purified cI (Meyer *et al.*, 1975; Walz *et al.*, 1976) or Cro (Johnson *et al.*, 1978; Takada, 1979). The binding of cI to *o*R1 and *o*R2 activates transcription from *prm* while repressing transcription from *p*R. *prm* is repressed when cI is bound to *o*R1, *o*R2 and *o*R3 (Mauer *et al.*, 1980). *In vitro* transcription from *prm* (the low level seen in the absence of cI) is repressed by low concentrations of Cro, which correspond to occupancy of *o*R3. At higher concentrations, Cro represses transcription from *p*R by filling *o*R2 and *o*R3.

P2 is the best known phage from another major group, the non-inducible phages. Transcription studies on P2 have

FIGURE 1.4

THE STRUCTURE OF λ *o*R REGION

The region of the λ genome between *c*I and *cro* is schematically drawn with the relative positions of the repressor-binding sites, *o*R1, *o*R2 and *o*R3, and the promoters, *p*RM and *p*R.


cI
oR3
oR2
oR1
cro
pRM
pR

been confined to *in vivo* studies whereby early transcription is found to originate exclusively from the right half of the phage DNA molecule. Later in infection, transcription of genes situated on the left half of the molecule is progressively induced (Geisselsoder *et al.*, 1973). Geisselsoder *et al.* (1973) also found that P2 genes A and B are both necessary to activate late gene transcription, since mutants in both fail to produce significant amounts of left half specific RNA.

The P2 specific immunity repressor, encoded by the C gene, regulates expression of the P2 early operon. The early operon, including the gene *cox*, is located to the right of the repressor gene (Figure 1.1).

So far, there is no evidence in P2 for genes comparable to the λ genes N, *c*II, *c*III or *cro*.

An investigation of the control mechanisms of 186, which is classed as a member of the P2 family by morphology and yet shares the induction properties of the lambdoid phages, will contribute to the understanding of the relatedness of phage and perhaps shed some light on their evolution. If 186 differs from λ in it's transcriptional control, then a contribution will be made to the field of control of gene expression which is relevant to both prokaryotes and eukaryotes.

6. Transcription studies

Transcription studies are important in the beginning of any gene control study and RNA assays are the most direct means for studying regulation of gene expression.

Coliphage have three definite stages of lytic development:

1) early
2) middle
3) late

The separation of early and middle stages can be inferred from measurement of RNA synthesis under conditions in which protein synthesis is inhibited, or *in vitro*.

For example, in λ, initial evidence for the role of the antiterminator, N, came from direct studies of transcripts *in vivo* and *in vitro*. The initial transcripts are delimited *in vitro* if the termination factor rho is added (Roberts, 1970) or *in vivo* in the absence of N protein (Echols, 1971; Kourilsky *et al.*, 1971). Transcription is extended rather than reinitiated in the presence of N or absence of rho (Portier *et al.*, 1972). Thus N protein acts *in vivo* to prevent termination.

Of immediate interest in our laboratory is the characterisation of 186 initial early transcription. Essentially, this is transcription in the absence of any phage protein synthesis. To accomplish this characterisation two approaches may be taken;

1) studies *in vivo* in the presence of chloramphenicol
2) studies *in vitro* in the absence of any translation machinery.

Finnegan and Egan (1981) followed the first approach and hybridized the RNA made upon induction of a 186 prophage in chloramphenicol treated *E. coli* to cloned 186 restriction fragments. The only detectable transcription was a low level

hybridizing to pEC16 and a higher level which hybridized to pEC35 (Figure 1.5). Since pEC35 was considered to be diagnostic for B gene transcription, and Finnegan and Egan had concluded from transcription analysis after induction of an A*am* prophage, that B gene transcription was dependent upon a functional A gene, the B gene was not expected to be active in chloramphenicol treated cells. Thus, it was presumed that the transcriptional activity of pEC35 originated from the interval to the right of Δ2 (74.2% to 77.3%). Finnegan and Egan (1981) therefore, postulated the existence of a primary control gene X to map in the region around 76% on the chromosome, whose protein product is necessary for any further phage transcription.

To obtain maximum information from RNA assays, specific RNA molecules must be studied, and thus transcriptional units and their individual regulation may be defined. The following are some of the methods used in this thesis to study RNA's.

a) Hybridization

DNA/RNA hybridization is a powerful technique for the analysis of phage-directed RNA transcripts. With the advent of restriction endonucleases, radiolabelled RNA can be hybridized to DNA restriction fragments which have been separated by gel electrophoresis, denatured and transferred onto nitrocellulose (Southern, 1975). With a restriction map of defined genetic content, a great deal of information may be gained about the origin of individual RNA species.

b) Transcription of cloned restriction fragments

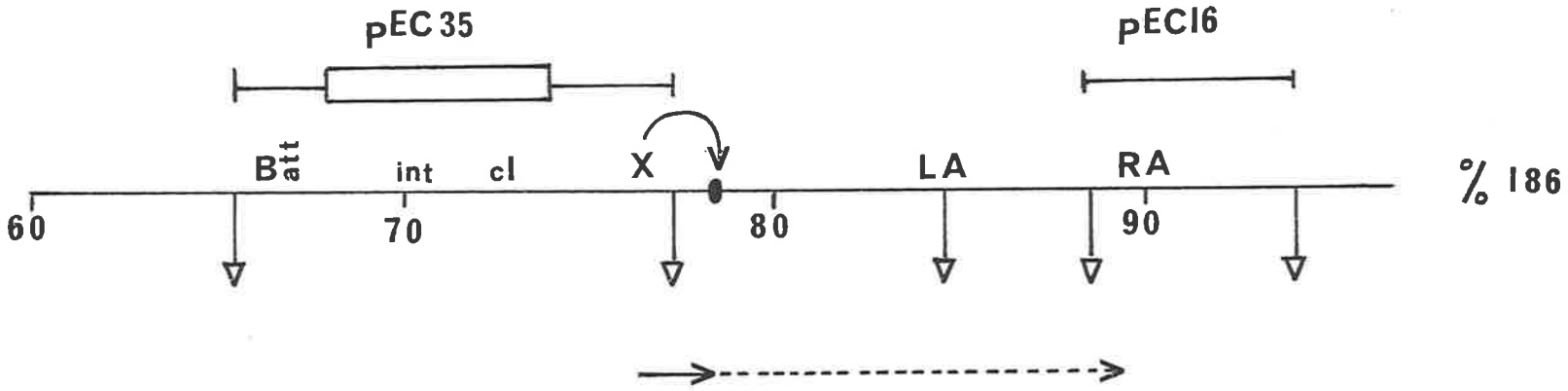
Once RNA's have been assigned to regions on the chromo-

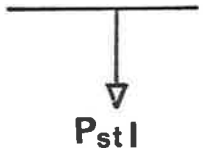
FIGURE 1.5

GENE X FUNCTION

A schematic representation of the position and function of gene X is shown. Gene X is transcribed and its product acts to extend rightward transcription through the replication genes.

The position of the Δ2 deletion in pEC35 is indicated by a box.


pEC35
pECI6
B_{att}
int
cI
X
LA
RA
% 186
60
70
80
90


P_{st} I

some by hybridization, then these regions may be cloned into plasmid vectors and used as templates in an *in vitro* transcription reaction. This technique allows the assignment of promoters as only DNA fragments with residing promoters will be transcriptionally active. "Run-off" transcription, whereby cloned fragments are digested with a restriction enzyme having an internal site, and then used as templates for *in vitro* transcription, can determine the direction of transcription, and also from the size of the run-off RNA, predict the location of promoters.

c) Sequencing

The most definitive approach to RNA analysis lies in the determination of the nucleotide sequence. The more recent developments in RNA sequencing methods have focused on the use of *in vitro* [^{32}P] end-labelling of RNA (Silberklang *et al.*, 1979). RNA's may be labelled at their 5'-termini by the incorporation into an *in vitro* transcription reaction of an initiating triphosphate labelled at the γ position with [^{32}P]. Alternatively, T4 polynucleotide kinase along with [γ-^{32}P]ATP may be used, following the removal of any pre-existing 5'-terminal phosphates (Chaconas, 1980). Once the RNA's are labelled, they may be sequenced by the enzymatic cleavage method of Donis-Keller (1977).

Another method of obtaining an RNA sequence is by extending a labelled DNA primer, which is hybridized to an internal region of the RNA, with reverse transcriptase in the presence of dideoxynucleoside triphosphates.

The RNA sequences of amber mutants may be obtained by transcribing a DNA carrying the mutation. This not only

allows confirmation of the gene content of an RNA species, but also the elucidation of the reading frame of the translation product along with the identification of the protein initiation and termination signals.

7. Aim of this work

A sound knowledge of the molecular biology of 186 is a prerequisite for any experiments on the control of it's gene expression.

A genetic map of 186 was constructed by Hocking (1982a) and a physical map by Finnegan and Egan (1979). Finnegan & Egan, (1981) investigated the temporal appearance of 186 messenger RNA by hybridizing RNA from 186 infected cells to 186 DNA restriction fragments. My aim was to provide a detailed immediate early transcription map of 186 by *in vitro* transcription studies employing the methods of RNA analysis just alluded to.

CHAPTER 2

MATERIALS AND METHODS

A. MATERIALS

All bacterial and bacteriophage strains used in this work are described below.

1. Bacterial strains

The bacterial strains used in this study are described in Table 2.1.

2. Bacteriophage strains

The bacteriophage used in this study are described below.

a) Phage 186 strains

186*c*I*ts*: a heat inducible mutant of 186 (Baldwin *et al.*, 1966)

186*c*I*ts*B*am*17: has an amber mutation in the B gene, allele 17 (Hocking and Egan, 1982b)

186*c*I*ts*B*am*57: has an amber mutation in the B gene, allele 57 (Hocking and Egan, 1982b).

186Δ1: phage with a deletion from 67.9% to 74.2%, phenotypically $int^{-}c\mathrm{I}^{-}$ (Dharmarajah, 1975)

b) Phage P2 strains

P2: wild-type from C1055 (P2) (Geisselsoder *et al.*, 1970)

P2*vir*22*ogr*52: *vir*22 has a 5.0% deletion from 72.2% to 77.2% and a 0.5% insertion, *ogr*52 is a mutation in the *ogr* gene permitting growth on an *E. coli gro* mutant (B. Sauer, pers. commun).

c) Other phage strains

Hy5: a P2/186 hybrid isolated in this laboratory (Bradley *et al.*, 1975)

TABLE 2.1 BACTERIAL STRAINS

Collection number	Strain	Genotype	Origin or reference
E. coli K12			
E536	W3350	Su^- F^- *galK galT strR*	Campbell (1965)
E508	C600	Su^+ F^- *thr leu thi lac*Y *ton*A *sup*E	Appleyard (1954)
E574	C600	(186*c*I*ts*)	This laboratory, from 186*c*I*tsp* described by Baldwin *et al.*, (1966)
E1017	E508	(186*c*I*ts*B*am*17)	This laboratory, from 186*c*I*tsp* after uv mutagenesis by Hocking and Egan (1982b)
E1057	E508	(186*c*I*ts*B*am*57)	This laboratory, after NNG mutagenesis of 186*c*I*tsp* by Hocking and Egan, (1982b)
E402	CGSC 5122 (Hy 5)	F^+ *thi thy uvr*A	Rupp *et al.* (1971)
E2106	E536	carrying pBR322	Bolivar *et al.* (1977)
E2107	GM31 (pBR325)	*his thr leu* B1 *dem*	Bolivar *et al.* (1977)
E2121	SR71	N100 *rec*A *gal* Su^- carrying pKC7	Rao and Rogers (1979)

E2234	E536	carrying pEC35	Finnegan and Egan, (1979)
E0605	JM101	Δ*lac pro sup*E *thi* F^+ *tra*D36 *pro*AB *lac*iq 2ΔM15	Messing *et al*, (1981a)
E2241	E699 (186*c*I^+)	carrying pEC400	This laboratory, H. Richardson
E. coli C			
E814	C1792	F^- *arg his trp* SuIII *str*R	Sunshine *et al*. (1971)
E273	C2121	*cia str*R *gro*109	Sunshine *et al*. (1971)

M13mp9: cloning vector with unique sites *Hin*dIII-*Pst*I-*Sal*I-*Bam*HI-*Sma*I-*Eco*RI (Messing, 1981b)

3. Plasmids

186 restriction fragments were cloned into pBR322 or pBR325 and pKC7, derivatives of pBR322.

pEC35: 186 *Pst*I fragment from 65.5-77.3% with a deletion from 67.9-74.2% cloned into pBR322 (Finnegan and Egan, 1981).

pEC400: 186 *Xho*I-*Bgl*II fragment from 67.5-79.6% with a deletion from 67.9-74.2% (Δ1) cloned into pKC7.

pEC200: This thesis, 186 *Eco*RI fragment from 92.0-2.3% cloned into pBR325.

These plasmids are represented diagrammatically in Figure 5.6.

4. Chemicals

Unless otherwise stated chemicals were analytical grade.

CsCl: Bethesda Research Laboratories

Urea: Sigma Chemical Co.

PEG6000: for phage preparation, from Sigma Chemical Co.

PEG6000: for sequencing, from BDH Chemicals Aust. Pty. Limited.

Phenol (Analar AR grade) was redistilled and stored in the dark at -15°C.

Bacto-Tryptone, Bacto-Agar and Yeast Extract were obtained from Difco Laboratories, U.S.A.

Amine A was obtained from Humpko Sheffield, U.S.A.

Brij58: Sigma Chemical Co.

Acrylamide: Sigma Chemical Co.

N,N'-Methylene-bis-Acrylamide: Sigma Chemical Co.

Diethyl-pryocarbonate: Sigma Chemical Co.

Agarose: Sigma Chemical Co.

L.G.T. agarose: Bethesda Research Laboratories.

IPTG: Sigma Chemical Co.

BCIG: Sigma Chemical Co.

ampicillin: Sigma Chemical Co.

tetracycline: gift from Upjohn Pty. Limited

chloramphenicol: Sigma Chemical Co.

ribonucleoside triphosphates: Sigma Chemical Co.

deoxynucleoside triphosphates: Sigma Chemical Co.

dideoxynucleoside triphosphates: Sigma Chemical Co.

5. Isotopes.

[α-^{32}P]-dCTP and [α-^{32}P]-dATP at specific activities of 1000 Ci/mmol, and [γ-^{32}P]-ATP at a specific activity of 2000 Ci/mmol, and [α-^{32}P]-GTP at a specific activity of 1000 Ci/mmol were prepared by Dr. R.H. Symons or more recently by Biochemical Research Enterprises of South Australia.

6. Enzymes

RNA polymerase: gift from W. Zillig at a concentration of 37 mg/ml.

Klenow fragment of *E. coli* DNA polymerase was obtained from Boehringer

Restriction endonucleases were obtained from New England Biolabs and used under the recommended conditions of the manufacturer.

T4 polynucleotide kinase: Boehringer

T4 DNA ligase: Boehringer

calf-intestional alkaline phosphatase: Sigma Chemical Co.

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

RNases A, T1 and U2 were obtained from the Sigma Chemical Co.

Phy M RNase and *B. cereus* RNase were gifts from J. Haseloff

lysozyme: Sigma Chemical Co.

7. Liquid media and buffers

All media and solutions of chemicals were prepared in glass-distilled water and were sterilised by autoclaving for 25 min at 120°C and 15 p.s.i.

L broth: 1% Bacto-tryptone, 0.5% yeast extract, 1% NaCl, pH 7.0

LGC broth: L broth supplemented with 0.1% glucose and 0.0024 M $CaCl_2$ (the glucose was made separately as a 20% sterile solution, the $CaCl_2$ was prepared as a 0.4 M solution and added to sterile broth).

YT broth: 8 g/l Bacto-tryptone, 5 g/l yeast extract 5 g/l NaCl (prepared as a 2 x solution).

M13 minimal media: 10.5 g/l K_2HPO_4, 4.5 g/l KH_2PO_4, 1.0 g/l $(NH_4)_2SO_4$, 0.5 g/l Na_3 citrate.$2H_2O$, this was autoclaved then the following was added from separately prepared sterile solutions; 0.8 ml of 1 M $MgSO_4$, 10 ml of 20% glucose, 0.5 ml 1% thiamine HCl.

Antibiotics were added to liquid media at the following concentrations; ampicillin 50 µg/ml, tetracycline 20 µg/ml.

TM: 0.01 M Tris-HCl, pH 7.1, 0.01 M $MgSO_4$

TE: 0.01 M Tris-HCl, pH 8.0, 0.001 M EDTA

$TE_{0.1}$: 0.01 M Tris-HCl, pH 8.0, 0.1 mM EDTA

TBE: 0.089 M Tris-HCl, 0.089 M boric acid, 2.7 mM EDTA prepared as a 10 x solution and pH to 8.3.

TAE: 0.04 M Tris-acetate, 0.02 M Na acetate, 1 mM EDTA prepared as a 10 x solution and pH to 8.2.

8. Solid Media

Soft agar: 1% Bacto-tryptone, 0.5% NaCl, 0.7% Bacto-agar

YT soft agar: YT broth + 0.7% Bacto-agar

Z plates: 1% Amine A, 0.5% NaCl, 1.2% Bacto-agar

YGC plates: 1% Amine A, 0.5% yeast extract, 1% NaCl, 1.5% Bacto-agar then add 5 ml 20% glucose and 6 ml 0.4 M $CaCl_2$ to one litre.

M13 cloning plates: add 1.5% Bacto-agar to M13 minimal media

Antibiotic plates: antiobiotics were added to YGC plates at the following concentrations; ampicillin 50 µg/ml, tetracycline 20 µg/ml, chloramphenicol 30 µg/ml.

Plates were prepared from 30 ml of the relevant mixture, dried with lids on at 37°C overnight and stored at 4°C until required.

9. Miscellaneous

Nitrocellulose: Schleicher and Schuell (pore size 0.45 µ)

BSA: Sigma Chemical Co. To remove nucleases the BSA was acetylated according to the method of Gonzalez *et al.* (1977).

t-RNA: Sigma Chemical Co. A soluiton was prepared and phenol extracted three times before use.

M13 specific 17-mer primer ($GTA_4CGACG_2C_2AGT$) was obtained

from New England Biolabs.

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

Positive/Negative Land Pack Film: Polaroid

Ethanol: redistilled under RNase free conditions and stored -20°C.

B. METHODS

1. Storage of bacteria and bacteriophage

Long term storage of bacterial cultures was in 40% glycerol frozen at -80°C. Bacterial stocks stored for limited periods were maintained at 4°C on either Z or YGC plates, or in the case of JM101, on M13 minimal media plates.

Low titre P2 phage stocks were maintained in LGC broth at 4°C after milliporing. High titre phage stocks were prepared by CsCl density centrifugation and the CsCl was removed by dialysis against 3 x 1 litre changes of TM and then stored at 4°C.

2. Growth of Bacterial Cultures

A stationary phase bacterial culture was prepared by inoculating broth with a loopful of bacteria from a plate stock and incubating overnight in capped flasks at 30°C or 37°C in a New Brunswick gyrotory water bath.

Log phase cultures were prepared by diluting a freshly grown stationary phase culture 50 fold into sterile broth, and incubating at 30°C or 37°C until the required cell density was reached. Cell density was measured by following OD_{600} using a Gilford 300 T-1 Spectrophotometer.

3. Titration of phage stocks

Phage were diluted into TM buffer and a 0.1 ml sample was added to 0.2 ml log phase indicator bacteria (log phase

culture). Three ml of soft agar (melted and maintained at 45°C) was added and the contents of the tube poured over an agar plate. When the agar had solidified the plates were inverted and incubated overnight at 37°C. Plaques were scored the following day.

For assays of 186, bacteria were grown in L broth and Z plates were used. Indicator bacteria were grown in LGC broth and YGC plates were used for assays of P2 and Hy5.

4. Preparation of phage stocks

High titre stocks of all phage were prepared with the aim of obtaining phage DNA.

a) 186 phage

186*cIts* as well as 186*cIts*B*am*17 and 186*cIts*B*am*57 were prepared by heat induction of log phage cultures.

Stationary phase cultures grown at 30°C were diluted fifty-fold into 500 ml L broth and incubated at 30°C with aeration to an OD_{600} = 0.8. Cultures were then transferred to a 45°C water bath and shaken by hand for fifteen minutes. They were then returned to 37°C and shaken until lysis was complete. 19 g/l NaCl was added and allowed to dissolve. The lysates were centrifuged (9,000 r.p.m., 4°C, 20 minutes) to remove bacterial debris and then 100 g/l PEG 6000 was added to the supernatant. The solution was stored at 4°C overnight and the precipitate was collected by centrifugation (9,000 r.p.m., 4°C, 20 minutes) and resuspended in a small volume of TM and then purified by centrifugation in CsCl.

b) P2 phage

High titre stocks of P2 and P2*vir*22*ogr*52 were prepared by infection of a large scale culture with a low titre P2

stock.

P2 was raised on E814 and P2*vir*22*ogr*52 was raised on E273. Low titre P2 stocks were prepared as follows. To 15 ml LGC broth 0.6 ml log phase bacteria were added and the culture grown at 37°C to an OD_{600} = 0.8. P2 phage leached from a single plaque into 1 ml LGC broth was added to the culture and the optical density followed. At the beginning of lysis 0.1 M EDTA, pH 8.0 was added to a final concentration of 6 mM and incubation continued till lysis was complete. Bacterial debris was removed by centrifugation and the supernatant millipored and stored at 4°C.

For high titre P2 stocks stationary phase bacteria were diluted a hundred-fold into 500 ml LGC broth. The culture was grown to OD_{600} = 0.2 and then infected (moi = 0.2) with P2 phage from a low titre stock. Incubation was continued at 37°C and 0.1 M EDTA, pH 8.0 added to a concentration of 6 mM at the onset of lysis. The phage were prepared for CsCl purification at the completion of lysis as in Section 2.B.4(a) except that NaCl was added to the lysate at 25 g/l instead of 19 g/l.

c) Hy5 phage

Stocks were prepared by heat induction of a Hy5 lysogen by the method described for 186 in Section 2.B.4(a) except that L broth was supplemented with 10 mM $MgSO_4$ and 20 µg/ml thymine.

d) Block gradient centrifugation

Two solutions of CsCl were prepared in TM of densities 1.6 g/ml and 1.35 g/ml. The block gradient was prepared by adding 3 ml of CsCl density 1.35 into a 10 ml polycarbonate

oak ridge tube and underlaying it with 1.5 ml of CsCl density 1.6. The phage suspension in TM was layered over the block gradient and the tube centrifuged (45,000 r.p.m., 8°C, 60 minutes) in a Beckman Ti-50 rotor.

The phage were recovered by piercing the bottom of the tube and collecting the opaque phage band. The phage were then further purified by reverse block centrifugation. An equal volume of saturated CsCl was added to the phage suspension and this was underlayed beneath 1.5 ml 1.6 density and 2.5 ml 1.35 density. The tube was topped with paraffin oil and centrifuged (45,000 r.p.m., 8°C, 90 minutes) in a Ti-50 rotor.

5. Extraction of bacteriophage DNA

An equal volume of TE was added to redistilled phenol, shaken and the phases allowed to separate. TE saturated phenol was taken from the lower phase.

A sample of a high titre phage stock (generally 0.2 or 0.3 ml) was diluted to 5.0 ml in TE and an equal volume of TE saturated phenol added. The mixture was shaken and the phases separated by centrifugation in Corex 30 ml glass tubes at 7,500 r.p.m., 4°C, 5 minutes. The aqueous phase was withdrawn and re-extracted twice with an equal volume of TE saturated phenol. The phenol layer was washed with an equal volume of TE. The aqueous phases were pooled and 3 x ether washed. 4 M NaCl was added to the aqueous phase to a final concentration of 0.2 M along with 2.5 volumes of ethanol. The DNA was allowed to precipitate at - 20°C overnight before being pelleted by centrifugation (10,000 r.p.m., 4°C, 15 minutes) in a Beckman JA-20 rotor. The DNA pellet was dried

in vacuo and resuspended in TE. To check the purity of the DNA, spectra of absorbance over the range 230-340 nm were determined on a Varian Super Scan 3 Ultra-Violet Spectrophotometer. $OD_{260/280}$ and $OD_{260/230}$ ratios were greater than 1.8 for all DNA used.

6. Plasmid DNA preparation

Two techniques were employed depending on the subsequent use of the DNA. For transcription *in vitro*, plasmid DNA was prepared by a gentle lysis procedure. For analysis of plasmids by restriction endonuclease digestion, the alkaline/SDS procedure of Birmboim and Doly (1979) was used. These techniques are described below.

a) Gentle lysis procedure

10 ml plasmid containing cells were grown in L broth + antibiotic at 37°C overnight. This was subcultured into 800 ml of L broth and grown with aeration to an OD_{600} = 1.0. The cells were amplified by adding chloramphenicol 160 mg/800 ml (for pBR325, amplification was omitted) and continuing incubation at 37°C overnight. The cells were collected by centrifugation and resuspended in 7.5 ml 25% sucrose, 50 mM Tris, pH 8.0. 2.0 ml of a 10 mg/ml freshly prepared lysozyme solution was added and the tube left on ice 5 minutes. Then 3 ml 0.25 M EDTA, pH 8.0, was added, mixed gently and kept on ice for 5 minutes. 12.0 ml of detergent solution (1% Brij 58, 0.4% Na deoxycholate, 50 mM Tris, 25 mM EDTA) was added and the solution kept on ice a further 10 minutes. The solution was then centrifuged (18,000 r.p.m., 4°C, 30 minutes) and 0.95 g CsCl and 20 µl of a 10 mg/ml solution of ethidium bromide added per ml of the supernatant. This

mixture was centrifuged (45,000 rp.m., 20°C, 36 hours) in a Ti-50 rotor. The plasmid band was recovered by piercing the bottom of the polycarbonate tube and recentrifuged after adding 6 ml of a 0.95 g/ml solution of CsCl (45,000 r.p.m., 20°C, 24 hours).

The ethidium bromide was removed by at least three extractions with propan-2-ol. The DNA was ethanol precipitated, washed with 70% ethanol, dried *in vacuo* and resuspended in TE.

b) Alkaline/SDS procedure

1 ml of a stationary phase plasmid containing culture was centrifuged in an eppendorf centrifuge for 2 minutes and the cells resuspended in 100 µl, 25 mM Tris, pH 8.0, 10 mM EDTA, 15% sucrose, to which 5 mg/ml lysozyme had been freshly added. After 40 minutes on ice, 200 µl 0.2 N NaOH, 1% SDS was added and the mixture left on ice 10 minutes. 150 µl 3 M Na acetate, pH 4.6 was added, the tubes kept on ice for 40 minutes and then centrifuged for 10 minutes. The supernatant was ethanol precipitated, washed with 70% ethanol, dried and redissolved in 39 µl water. 1 µl of a 5 mg/ml RNase A solution, which had been heated at 80°C 10 minutes to inactivate DNases, was added and the solution incubated for 20 minutes at 37°C. Restriction endonuclease digestion was carried out directly.

7. Preparation of the replicative form of M13mp9

A 50 ml stationary culture of JM101 grown in M13 minimal media was subcultured into 800 ml 2X YT broth and grown at 37°C to an OD_{600} of 0.8. At the same time a 5 ml culture of JM101 was grown to an OD_{600} of 0.6 and 200 µl of a single-

strand M13mp9 phage supernatant added (see Section 2.B.17(a). Incubation was continued at 37°C for 1 hour. This 5 ml culture was then added to the 800 ml culture and grown overnight at 37°C. The gentle lysis procedure (Section 2.B.6(a)) including CsCl purification was then employed for the isolation of replicative form M13mp9.

8. Cloning of 186 restriction fragments into plasmids

a) Ligation

Donor and vector DNA digested with the appropriate restriction endonucleases were ethanol precipitated and resuspended in TE. The insertion of specific fragments into plasmid vectors was optimised by adjusting the concentration of both insert and vector DNA according to the mathematical treatment of Dugaiczuk *et al*. (1975). Ligation was carried out in 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP with 0.2 units of T4 DNA ligase. The reaction mix was incubated at 14°C for 16 hours.

b) Transformation

Stationary phase E536 bacteria were diluted into L broth to a starting OD_{600} = 0.01, and grown at 37°C with shaking to an OD_{600} = 0.3. The culture was chilled on ice for 10 minutes and then the cells were washed once in one half volume of cold 0.1 M $CaCl_2$ and left on ice for 20 minutes. The cells were finally resuspended in one-tenth original volume of cold 0.1 M $CaCl_2$ and left on ice for at least two hours. Aliquots of ligation mix were added to 0.1 ml competent cells. This mixture was chilled for 10 minutes then placed at 37°C for 5 minutes before 2 ml L broth was added. The cells were shaken at 37°C for 2 hours and then 0.1 ml aliquots plated onto YGC + antibiotic plates.

Plates were incubated at 37°C overnight and then recombinants were selected by screening for sensitivity to the second antibiotic (i.e., the determinant destroyed by insertion of the donor fragment).

9. Transfer of DNA from a gel to nitrocellulose

Urea/sucrose loading buffer (4 M urea, 50% sucrose, 50 mM EDTA, 0.1% bromophenol blue) was added to restriction endonuclease digests of phage DNA and the fragments separated on 1% agarose gels containing TAE buffer. The method used for transferring the DNA from the agarose gel to nitrocellulose was essentially that described by Southern (1975) and modified by Wahl *et al.* (1979).

The DNA was first denatured by bathing the gel twice for 15 minutes in 0.25 N HCl. The gel was then rinsed in water then shaken in 0.5 N NaOH, 1 M NaCl for a further two 15 minute periods. It was rinsed once again then neutralized in 0.5 M Tris, pH 7.4, 3 M NaCl twice for 15 minutes. Finally, the gel was bathed in 1 M NH_4 acetate, 20 mM NaOH for 30 minutes. The nitrocellulose was wetted then bathed for 30 minutes in 1 M NH_4 acetate, 20 mM NaOH before being layed over the gel. A wad of paper towels were placed on top of the nitrocellulose and held in place by a light weight. The following morning the nitrocellulose was air dried and baked *in vacuo* for three hours to immobilize the DNA.

10. Hybridization to transferred DNA

The nitrocellulose was cut into strips 1 cm wide and sealed in a plastic bag along with the radioactive probe in 2 x SSC (0.3 M NaCl, 0.03 M Na_3 citrate, pH 7.4), 0.1% SDS.

The bags were immersed in water to prevent drying and incubated at 60°C for 24 hours.

Upon removal, the nitrocellulose strips were washed extensively as follows; twice for 5 minutes in 2 x SSC at room temperature followed by three 5 minute washes at room temperature in 2 x SSC, 0.1% SDS. Then twice for 15 minutes in 0.1 x SSC, 0.1% SDS at 50°C and finally two 45 minute washes in 0.01 x SSC, 0.1% SDS at 50°C.

The nitrocellulose strips were air dried then auto-radiographed.

11. Extraction of DNA from L.G.T. agarose

Some of the methods to follow involve the manipulation of isolated restriction fragments. When fragments were to be isolated the DNA was electrophoresed in low gelling temperature agarose.

The gel was stained with acridine orange (100 µg/ml) and the DNA bands, visualised by eye, excised. The gel slice (of volume 100 µl when melted) was incubated at 65°C for 5 minutes and 200 µl 0.2NETM (0.2 M NaCl,/mM EDTA, 10 mM Tris, pH 7.4, 10 mM $MgCl_2$) was added. The tube was removed from 65°C and an equal volume of TE saturated phenol added. The solution was mixed and the phases separated by centrifugation. The aqueous phase was removed and the phenol phase washed with a further 100 µl 0.2NETM. The aqueous phases were pooled, ethanol precipitated, washed with 70% ETOH, dried *in vacuo* and resuspended in $TE_{0.1}$. The DNA thus prepared was suitable for further digestion by restriction endonucleases and for cloning.

12. Binding single-stranded DNA to nitrocellulose

Single-stranded 186 DNA was prepared by B. Kalionis using the method of Szybalski *et al.* (1971) and the single-strands bound to nitrocellulose essentially as described by Kafatos (1980). Nitrocellulose was pre-soaked in water and placed on top of a wad of dry paper towels and a moist nitrocellulose strip. An equal volume of 2 M ammonium acetate was added to the DNA and this was then spotted onto nitrocellulose using a drawn out capillary. The DNA spot was washed through with drops of 1 M ammonium acetate, then the filters were air dried and baked for 2 hours *in vacuo* at 80°C.

Hybridization, washing of filters and autoradiography was exactly as described in Section 2.B.10.

13. *In vitro* transcription

The optimal conditions for 186 *in vitro* transcription were determined as described in Chapter Three, therefore I will only record the optimal conditions here, any variation is given in the figure legends.

2.5 µg DNA (phage or plasmid) was added to a reaction mix containing 25 mM Tris-HCl, pH 8.0, 0.15 M KCl, 1.6 mM Na_2 EDTA, 1.0 mM DTT (Added freshly), 0.2 mM each of ATP, CTP, UTP and 0.02 mM GTP in a final reaction volume of 24 µl. This mix was added to 2 µCi of [α-^{32}P]-GTP which had been dried *in vacuo*, then 4.6 µg of RNA polymerase was added. The reaction was incubated at 37°C 5 minutes to allow pre-binding of RNA polymerase to promoters, then the reaction was initiated by the addition of $MgCl_2$ to a final concentration of 10 mM. RNA synthesis was allowed to proceed for

10 minutes then the reactions were stopped by the addition of 2.5 volumes ethanol. 20 µg tRNA was added as carrier and the precipitated RNA washed in 70% ethanol, dried *in vacuo* and resuspended in water. An equal volume of formamide loading buffer (95% deionized formamide, 10 mM Na_2 EDTA, 0.02% (w/v) bromophenol blue, 0.02% (w/v) xylene cyanol FF) was added, and the tubes heated at 65°C for 5 minutes and snap cooled on ice before electrophoresis. The samples were loaded onto either 20 x 40 x 0.025 cm or 16 x 16 x 0.05 cm 5% polyacrylamide gels containing TBE and 8 M urea. For electrophoresis of transcription reactions, the acrylamide:N,N'-methylene-bis-acrylamide ratio was 30:0.8. After electrophoresis the gel was autoradiographed.

a) Preparative RNA synthesis

For preparative RNA synthesis the transcription reactions were scaled up to 100 µl, the reactions pooled and loaded onto a preparative track of a polyacrylamide gel. If unlabelled RNA was required the [α-^{32}P]-GTP was omitted and replaced by 0.2 mM GTP.

b) [γ-^{32}P]-ATP labelling of transcripts

In vitro transcription reactions were prepared on a preparative scale as described above except that the radiolabelled species was [γ-^{32}P]-ATP and the other three nucleotide triphosphates were at 0.2 mM. Unlabelled ATP was omitted but after the addition of $MgCl_2$ followed by 10 minutes at 37°C, the reaction was "chased" by the addition of 0.2 mM ATP and incubation continued for a further 10 minutes.

14. Extraction of DNA/RNA from polyacrylamide gels

Whenever nucleic acids were to be excised, gels of 0.05 cm thickness were run. After autoradiography at room temperature for 20-90 minutes, the resultant autoradiograph was used as a template to locate and excise the ^{32}P-labelled fragments. Excised bands were eluted by soaking overnight (at room temperature for RNA or at 37°C for DNA) in 500 µl of 0.5 M ammonium acetate, 1 mM Na_2 EDTA, 0.1% SDS, which contained 60 µg *E. coli* tRNA as carrier if the RNA bands were to be sequenced using the partial enzymatic cleavage technique. After soaking, the elution buffer was removed and the fragments precipitated by the addition of 1 ml of ethanol. After centrifugation at 10,000 g for 10 minutes at 4°C, the pelleted fragments were washed with 70% ethanol, dried *in vacuo* and resuspended in water.

15. 5'-^{32}P-labelling of RNA

a) Removal of 5'-phosphates

RNA fragments were suspended in 50 µl water and 6 µl 1 M Tris, pH 9.0 added along with 2 µl 5% SDS and 0.3 units calf intestinal alkaline phosphatase. The reaction was incubated at 37°C for 30 minutes, heated at 65°C for 10 minutes and then extracted with 50 µl TE saturated phenol. The aqueous phase was removed and sodium acetate added to 0.3 M. The RNA was then precipitated with 150 µl ethanol, washed with 70% ethanol, dried *in vacuo*.

This method was also employed to remove the 5'-phosphates from pBR325 prior to cloning.

b) Kinasing of RNA

The precipitated, dried RNA was resuspended in 20 µl

water and 5 µl of 5 x polynucleotide kinase buffer (125 mM Tris-HCl, pH 9.0, 50 mM $MgCl_2$, 50 mM DTT) added. This mix was transferred to a tube in which 10 µCi [γ-^{32}P]-ATP had been dried and 5.0 units of polynucleotide kinase was added. The reaction was incubated at 37°C for 30 minutes then ammonium acetate added to 0.3 M followed by sodium acetate to 0.3 M. The RNA was ethanol precipitated, washed with 70% ethanol, dried and resuspended in 5 µl water plus 5 µl formamide loading buffer (95% deionized formamide, 10 mM EDTA, 0.02% xylene cyanol FF, 0.02% bromophenol blue). The tubes were heated at 65°C for 5 minutes then loaded onto a 20 x 40 x 0.05 cm 5% polyacrylamide gel containing TBE and 8 M urea. After electrophoresis and autoradiography the labelled RNA was excised and eluted from the gel slice as described in Section 2.B.14.

16. Sequence determination of RNA using the partial enzymatic cleavage technique

The procedure is essentially that of Donis-Keller *et al.* (1977). 5'-^{32}P-labelled RNA with 60 µg *E. coli* tRNA, was resuspended in 30 µl water and six aliquots of 5 µl dispensed. The RNA was dried *in vacuo* and then resuspended in the appropriate buffer as follows:

Tube N	6 µl	20 mM Na citrate, pH 5.0
(No enzyme)		1 mM Na_2 EDTA
		7 M urea
Tube T	6 µl	20 mM Na citrate, pH 5.0
(RNase T1)		1 mM Na_2 EDTA
		7 M urea

Tube U	6 µl	20 mM Na citrate, pH 3.5
(RNase U2)		1 mM Na_2 EDTA
		7 M urea
Tube L	5 µl	50 mM $Na_2CO_3/NaHCO_3$,
(alkali ladder)		pH 9.0
Tube P	6 µl	20 mM Na citrate, pH 5.0
(RNase Phy M)		1 mM Na_2 EDTA
		7 M urea
Tube B	5 µl	20 mM Na citrate, pH 5.0
(*Bacillus cereus* RNase)		1 mM Na_2 EDTA

Tubes N, T, U, P, and B were heated at 80°C for 1 minute, snap cooled on ice, and the ribonucleases added.

Tube N -

Tube T 0.5 µl 10,000 units/ml RNase T1; 5 units

Tube U 2 µl 200 units/ml RNase U2; 0.4 units

Tube L -

Tube P 1 µl RNase Phy M extract

Tube B 2.5 µl *B. cereus* RNase extract

Tube L was heated at 100°C for 90 seconds while the remaining tubes were incubated at 50°C for 20 minutes. At completion of the sequencing reactions formamide loading buffer (95% deionized formamide, 10 mM EDTA, 0.02% xylene cyanol FF, 0.02% bromophenol blue) was added to the samples to a final volume of 9 µl. Samples were heated at 80°C for 1 minute and snap cooled on ice before being loaded onto a 20 x 40 x 0.05 cm 20% or 6% polyacrylamide gel (ratio acrylamide to N-N'-methylene-bis-acrylamide was 19:1) containing TBE and 8 M urea. The bands were visualised by autoradiography at -80°C, using calcium tungstate intensifying screens.

17. Sequence determination of RNA using cloned DNA primers

The DNA primers used in this work were fragments transcribed using recombinant phage M13 single-strand DNA as a template. The recombinant M13 containing 186 sequences of the same polarity as the RNA sequence was transcribed using an M13 specific oligonucleotide primer and the Klenow fragment of *E. coli* DNA polymerase 1 with [α-^{32}P]-dCTP essentially as described by Bruening *et al.* (1982). The partially double-stranded DNA molecules were subjected to restriction enzyme digestion, and the labelled fragments fractionated by polyacrylamide gel electrophoresis. The primer band was excised and eluted from the gel slice as described in Section 2.B.14.

a) RNA-DNA hybridization

RNA-DNA hybrids were prepared as follows. The purified primer and the appropriate RNA molecule were resuspended in 25 µl water and 25 µl of 2 x hybridization buffer (0.36 M NaCl, 20 mM Tris-HCl, pH 7.0, 2 mM EDTA, 0.1% SDS) was added. The tube was heated at 100°C for 1½ minutes, then incubated at 60°C for 1-2 hours. The RNA-DNA hybrids were ethanol precipitated, washed with 70% ethanol and dried *in vacuo*.

b) Reverse transcription

The RNA-DNA hybrids were resuspended in 10 µl of water and 2 µl dispensed into each of four tubes, A, C, G, T containing 2 µl of the following triphosphate mixes in 1 x reverse transcriptase buffer (50 mM Tris-HCl, pH 8.3, 50 mM KCl, 10 mM $MgCl_2$, 10 mM DTT).

A: 0.02 mM dATP, 0.2 mM dCTP, dGTP, dTTP, 0.01 mM ddATP

C: 0.02 mM dCTP, 0.2 mM, dATP, dGTP, dTTP, 0.002 mM ddCTP

G: 0.02 mM dGTP, 0.2 mM dATP, dCTP, dTTP, 0.002 mM ddGTP

T: 0.02 mM dTTP, 0.2 mM dATP, dCTP, dGTP, 0.01 mM ddTTP.

To this mix 2 units of avian myeloblastosis virus reverse transcriptase in 50 mM Tris-HCl, pH 8.3, 50 mM KCl, 10 mM $MgCl_2$, 10 mM DTT was added to give a final volume of 5 µl. The tubes were incubated at 37°C for 30 minutes then 5 µl formamide loading buffer (95% deionized formamide, 10 mM EDTA, 0.02% xylene cyanol FF, 0.02% bromophenol blue) was added to each reaction and the tubes heated at 100°C for 3 minutes before electrophoresis on 20 x 40 x 0.025 cm 5% polyacrylamide gels, (acrylamide to N-N'-methylene-bis-acrylamide was 19:1 ratio) containing TBE and 8 M urea. The gels were fixed by washing in 20% ethanol, 10% acetic acid and dried at 120°C before autoradiography.

18. Sequence determination of DNA using the dideoxynucleotide technique

a) Cloning of DNA into M13mp9

To obtain a DNA sequence using the dideoxynucleotide technique of Sanger *et al.* (1977) phage DNA was firstly recombined into the replicative form of M13mp9.

Preparative amounts of the DNA fragment to be sequenced were isolated from a low gelling temperature agarose gel (Section 2.B.11) and mixed in a donor to vector ratio of 10 ng/kb donor to 20 ng mp9 vector digested with the appropriate restriction enzymes. The ligation reaction was carried out in a total volume of 10 µl containing 10 mM $MgCl_2$, 1 mM ATP, 10 mM DTT, and 0.2 units of T4 DNA ligase at 14°C overnight.

Competent JM101 were prepared by inoculating 2 x YT

broth with a stationary phase JM101 grown overnight in M13 minimal media at 37°C. Cells were grown to OD_{600} of 0.6-0.8 and the culture chilled for 10 minutes on ice. The cells were resuspended after centrifugation at 5,000 r.p.m. at 4°C in one fifth the original volume of cold 0.1 M $CaCl_2$ and kept on ice for at least 2 hours.

200 µl competent JM101 were transfected with 1-5µl of the ligation mix and left on ice for 40 minutes. The cells were heat shocked at 45°C for 2 minutes and added to a tube containing 3 ml YT soft agar, 200 µl log phase JM101, 20 µl IPTG (24 mg/ml) and 20 µl BCIG (20 mg/ml). This tube was poured onto a M13 cloning plate, allowed to set and incubated at 37°C overnight.

Colourless M13 plaques were toothpicked into 2 ml 2 x YT broth inoculated with a stationary phase JM101, and grown at 37°C with shaking for 6 hours. The DNA was prepared from the supernatant as follows. The cells were pelleted and 270 µl 20% PEG, 0.25 M NaCl added to 1 ml of the supernatant. This was left at room temperature for 15 minutes then centrifuged 10,000 g for 5 minutes. The resultant pellet was resuspended in 200 µl $TE_{0.1}$ and extracted with 100 µl TE saturated phenol. The aqueous phase was ethanol precipitated after adding 6 µl 3 M ammonium acetate, pH 4.6. The DNA pellet was washed with 95% ethanol, dried *in vacuo* and resuspended in 18 µl water. DNA thus prepared was stored at -20°C.

b) Sequence determination

In a final volume of 10 µl, 1 µl of the M13 specific

17-mer was mixed with 6 µl of the recombinant DNA prepared above and 1 µl of 10 x TM. This mixture was incubated at 60°C for 60 minutes, then at room temperature for 10 minutes. This annealed DNA was added into a tube in which 3.2 µM [α-^{32}P]-dCTP and 2 µl of 15.75 µM dCTP had been dried, and then 2 µl of this DNA-label dispensed into four tubes containing the remaining deoxynucleotides and a single dideoxynucleotide species at the following final concentrations;

A: 2.8 µM dATP, 43.5 µM dGTP, dTTP, 100 µM ddA

C: 30 µM dATP, dGTP, dTTP, 20 µM ddC

G: 43 µM dATP, dTTP, 3.2 µM dGTP, 60 µM ddG

T: 43 µM dATP, dGTP, 3.2 µM dTTP, 160 µM ddT

0.25 units Klenow was added to each tube and the reactions incubated at 37°C for 15 minutes. After synthesis the reactions were "chased" with 2 µl of a solution containing all four deoxynucleotides at 0.25 mM each for 15 minutes at 37°C. 4 µl of formamide loading buffer (95% deionized formamide, 10 mM EDTA, 0.02% bromophenol blue, 0.02% xylene cyanol FF) was added and the tubes heated at 100°C for 3 minutes before being loaded onto 20 x 40 x 0.025 cm 5% polyacrylamide gels (acrylamide to N-N'-methylene-bis-acrylamide was 19:1) containing TBE and 8 M urea. After electrophoresis the gels were fixed in 20% ethanol and 10% acetic acid before autoradiography at room temperature.

The work described in this thesis is divided into three sections. The first section, Chapter 3, describes the establishment of an *in vitro* transcription system for 186 DNA. The second section, Chapters 4 to 7, are concerned with the characterisation of the transcripts. The final section describes the characterisation of the P2 *ogr* gene transcript with the aim of comparing it to the 186 B gene.

CHAPTER 3

ESTABLISHMENT OF 186 *IN VITRO* TRANSCRIPTION

A. INTRODUCTION

This chapter comprises Section One, it describes the establishment of a 186 *in vitro* transcription system.

A prerequisite for faithful, specific transcription is RNA polymerase holoenzyme. The sigma subunit confers specificity on the enzyme by recognising promoter sequences on the DNA template. Core polymerase is capable of initiating synthesis at biologically incorrect sites on a template, for example at single-strand nicks. W. Zillig (Max-Planck-Institute for Biochemistry, Munich, Germany) sent RNA polymerase holoenzyme which was used in all the *in vitro* transcription experiments.

Technical difficulties were encountered before a system was established which enabled visualisation of 186 *in vitro* transcripts. Obtaining transcription of 186 DNA was not a problem as was indicated by the incorporation of labelled precursor into nucleic acid which was retained by TCA precipitation on GF/A filters. The problems were encountered in the treatment of the RNA prior to gel electrophoresis and in the choice of gel system.

B. RESULTS AND DISCUSSION

1. Technical Difficulties

Initially, 186 DNA was transcribed using the conditions described by Dunn and Studier (1973) for analysis of T7 *in vitro* transcripts. The reaction mix contained 20 mM Tris-HCl, pH 7.9, 10 mM $MgCl_2$, 50 mM KCl, 0.1 mM EDTA, 0.1 mM DTT, 0.1 µg/µl BSA which had been acetylated to remove nucleases, 0.185 µg/µl RNA polymerase, 0.1 µg/µl 186 DNA

and 0.2 mM each of ATP, CTP, UTP and 0.02 mM GTP plus 2 μCi [α-^{32}P]-GTP. The reaction mixes were prepared at 0°C and a 2 μl sample removed and spotted onto a GF/A filter. The filter was not TCA precipitated and so represented total counts in the reaction mix. The tubes were transferred to 37°C for 15 minutes and an equal volume of 50 mM EDTA was added to stop the reaction. A 5 μl aliquot was spotted onto GF/A filters which were immediately plunged into ice-cold 10% TCA + 2% NaH_2PO_4 + 2% $Na_4P_2O_7$. The filters were washed four times and finally washed in ether/acetone (v:v; 1:1), then dried and Cherenkov counted. A control reaction was included which contained no RNA polymerase. A representative comparison of TCA precipitable counts obtained with 186 DNA in the presence and absence of RNA polymerase is presented in Table 3.1. It was evident from the counts retained above the control that the system was capable of 186 *in vitro* transcription.

For gel electrophoresis the *in vitro* transcription reactions were ethanol precipitated, the samples dried *in vacuo* and resuspended in sample buffer (1 x TBE, 30% glycerol, 1% SDS, 0.1% bromophenol blue). They were then heated for 2 minutes in a boiling water bath according to Dunn and Studier (1973) prior to electrophoresis on 2% polyacrylamide/ 0.5% agarose composite gels containing TBE and 0.1% SDS as described by Golomb and Chamberlin (1974). The gel is shown in Figure 3.1(a). A smear, with no visible 186 transcripts resulted.

Later experiments involved electrophoresis of 186 transcription reactions prepared according to Dunn and Studier

TABLE 3.1

186 TCA PRECIPITABLE COUNTS

	(Control) Minus RNA polymerase	plus RNA polymerase
Total c.p.m.	269,990	310,880
TCA precipitable c.p.m.	1,050	20,246

The counts are calculated per 10 µl aliquot of the *in vitro* transcription reaction.

(1973) on 1.5% agarose gels containing 2.2 M formaldehyde (Lehrach *et al.*, 1977). RNA samples were heated in 2.2 M formaldehyde, 50% formamide, 0.018 M Na_2HPO_4 - 0.002 M NaH_2PO_4, for 5 minutes at 60°C. The gel is presented in Figure 3.1(b). Visible above a background smear are three ill defined 186 transcripts. Although RNA breakdown was still apparent, this combination of sample preparation and gel system did permit the visualisation of 186 transcripts.

A literature survey revealed numerous *in vitro* transcription reaction mixes, all containing similar components but the concentrations of these components varied. A reaction mix was devised and successfully employed for the *in vitro* transcription of 186 DNA. This method is detailed in the Materials and Methods section in Chapter 2, except that transcription was terminated by the addition of 6 µl of stop buffer (1 x TBE, 22.5 mM EDTA, 0.1% SDS, 8 M urea, 1% bromophenol blue, 1% xylene cyanol FF). The samples were loaded directly onto a 16 x 16 x 0.2 cm 5% polyacrylamide gel containing TBE and 8 M urea after heating at 65°C for 5 minutes. The result is shown in Figure 3.1(c). Three discrete transcripts were evident.

Resolution of the transcripts was improved by ethanol precipitation of the RNA prior to electrophoresis and then resuspending the transcripts in formamide loading buffer. Gels of 0.05 or 0.025 cm thickness also improved band resolution.

This procedure was followed for all subsequent *in vitro* transcription reactions and is recorded in the Materials and

FIGURE 3.1

TECHNICAL DIFFICULTIES

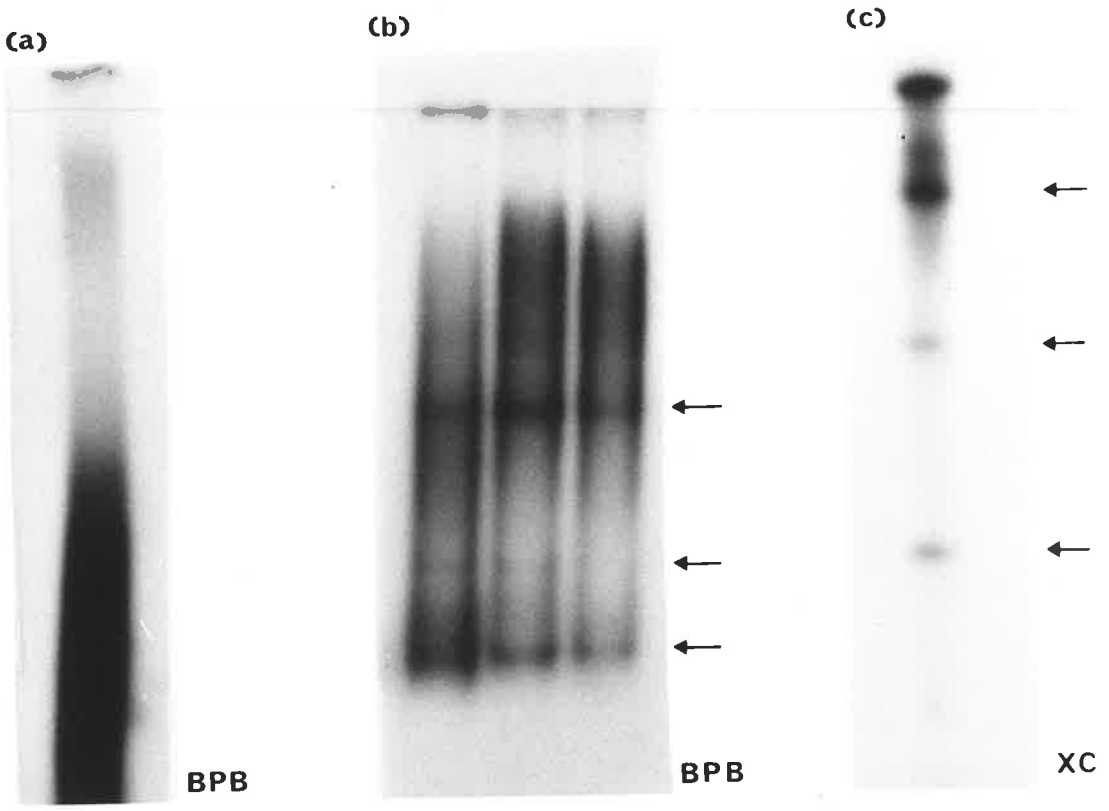
a) 2% acylamide/0.5% agarose composite gel containing TBE and 0.1% SDS. Samples were boiled for 2 minutes before electrophoresis.

b) 1.5% agarose containing 2.2 M formaldehyde. The 3 186 transcripts are indicated by arrows.

The transcription reactions were as described by Dunn and Studier (1973).

c) 5% polyacrylamide gel containing TBE and 8 M urea. (16 x 16 x 0.2 cm). The transcription reaction was as described in the Materials and Methods except stop buffer was added instead of formamide loading buffer before electrophoresis.

The three transcripts are designated.


(a)
(b)
(c)
BPB
BPB
XC

Methods section.

In summary then, before 186 *in vitro* transcripts could be visualised, three technical difficulties were surmounted. Most importantly, was the abolition of boiling the RNA prior to electrophoresis as boiling of samples resulted in breakdown of the RNA manifested by a smear on the gel. Heating at 65°C for 5 minutes in a denaturant such as formamide is sufficient for disrupting RNA aggregates and RNA secondary structures. The choice of gel system was also an important aspect to consider, as resolution of the transcripts on agarose/formaldehyde or acrylamide/SDS gels (even without boiling the samples) was inferior to that of polyacrylamide/urea gels. The choice of loading buffer also contributed to the resolution of the transcripts. Superior resolution was obtained with formamide as the denaturant rather than urea.

2. Optimal conditions for 186 transcription

Conditions had been established for 186 *in vitro* transcription. The following section is concerned with the optimisation of these conditions.

From a literature survey, the two components influencing the specificity of transcription initiation were the concentration of RNA polymerase and the ionic strength.

It had been found empirically that dilution of RNA polymerase into a storage buffer before addition to a reaction mix caused a loss of enzyme activity, so the smallest practicable amount (0.125 µl) of RNA polymerase was added from the original 37 µg/µl stock to the *in vitro*

reaction mix attaining a final enzyme concentration of 0.185 µg/µl, which fortuitously produced discrete 186 transcripts. The amount of RNA polymerase was varied to see the effect, if any, on the 186 *in vitro* transcription pattern to empirically determine the optimal concentration of RNA polymerase for 186 transcription. 2.5 µg of 186 DNA was transcribed at 100 mM KCl whilst varying the amount of RNA polymerase from 0.185 µg/µl to 1.48 µg/µl (the volume varied from 0.125 µl to 1 µl). Figure 3.2 shows the result of such an expériment. As the RNA polymerase concentration increased from 0.185 µg/µl to 1.48 µg/µl the level of transcription decreased. This is difficult to explain as, above an optimum RNA polymerase concentration, I would have expected to see an increase in non-specific transcription due to initiation from minor promoters, which only exhibit initiation when RNA polymerase is added in excess of that needed to saturate the stronger early promoters (Stahl and Chamberlin, 1977), and non-specific initiation. This total decrease in the level of transcription was baffling but was not pursued further. Future transcription reactions used 0.185 µg/µl RNA polymerase.

To determine the optimal salt conditions for 186 transcription, 186 DNA was transcribed, keeping other reaction components constant, but varying the KCl concentration. Figure 3.3 shows an autoradiogram of 186 transcription reactions at KCl concentrations ranging from 50 to 200 mM. As the KCl concentration increased so the background of non-specific transcription decreased. There is a marked

FIGURE 3.2

VARYING THE RNA POLYMERASE CONCENTRATION

2.5 µg 186 DNA was transcribed at a KCl concentration of 100 mM and various RNA polymerase concentrations as indicated above the autoradiogram. The transcripts were electrophoresed on a 16 x 16 x 0.05 cm 5% polyacrylamide gel containing 8 M urea and TBE.

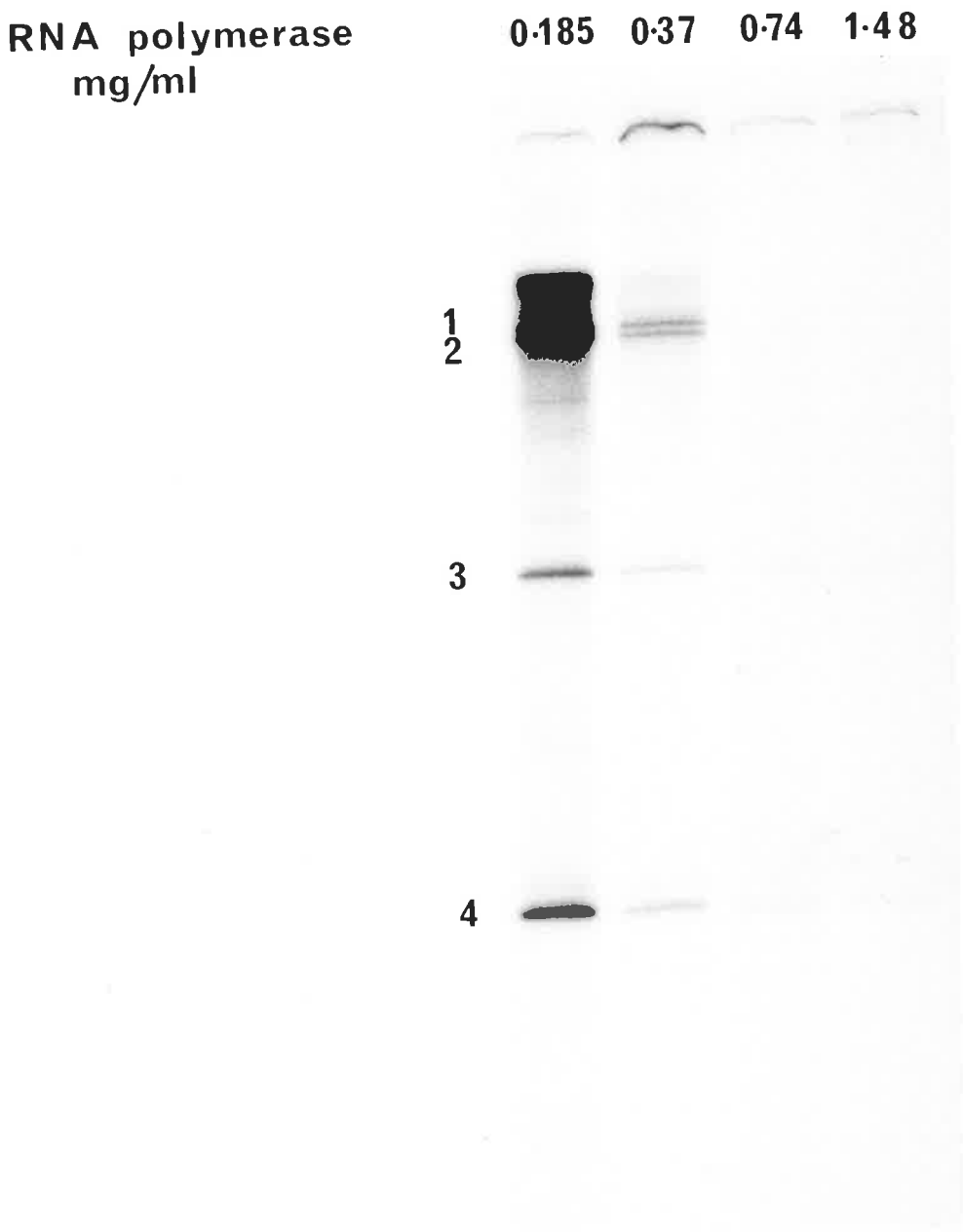
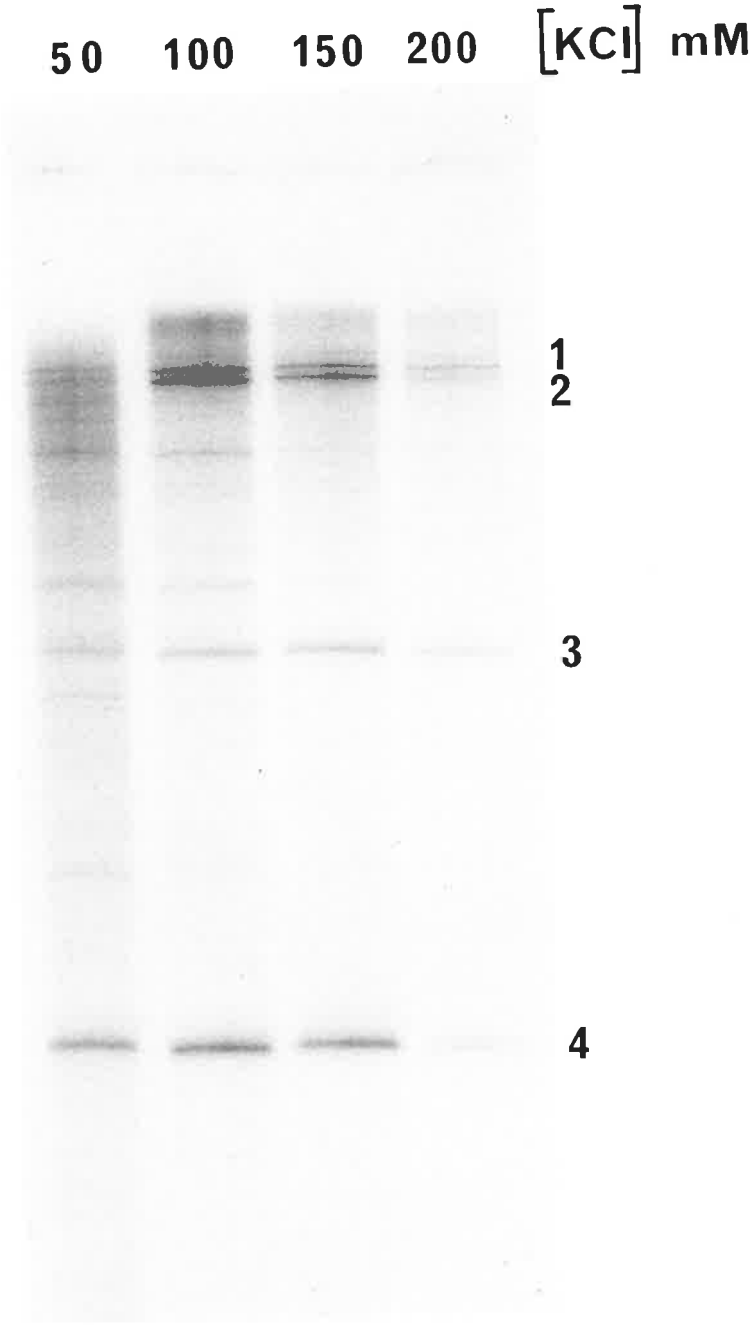

RNA polymerase
mg/ml
0·185
0·37
0·74
1·48
1
2
3
4

FIGURE 3.3

VARYING THE CONCENTRATION OF KCl

2.5 µg 186 DNA was transcribed with 0.185 µg/µl RNA polymerase whilst varying the concentration of KCl as indicated above the autoradiogram. The transcripts were separated on a 16 x 16 x 0.05 cm 5% acrylamide/8 M urea gel containing TBE.


50
100
150
200
[KCl] mM
1
2
3
4

decrease in the intensity of the four major 186 transcripts and an increase in the background transcription level at a KCl concentration of 50 mM compared with 150 mM. All future transcriptions contained 150 mM KCl.

Figure 3.4 shows 186 *in vitro* transcripts generated under optimal RNA polymerase and KCl concentrations. Resolution of the bands is maximized by loading the RNA in formamide loading buffer onto a 20 x 40 x 0.025 cm sequencing gel. There are four major transcripts, designated Bands 1, 2, 3 and 4.

3. Sizing of 186 *in vitro* transcripts

The four major 186 *in vitro* transcripts were sized on methyl mercury/agarose gels (Bailey and Davidson, 1976). Under the complete denaturation conditions of methyl mercury, it was possible to use DNA as size markers as RNA size markers were unavailable. The DNA used was a *Hae*III digest of ϕX174 and a *Hin*dIII digest of λ. A plot of $\sqrt{\text{molecular weight}/2}$ versus log mobility (Lehrach *et al.*, 1977) was prepared for the DNA size markers and a line of best fit drawn by eye as shown in Figure 3.5. The sizes of the 186 transcripts were calculated from this line and as such are only approximate. The sizes are as follows:

Band 1 1,540 bases

Band 2 1,450 bases

Band 3 590 bases

Band 4 290 bases.

4. Molar ratio of the transcripts

Once the sizes of the transcripts were known their molar ratios could be determined. 186 DNA was transcribed

FIGURE 3.4

186 *IN VITRO* TRANSCRIPTS

2.5 µg 186 DNA was transcribed *in vitro* at a KCl concentration of 150 mM and 0.185 µg/µl RNA polymerase. The transcripts were electrophoresed on a 20 x 40 x 0.025 cm 5% polyacrylamide gel containing TBE and 8 M urea. The xylene cyanol (XC) was run to the bottom.

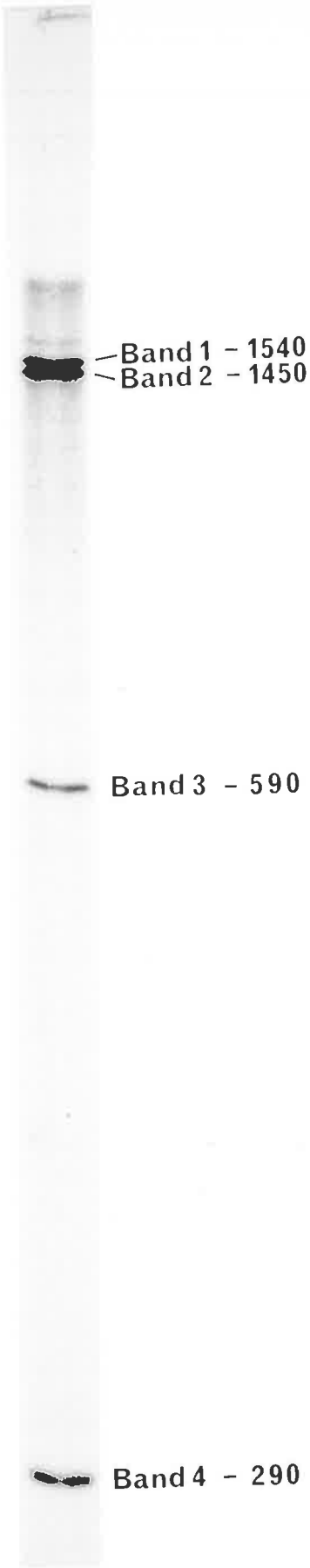
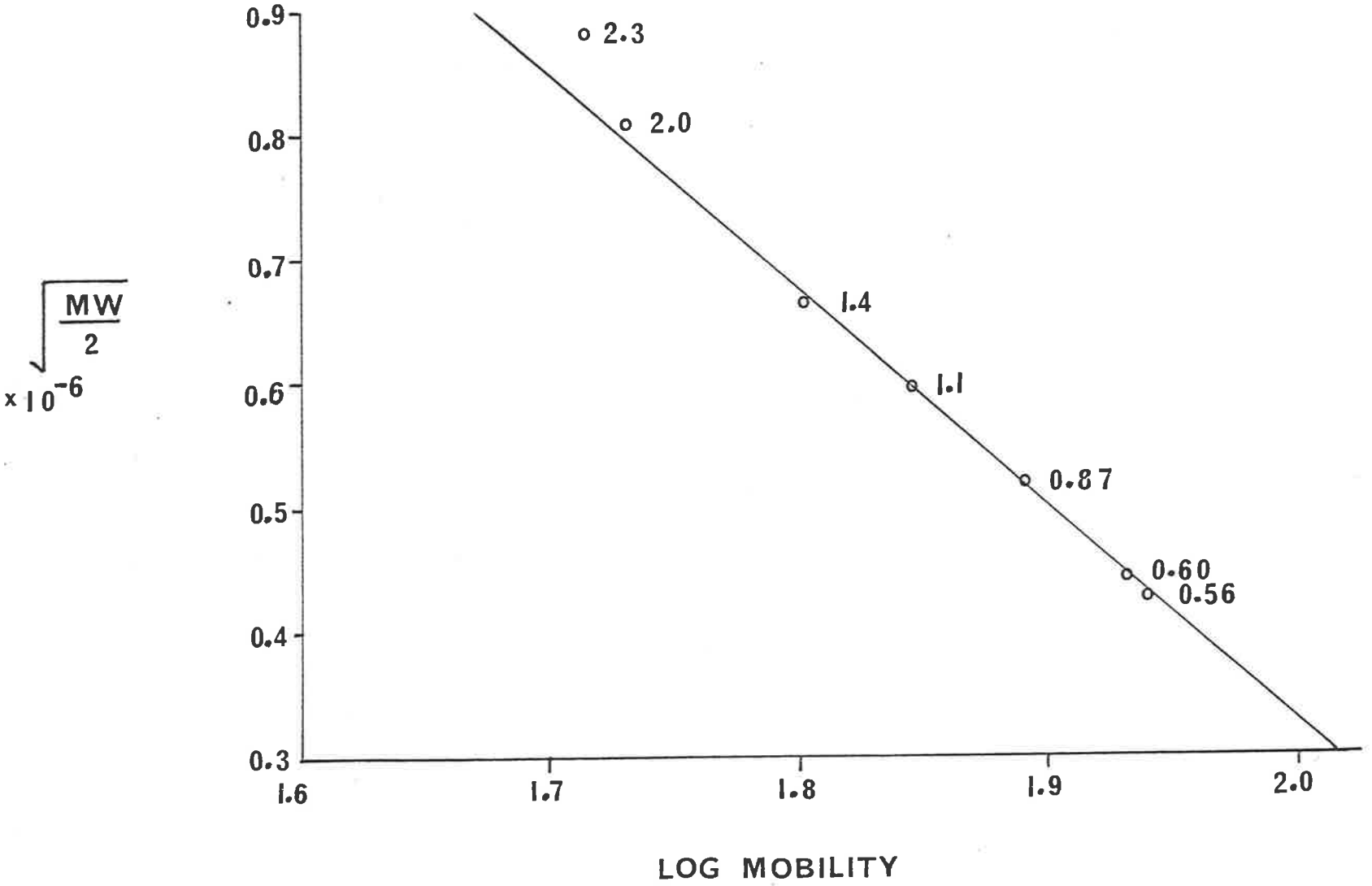

Band 1 - 1540
Band 2 - 1450
Band 3 - 590
Band 4 - 290

FIGURE 3.5

STANDARD CURVE FOR THE SIZE DETERMINATION OF 186 TRANSCRIPTS

186 *in vitro* transcripts were electrophoresed along with a *Hae* III digest of ϕX174 DNA and a *Hind* III digest of λ DNA on methyl mercury gels. A plot of $\sqrt{\text{molecular weight}/2}$ versus log mobility is shown. The molecular weight is divided by 2 because DNA was used as standard size markers. A line of best fit was drawn by eye and the sizes of the 186 *in vitro* transcripts estimated.

The sizes of the DNA markers are indicated in kb. The sizes of the ϕX174 and λ markers in base pairs are: ϕX174: 1,353, 1,078, 872, 603; λ: 2,322, 2,028, 564.


$\sqrt{\frac{MW}{2}} \times 10^{-6}$
0.9
0.8
0.7
0.6
0.5
0.4
0.3
2.3
2.0
1.4
1.1
0.87
0.60
0.56
1.6
1.7
1.8
1.9
2.0
LOG MOBILITY

under optimal conditions and the transcripts separated by electrophoresis. The bands were excised and the gel slices Cherenkov counted in a Packard scintillation spectrometer. From the counts the molar ratios were calculated.

band	1	2	3	4
gel slice c.p.m.	4,398	18,089	6,364	12,339
RNA size (base pairs)	1,540	1,450	590	290
molar ratio	1	4.4	3.8	14.9

Band 4, the smallest of the transcripts was the most actively transcribed, being 3.4 times more abundant than Band 2 and almost 15 times more abundant than Band 1.

CHAPTER 4

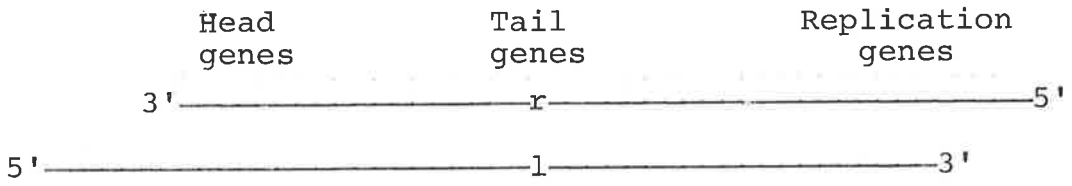
CHARACTERISATION OF TRANSCRIPTS:DIRECTION OF TRANSCRIPTION

This is the first in a series of Chapters (4 - 7) which comprise the second section of this thesis, i.e., the characterisation of the four 186 *in vitro* transcripts. These chapters record the sequence of experiments which initially permitted the approximate assignment of the transcripts to regions on the 186 genome, and later progressively focused on the precise location of the 5'-ends of the transcripts.

A. INTRODUCTION

To determine the direction of transcription, 186 *in vitro* transcripts were hybridized to separated strands of 186 DNA.

Kalionis and Egan (1981) oriented the strands of 186 DNA. By convention the 186 genetic map is drawn with head and tail genes to the left and early genes to the right.



The l strand is the template for left message.

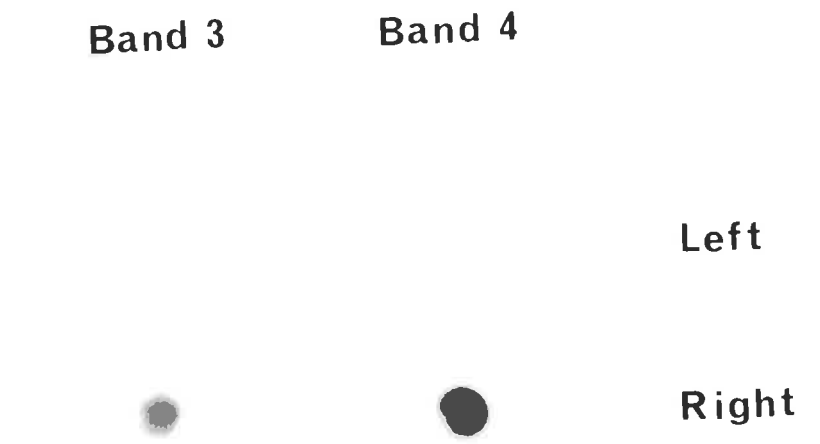
B. RESULTS AND DISCUSSION

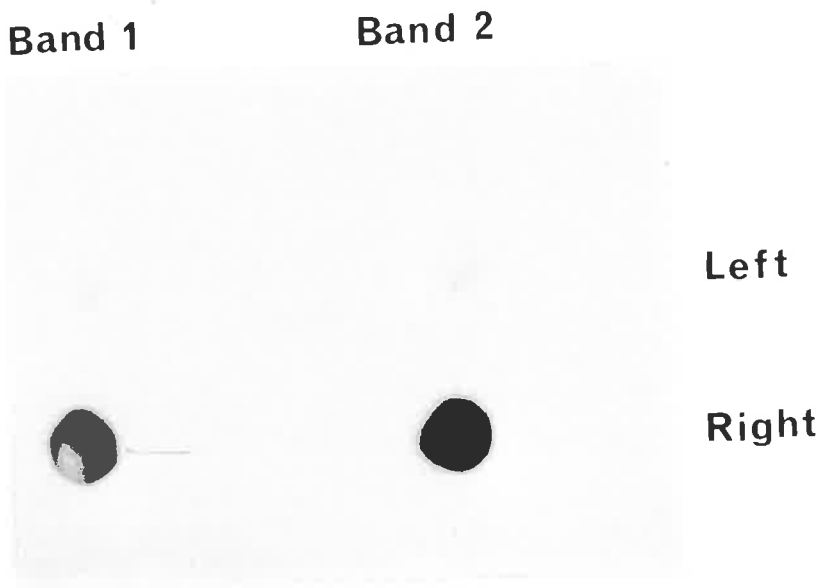
The l and r strands of 186 DNA were bound to nitrocellulose and probed with [α-^{32}P]-GTP labelled Bands 1, 2 3 and 4 which had been isolated from polyacrylamide/urea gels. The hybridization results are shown in Figure 4.1.

FIGURE 4.1

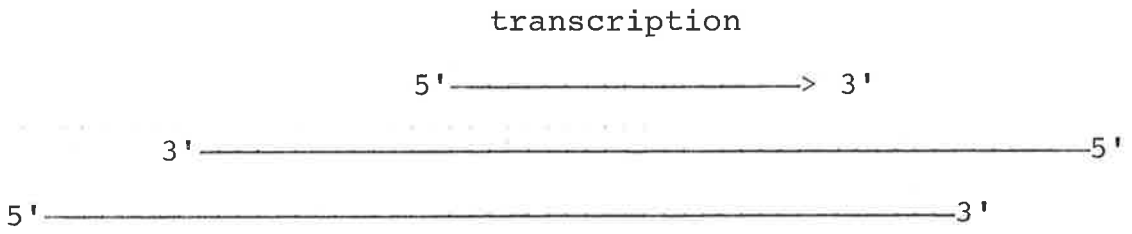
DIRECTION OF TRANSCRIPTION

2 µg each of 186 left and right strands were dot-blotted onto nitrocellulose. 2 100 µl *in vitro* transcription reactions were electrophoresed on 20 x 40 x 0.05 cm 5% polyacrylamide gels and Bands 1, 2, 3 and 4 excised and eluted. The transcripts were hybridized to the nitrocellulose strips and autoradiographed.


Band 3
Band 4
Left
Right


Band 1
Band 2
Left
Right

All the transcripts were synthesised from left to right relative to the 186 genetic map.



CHAPTER 5

CHARACTERISATION OF TRANSCRIPTS:

PRELIMINARY MAPPING OF *IN VITRO* TRANSCRIPTS

Chapter 5 is divided into two parts. Part 1 describes the preliminary mapping of the *in vitro* transcripts by hybridization to 186 restriction fragments. Part 2 describes the transcriptional analysis of cloned restriction fragments. Based on the hybridization data, plasmid clones were chosen which contained the DNA regions to which the transcripts hybridized, and these were transcribed *in vitro*. In addition, to confirming the hybridization data, specific regions of 186 DNA could be studied in isolation and the fragments containing promoters identified by virtue of their ability to initiate *in vitro* transcription. It was envisaged that this approach would assign an entire transcript to a restriction fragment and then by choosing restriction sites within the coding sequences, "run-off" transcription would position the 5'-ends of the transcripts.

PART 1. HYBRIDIZATION OF 186 *IN VITRO* TRANSCRIPTS TO 186 AND HY5 DNA DIGESTED WITH RESTRICTION ENDONUCLEASES

A. INTRODUCTION

186 early *in vivo* transcripts originate to the right of 64% (Finnegan and Egan, 1981) and it was the expectation that *in vitro* transcripts would also originate from within this region. The fragments generated upon a *Bam*HI digestion of 186 DNA would readily indicate the validity of this expectation as most of the 186 early region resides on one fragment of 11.2 kb (58.4-95.7%). The early region can be partitioned into well defined fragments using a *Pst*I digestion which would be useful for locating the transcripts more precisely.

However, the late region of 186 possesses numerous *Pst*I recognition sites which are not separated on agarose gels from the early fragments and would thus make the assignment of a transcript to a particular fragment impossible. To circumvent this problem, Hy5 , a P2/186 hybrid, containing the early genes of 186 and the morphogenic genes of P2, was used since a *Pst*I digest of Hy5 generates fragments readily separated by agarose gel electrophoresis.

A physical map of 186 showing the Hy5 crossover point is presented in Figure 5.1. The positions of *Bam*HI and *Pst*I restriction sites are indicated.

B. RESULTS AND DISCUSSION

Digested 186 and Hy5 DNA was separated by agarose gel electrophoresis and transferred onto nitrocellulose. The filters were then probed with [α-^{32}P]-GTP labelled 186 transcripts which had been isolated from preparative gels.

1. Location of Band 3

The hybridization pattern of Band 3 to *Bam*HI digested 186 DNA is presented in Figure 5.2. Band 3 hybridized predominently to the 11.2 kb fragment (58.4-95.7%) and also to the adjacent 4.2 kb fragment (186 cohesive ends; 95.7-10.0%). Thus Band 3 traversed the *Bam*HI site at 95.7%. The length of Band 3 is 590 bases and its direction of transcription is rightward, so it must originate to the left of 95.7% and is precluded by its size from traversing the cohesive ends.

Band 3 hybridized to the 6.3 kb annealed 186/P2 ends fragment of a *Pst*I digest of Hy5 (Figure 5.3). This result confirmed the location of this band to the righthand end of

FIGURE 5.1

PHYSICAL MAP OF 186

This map shows the positions of 186 genes and the crossover position of Hy5 in relation to the *Bam*HI and *Pst*I restriction sites.

The sizes of the fragments generated are given above the lines.

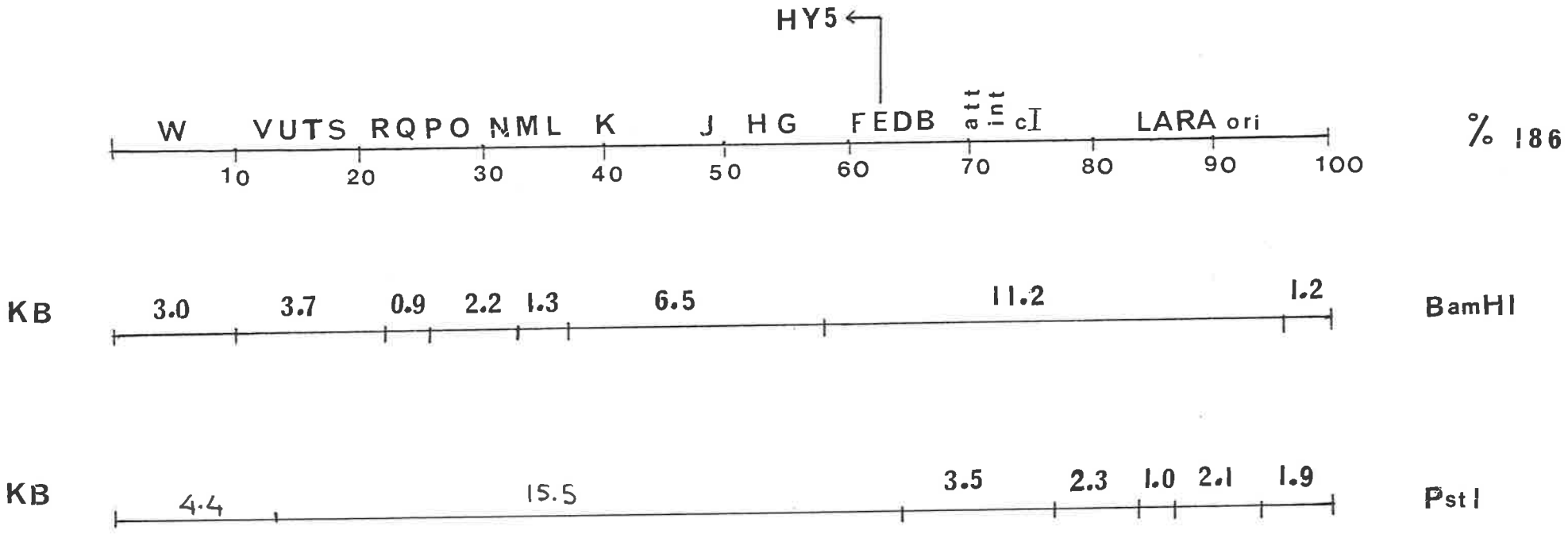
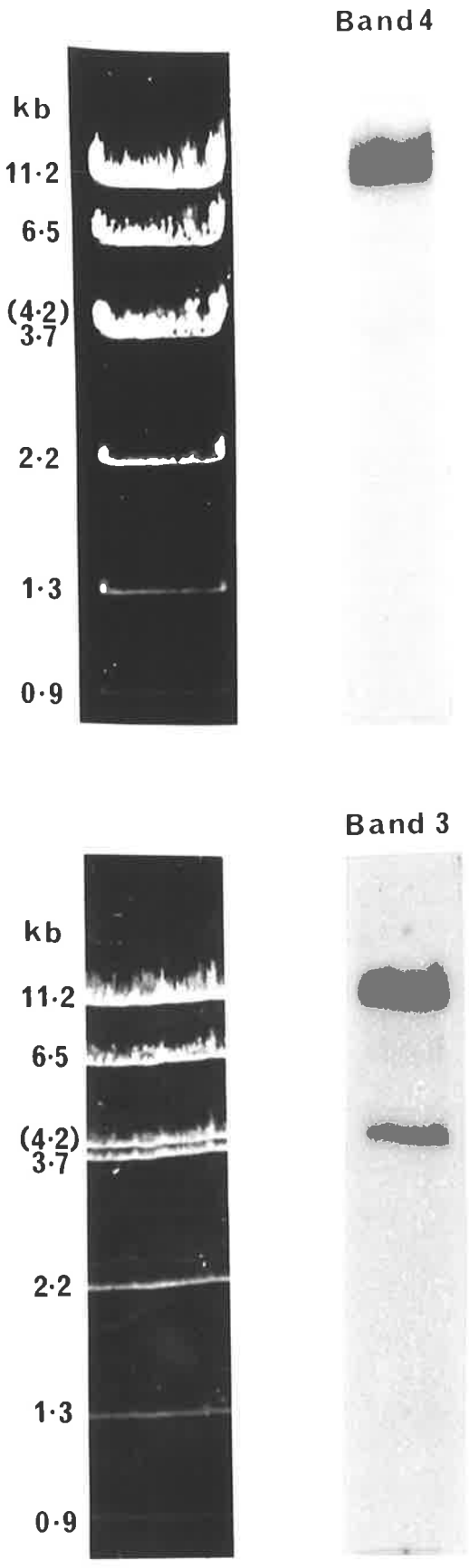

HY5
W VUTS RQPO NML K J HG FEDB att int cI LARA ori
% 186
10 20 30 40 50 60 70 80 90 100
KB
3.0 3.7 0.9 2.2 1.3 6.5 11.2 1.2
BamHI
KB
4.4 15.5 3.5 2.3 1.0 2.1 1.9
PstI

FIGURE 5.2

HYBRIDIZATION OF BANDS 3 AND 4 TO *Bam*HI DIGESTED 186

5 µg 186 DNA was digested with *Bam*HI, the fragments separated by 1% agarose gel electrophoresis and transferred onto nitrocellulose. Before transfer the gel was stained with 2 µg/ml ethidium bromide and photographed. 2 100 µl *in vitro* transcription reactions were loaded onto a preparative 5% polyacrylamide gel containing TBE and 8 M urea. The relevant bands were excised, eluted and hybridized to the nitrocellulose strips. The sizes of the restriction fragments are given next to the photograph of the gel. The size of the annealed ends fragment is in brackets.


Band 4
kb
11·2
6·5
(4·2)
3·7
2·2
1·3
0·9
Band 3
kb
11·2
6·5
(4·2)
3·7
2·2
1·3
0·9

the genome.

2. Location of Band 4

Band 4 was hybridized to a *Bam*HI digest of 186 (Figure 5.2). This transcript originated within the 11.2 kb fragment (58.4-95.7%) encoding the early genes.

The hybridization pattern of Band 4 to *Pst*I digested Hy5 is shown in Figure 5.3. Band 4 hybridized to the 3.5 kb fragment spanning 65.5% to 77.3% of the 186 genome, so reducing the size of the region within which the transcript originated from 11.2 kb to 3.5 kb.

It was anticipated that Bands 1 and 2 would also originate from the early region, so these transcripts were hybridized to a *Pst*I digest of Hy5 DNA directly, bypassing the hybridization to *Bam*HI digested 186.

3. Location of Band 1

Band 1 annealed to the same *Pst*I fragment as Band 3, i.e., the 1.9 kb rightmost fragment (Figure 5.4). Since Band 1 is synthesised in a rightward direction it must originate within this rightmost fragment with transcription towards the cohesive ends. It may be encoded entirely on this fragment or it may traverse the cohesive ends.

4. Location of Band 2

Figure 5.4 shows the hybridization of Band 2 to a *Pst*I digest of Hy5 DNA.

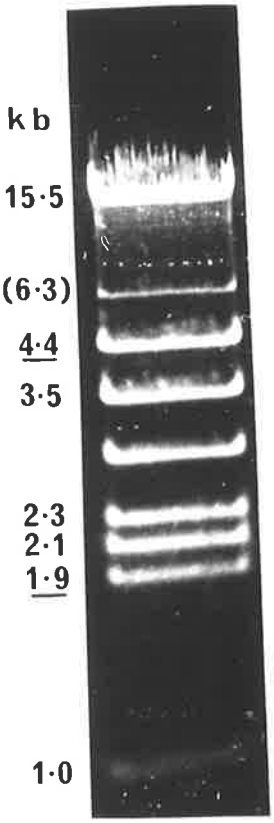
Band 2 hybridized to the adjacent 3.5 kb (65.5-77.3%) and 2.3 kb (77.3-84.6%) fragments indicating that this transcript originated within the 3.5 kb fragment and was synthesised in a rightward direction across the *Pst*I site at 77.3% into the 2.3 kb fragment. Band 2 also hybridized

FIGURE 5.3

HYBRIDIZATION OF BANDS 3 AND 4 TO *Pst*I DIGESTED HY5

5 µg Hy5 DNA was digested with *Pst*I, the fragments separated on a 1% agarose gel and transferred onto nitrocellulose. The DNA probed by Band 4 underwent end denaturation by heating at 80°C for 7 minutes prior to electrophoresis. The sizes of the end fragments are underlined and the size of the annealed ends is in brackets. The DNA probed by Band 3 was not end denatured. Experimental details are given in the legend to Figure 5.2.

Band 4



Band 3

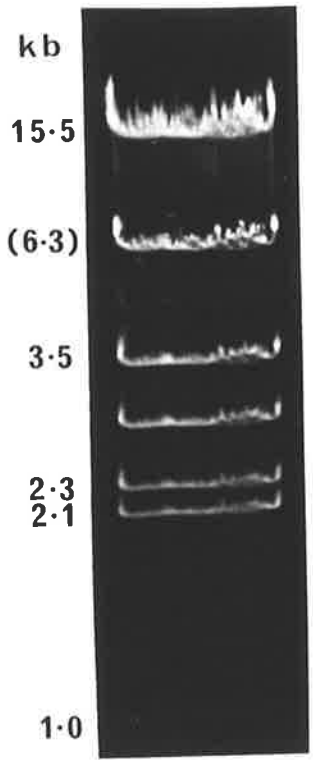
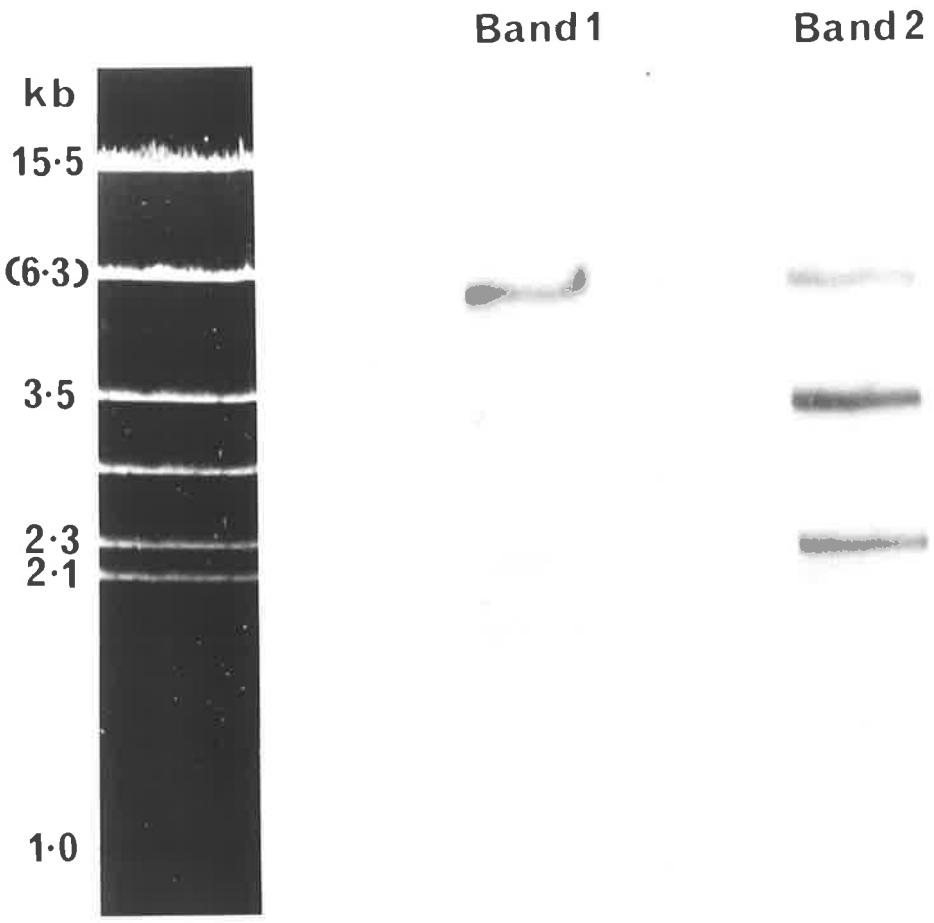


FIGURE 5.4

HYBRIDIZATION OF BANDS 1 AND 2 TO *PstI* DIGESTED HY5

Experimental details are exactly as described in the legend to Figure 5.3. Hy5 DNA was not end denatured before electrophoresis.


Band 1
Band 2
kb
15·5
(6·3)
3·5
2·3
2·1
1·0

to a lesser extent to the annealed ends. This observation is difficult to explain as the regions to which hybridization occurred are separated by more than 5 kb. There are two possible explanations for the observed hybridization to the rightmost *Pst*I fragment. Band 2 may be composed of two RNA species of similar size, which are not resolved on the gel system used, one originating from the rightmost end, the other originating in the 3.5 kb *Pst*I fragment. Alternatively, and probably most likely, the hybridization to the rightmost end results from contamination of Band 2 by the Band 1 transcript. These RNAs differ in size by less than 100 bases and remain in the log region of a 5% polyacrylamide/8 M urea gel, hence they are difficult to excise without cross-contamination. This is supported by the observed hybridization of Band 1 to the 3.5 kb and the 2.3 kb fragments after a long period of autoradiography.

A summary of the gross location of the four major transcripts is presented in Figure 5.5.

PART 2. *IN VITRO* TRANSCRIPTION OF CLONED 186 RESTRICTION FRAGMENTS

A. INTRODUCTION

Bands 2 and 4 hybridized to the 3.5 kb *Pst*I fragment (65.5-77.3%). Plasmid clones were available which contained sequences from this region, but the entire sequence was not represented as these clones were derived from 186 deletion phage.

Earlier attempts to clone the 3.5 kb fragment as well as

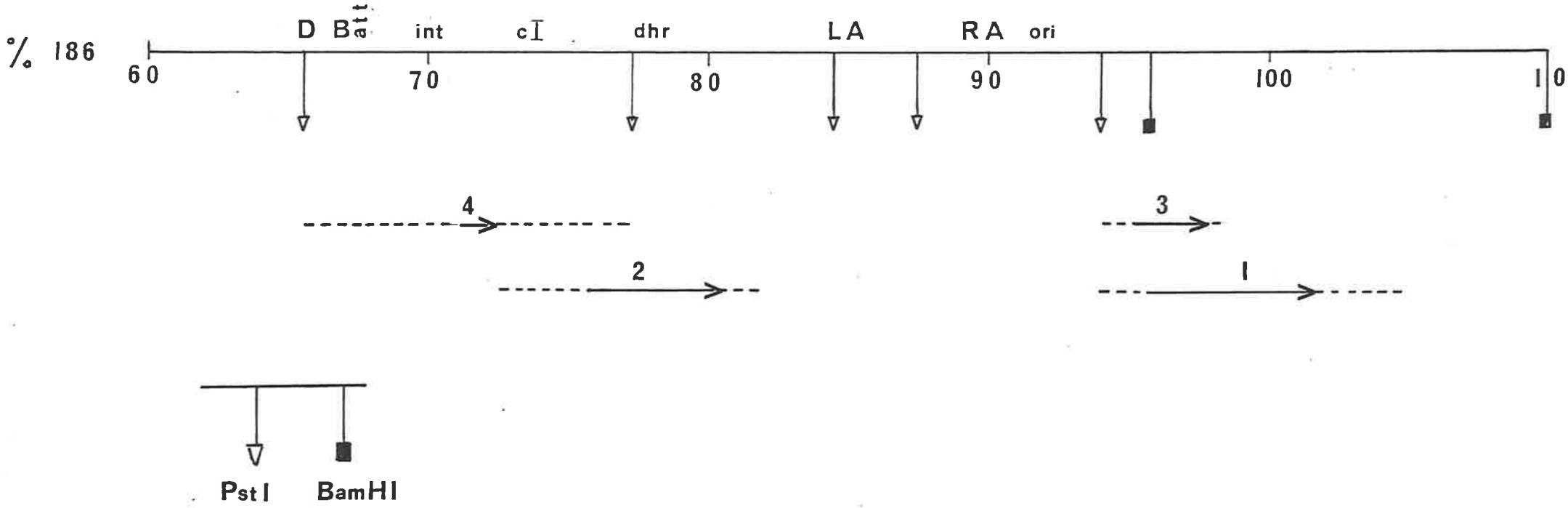
FIGURE 5.5

LOCATION BY HYBRIDIZATION OF THE 186 *IN VITRO* TRANSCRIPTS

The 186 *in vitro* transcripts located by hybridization are shown below the 186 physical map. The map is drawn from 60% across the cohesive ends to 10%.

The diagram is drawn to scale and the transcripts are within the fragments to which they hybridized.

The dashed lines represent the 5'- or 3'-extremeties of the transcripts.


% 186
D
B att
int
cI
dhr
LA
RA
ori
60
70
80
90
100
110
4
3
2
1
PstI
BamHI

an attempt during this work to clone the *Xho*I-*Bgl*II fragment (67.6-79.6%) into plasmids were unsuccessful. Unstable plasmids were produced, probably due to the presence of the *int* gene on these fragments whose product would mediate recombination. An apparent enigma exists because the 3.5 kb *Pst*I fragment can be cloned stably into M13mp7 (B. Kalionis, pers. commun.).

186Δ2 isolated by Dharmarajah (1975) has a 1.8 kb deletion from 67.9% to 74.2% and a 0.4 kb insertion in the *Pst*I-*Bgl*II fragment (77.3-79.6%). 186Δ1 also isolated by Dharmarajah (1975) has been shown, within the accuracy of heteroduplex mapping (R. O'Connor, pers. commun.) and restriction analysis (B. Kalionis, pers. commun.) to have the same deletion as 186Δ2 which removed the *att* site, *int* and *c*I genes.

Finnegan and Egan (1981) cloned the 2.7 kb *Pst*I fragment from 186Δ2 into pBR322 (the sizes have been modified since Finnegan and Egan by sequence analysis). H.R. Richardson (pers. commun.) cloned the *Xho*I-*Bgl*II fragment (67.6-79.6%) from 186Δ1 into pKC7, thus both of these clones contained sequences from the 3.5 kb *Pst*I fragment to which Bands 2 and 4 hybridized.

Bands 1 and 3 hybridized to the right of the *Pst*I site at 87.5%. A plasmid clone containing this region was not available so I cloned the rightmost *Eco*RI fragment from 92.0% which extends across the cohesive ends to 2.3% and named this clone pEC200.

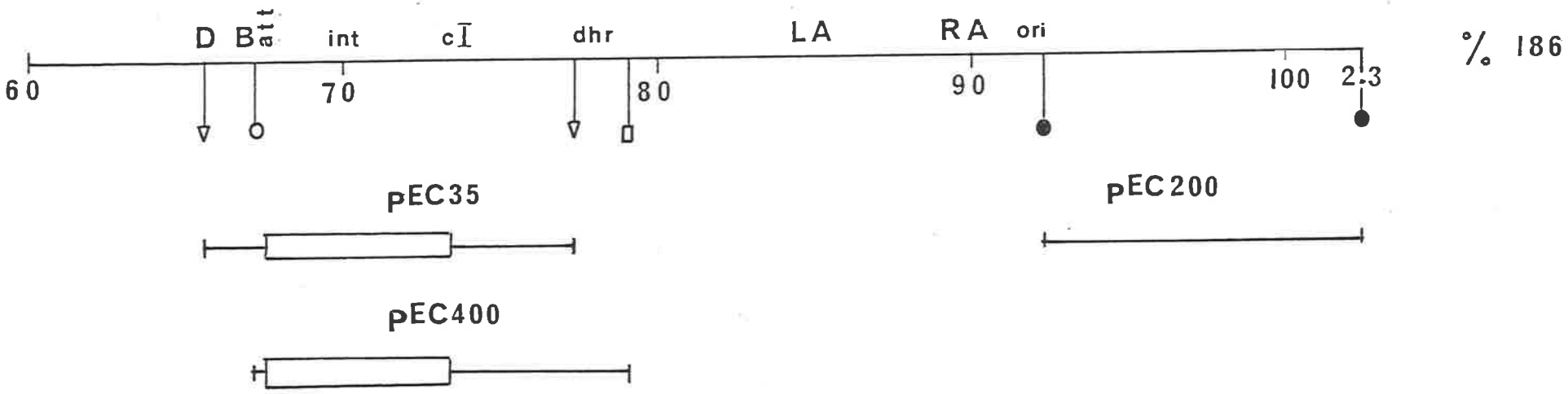
The 186 fragments cloned into plasmids and used in *in vitro* transcription experiments are shown in Figure 5.6.

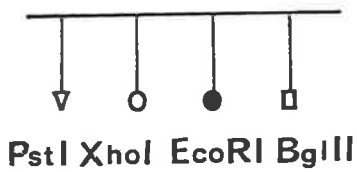
FIGURE 5.6

MAP SHOWING THE 186 FRAGMENTS CLONED TO GIVE pEC35, pEC400, pEC200

The position of the 1.8 kb deletion has been determined by sequence analysis (B. Kalionis, pers. commun.) and is represented by a box. The sites for *Pst*I, *Eco*RI, *Xho*I and *Bgl*II are shown to the right of 60%, across the cohesive ends to 2.3%.

The fragments cloned to give the plasmids pEC35, pEC400 and pEC200 are indicated below the map.


D
B
att
int
cI
dhr
LA
RA
ori
60
70
80
90
100
23
% 186
pEC35
pEC200
pEC400


PstI XhoI EcoRI BglII

B. RESULTS AND DISCUSSION

To avoid fusion transcripts, the plasmids were digested to release the cloned fragment. This mixture of linearised plasmid and 186 restriction fragment was then transcribed *in vitro*. The parent plasmid digested with the same restriction enzymes was transcribed as a control to discern plasmid transcripts from phage transcripts.

1. Bands 1 and 3

a) Cloning of 92.0-2.3% (pEC200)

The cohesive ends of 186 were ligated prior to digestion with *Eco*RI. The three *Eco*RI fragments produced (2.3-13.3% , 3.3 kb; 13.3-92.0%, 23.6 kb; 92.0-~~0.7~~ 2.3%, 3.1 kb) were "shot-gun" cloned into pBR325, which had been digested with *Eco*RI and phosphatased to prevent recircularisation of the plasmid. The 23.6 kb fragment was too large to be cloned into a plasmid so the only recombinants would contain the 3.3 kb or the 3.1 kb inserts.

Detection of recombinants containing the 3.1 cohesive end insert could be accomplished by screening for those clones unable to rescue genes V and W which are encoded on the 3.3 kb fragment. Recombinant cells were grown in L broth plus ampicillin to stationary phase and used to seed a plate lawn. 186 W*amber* and V*amber* lysogens were spotted onto the lawns, and those recombinants which showed no clearing in the spots further tested for the 3.1 kb cohesive ends insert.

Plasmids were prepared by the alkaline/SDS method and the size of the insert determined by digestion with *Eco*RI. Figure 5.7(a) shows an *Eco*RI digestion of three of the recombinant plasmids. All three contained a plasmid band of

FIGURE 5.7

RESTRICTION ANALYSIS OF THE COHESIVE ENDS INSERT

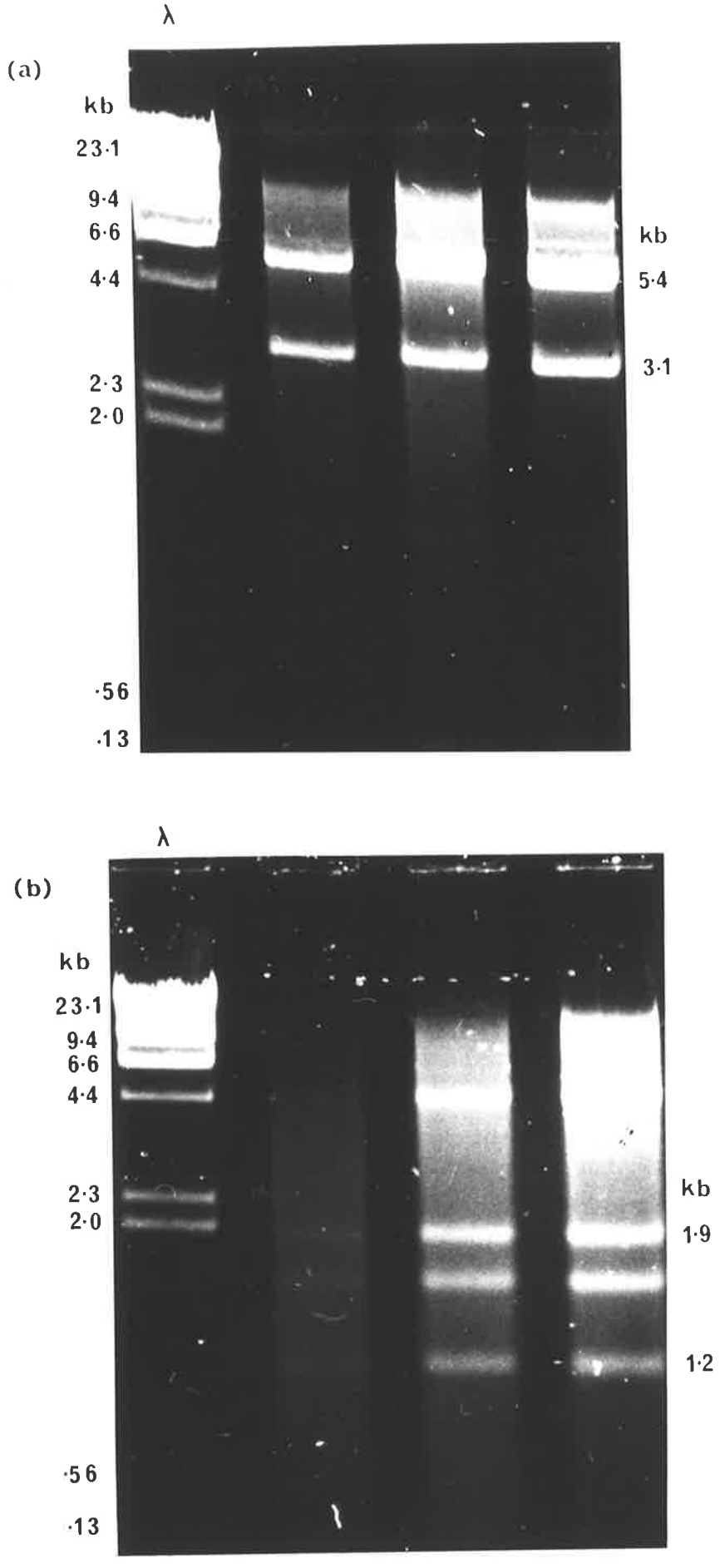
a) *Eco*RI digest of pEC200

The sizes of the 5.4 kb linearised plasmid and the released insert are given next to the photograph.

b) *Eco*RI/*Bam*HI double digestion of pEC200

The 3.1 kb cohesive ends insert gave two fragments of 1.2 and 1.9 kb when digested with *Bam*HI. The other 2 bands are plasmid.

The sizes of the λ *Hin*dIII markers are indicated beside the λ track. The separation was in 1% agarose containing TAE.


(a)
λ
kb
23·1
9·4
6·6
4·4
2·3
2·0
·56
·13
kb
5·4
3·1
(b)
λ
kb
23·1
9·4
6·6
4·4
2·3
2·0
·56
·13
kb
1·9
1·2

5.4 kb and a smaller insert. The smaller band ran between the 2.32 and 4.37 kb λ *Hin*dIII markers compatible with the expected size.

The 3.1 kb cohesive ends fragment has a diagnostic *Bam*HI site. The fragment sizes produced upon *Bam*HI digestion are shown in Figure 5.7(b). These fragments confirmed that the recombinants contained the cohesive ends insert. The clone was named pEC200.

b) *In vitro* transcription of pEC200

pEC200 and pBR325 were digested with *Eco*RI and transcribed *in vitro*. The transcription pattern is presented in Figure 5.8. Bands at the positions of 1 and 3 were visible in the pEC200 track but not in the pBR325 control track, indicating that these transcripts were of phage origin, and supporting the assignment of Bands 1 and 3 to the righthand end of the genome by hybridization.

2. Bands 2 and 4

a) *In vitro* transcription of 186Δ1

186Δ1 DNA was transcribed *in vitro* to determine the suitability of plasmids pEC35 and pEC400 for further characterisation of Bands 2 and 4. If promoter sequences for Bands 2 and 4 were deleted then these plasmids would prove unsuitable. Figure 5.9 shows an *in vitro* transcription pattern of 186Δ1. All four major transcripts were observed; the deletion did not alter the transcription pattern.

It is necessary to point out here that although all four *in vitro* transcripts were observed, these bands were not further characterised. The deletion would have no effect on Bands 1 and 3 as their sequences are remote to

FIGURE 5.8

IN VITRO TRANSCRIPTION OF pEC200

2.5 µg of *Eco*RI digested pEC200 and pBR325 were transcribed *in vitro* in a total reaction volume of 25 µl. 2.5 µg 186 was also transcribed. The transcripts were separated on 20 x 40 x 0.025 cm 5% polyacrylamide gels containing TBE and 8 M urea. The xylene cyanol (XC) was run to the bottom. The tracks are indicated above the autoradiograph and the bands identified at the side.

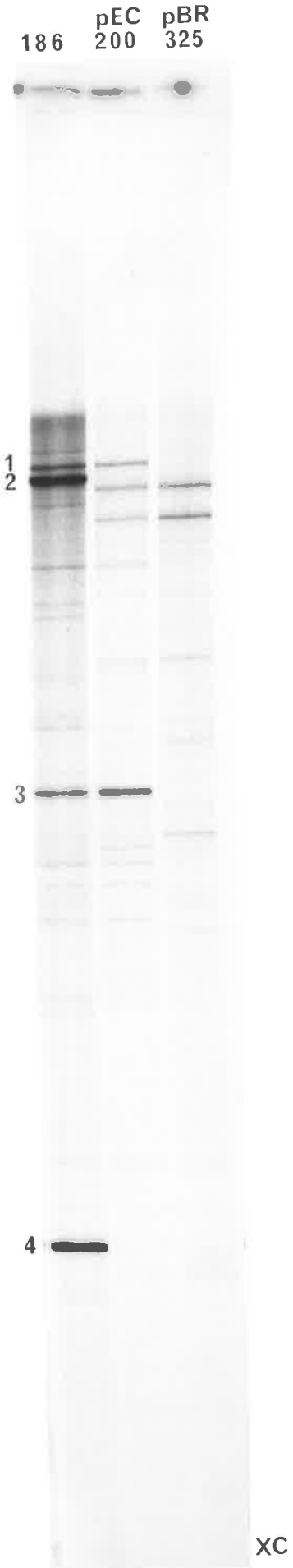

186
pEC 200
pBR 325
1
2
3
4
XC

FIGURE 5.9

IN VITRO TRANSCRIPTION OF 186Δ1

2.5 µg 186Δ1 DNA was transcribed *in vitro* and the transcripts electrophoresed as described in the legend to Figure 5.8.

186 186△1

the deleted DNA. However, the possibility that Bands 2 and 4 were fortuitous products of 186Δ1 transcription cannot be discounted. For instance, the rearrangement of DNA around the deleted region may have destroyed the Band 4 wild-type sequences but created a promoter or a terminator which fortuitously generates a transcript the same size as Band 4.

Whilst mindful of this reservation, transcription of pEC400 and pEC35 was initiated.

b) *In vitro* transcription of pEC400

*Xho*I/*Bgl*II digested pEC400 was transcribed *in vitro*. The resultant transcription pattern is shown in Figure 5.10. A transcript the size of Band 2 was evident as a phage encoded transcript.

c) *In vitro* transcription of pEC35

Figure 5.11 shows an *in vitro* transcription of pEC35 digested with *Pst*I. There were two prominent transcripts which did not have corresponding bands in the pBR322 track. One was at the position of Band 4. The larger band, subsequently sized on methyl mercury gels, was approximately 860 bases in length. There was no transcript at the position of Band 2. Since it had already been established that Band 2 traversed the *Pst*I site at 77.3% in a rightward direction, this 860 base transcript was a candidate for a run-off product of Band 2.

A transcript the size of Band 4 was not observed when pEC400 was transcribed. Band 4 did appear upon transcription of pEC35, therefore this transcript must originate to the left of the *Xho*I site at 67.6% and is thus located entirely

FIGURE 5.10

IN VITRO TRANSCRIPTION OF pEC400

2.5 µg pEC400 and pKC7 were digested with *Xho*I and *Bgl*II and then transcribed *in vitro*. 186 DNA was also transcribed. Electrophoresis was as described in the legend to Figure 5.8.

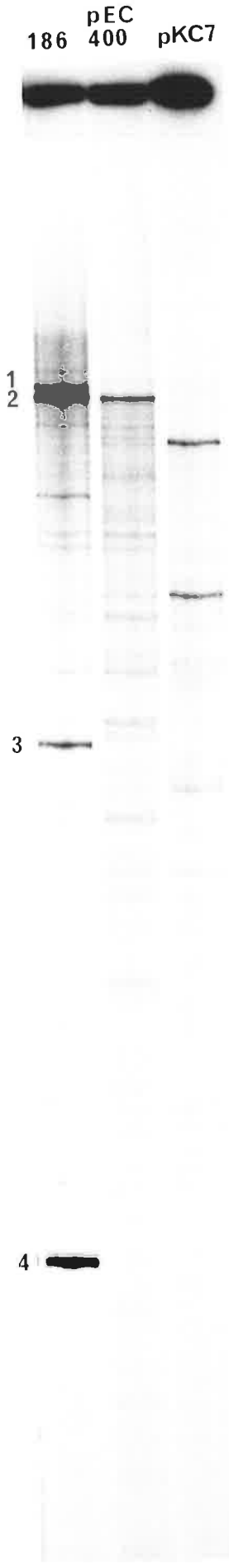
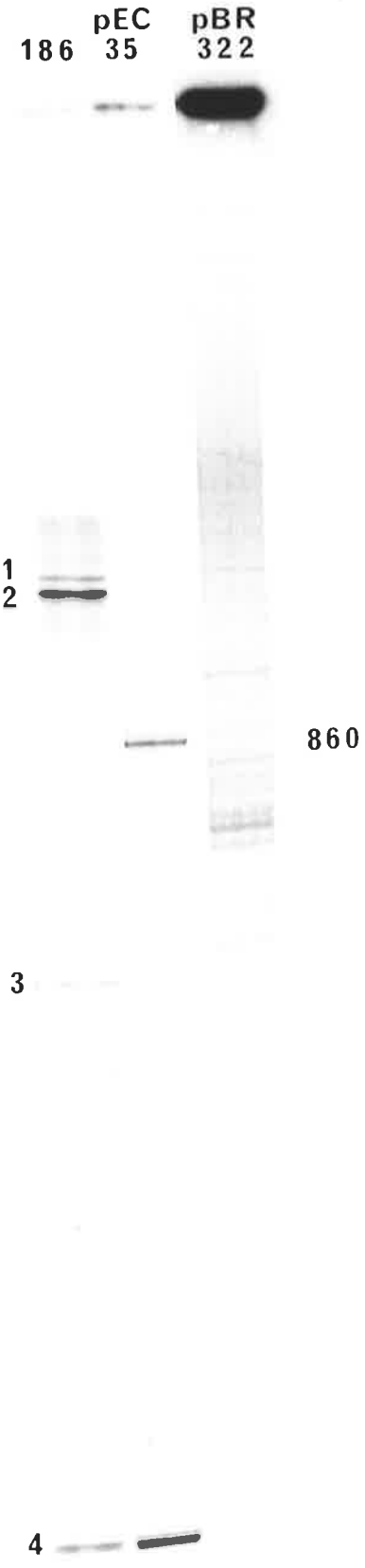

186
pEC 400
pKC7
1
2
3
4

FIGURE 5.11

IN VITRO TRANSCRIPTION OF pEC35

2.5 µg pEC35 and pBR322 were digested with *Pst*I and transcribed *in vitro*. 2.5 µg of 186 was also transcribed. Electrophoresis was as described in the legend to Figure 5.8.


186
pEC 35
pBR 322
1
2
860
3
4

on the 710 base pair fragment between the *Pst*I site at 65.5% and the left boundary of the deletion.

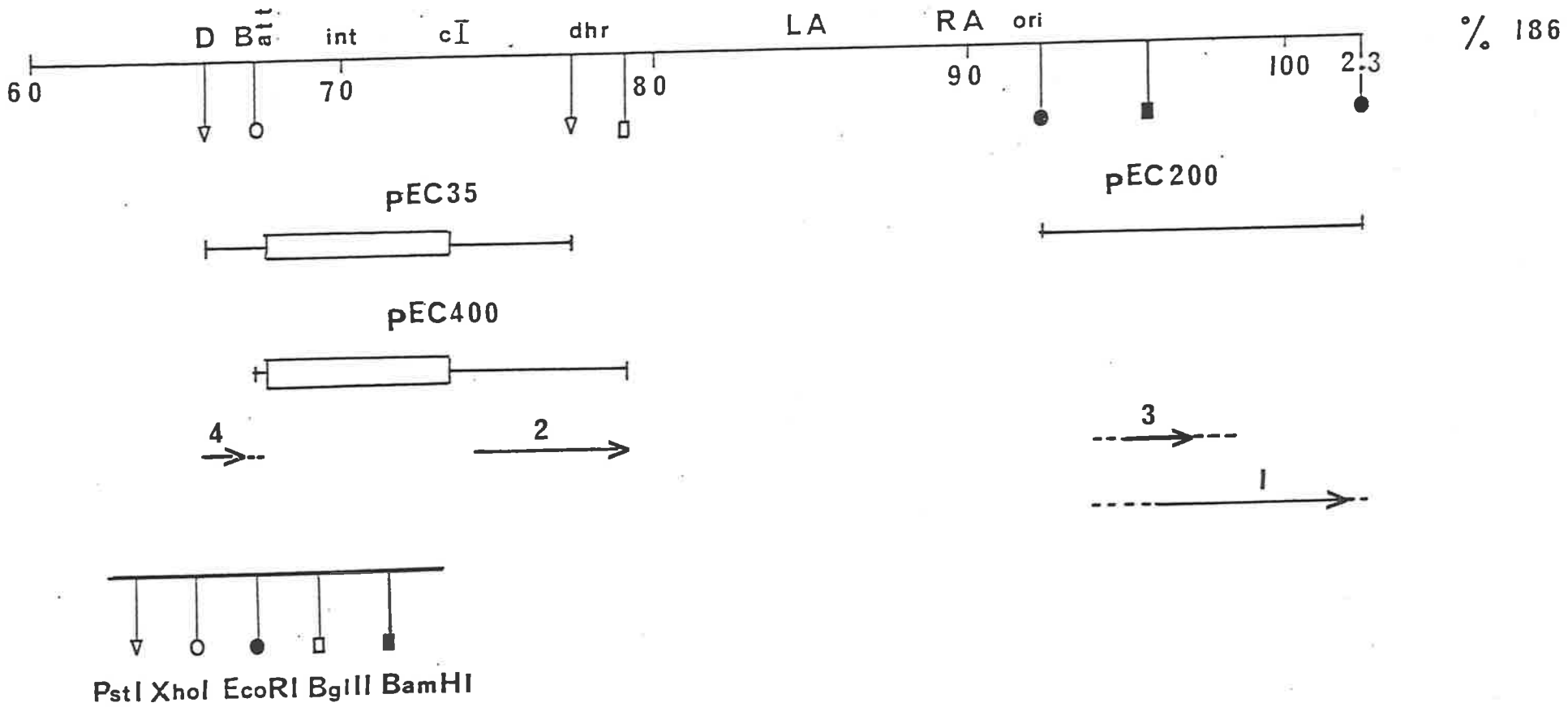
A diagram summarising what has thus far been established about the location of the 186 *in vitro* transcripts is presented in Figure 5.12.

FIGURE 5.12

SUMMARY OF TRANSCRIPT LOCATION

The diagram represents a summary of the location of the four 186 *in vitro* transcripts from the data thus far attained.

The 860 base transcript is considered here as a run-off product of Band 2, so the 5'-end of Band 2 has been drawn 860 bases to the left of the *Pst*I site at 77.3%. The positions of the 5'-ends of the other 3 transcripts are unknown as indicated by the dashed lines.


D
Bam
int
cI
dhr
LA
RA
ori
% 186
60
70
80
90
100
2.3
pEC35
pEC200
pEC400
4
2
3
1
PstI XhoI EcoRI BglII BamHI

CHAPTER 6

CHARACTERISATION OF TRANSCRIPTS:

FINER MAPPING OF *IN VITRO* TRANSCRIPTS

A. INTRODUCTION

Restriction fragments encoding each transcript in its entirety had been identified, and now, experiments designed to map the transcripts more precisely were initiated. Two approaches were pursued. One involved hybridization of the transcripts to subfragments of cloned 186 DNA, the other analysed the run-off products generated upon transcription of 186 DNA fragments which had been cleaved by restriction endonucleases within the transcript coding sequences. The sizes of the run-off products would locate the 5'-ends of the transcripts to a region within the accuracy of size estimation.

These techniques were applied to bands 1, 2 and 3. DNA sequence analysis of the region to which Band 4 had been assigned was in progress, so it was decided to locate Band 4 directly by sequencing as described in Chapter 7.

B. RESULTS AND DISCUSSION

1. Band 2

As discussed in the previous Chapter, the transcription data from pEC35 identified an 860 base transcript, which was considered a candidate for the run-off transcription product of Band 2. However, the transcription pattern of *Pst*I digested pBR322 (the parent plasmid of pEC35) had been inexplicably variable throughout this work. Indeed, the examples of pBR322 transcription in Figures 5.11, and 6.2 are rather different, regardless of being produced from the same plasmid preparation, albeit a month apart. This variation was not apparent with other plasmid templates although some differences, particularly in the levels of

background transcription were evident depending on the age of the template and of radioactive label.

Due to this variability in pBR322 transcription, hybridization of the 860 band to 186 sequences was performed to verify that the band was of phage origin.

a) Hybridization of the 860 base transcript to phage derived DNA

The 3.5 kb *Pst*I fragment (65.5%-77.3%) had been cloned into the M13 vector mp7 by B. Kalionis (this laboratory). This fragment was excised and separated from the replicative form by 1% low gelling temperature agarose gel electrophoresis. The fragment was extracted and then digested with *Sal*I (73.0%), *Sac*I (70.2%), *Kpn*I (69.1%) and *Xho*I (67.6%). Figure 6.1(a) shows a cleavage map of the 3.5 kb fragment. The fragments were separated on a 2% agarose gel and the DNA transferred to nitrocellulose. The 860 band from pEC35 transcription was hybridized to the filter. The result is shown in Figure 6.1(b). The 860 base transcript was phage derived and hybridized to the right hand end 1290 base pair fragment, a location expected for a run-off product of Band 2.

b) Run-off transcription of Band 2

If the 860 base RNA was a Band 2 run-off transcript, then 186 DNA digested with *Pst*I should produce a similar sized product.

To ascertain that 186 DNA digested with restriction enzymes was a suitable template for *in vitro* transcription, the DNA was firstly digested with *Hin*dIII. Figure 6.2 shows a *Hin*dIII cleavage map of 186 DNA. The four major *in vitro*

FIGURE 6.1

HYBRIDIZATION OF THE 860 RUN-OFF TRANSCRIPT TO THE 3.5 KB *Pst*I FRAGMENT (65.5 to 77.3%)

a) The positions of the *Sac*I, *Sal*I, *Kpn*I and *Xho*I restriction sites are shown for the 3.5 kb *Pst*I fragment from 65.5-77.3%. The sizes of the fragments generated are indicated in base pairs.

b) 2 100 µl *in vitro* transcription reactions of pEC35 digested with *Pst*I were electrophoresed on a preparative 5% polyacrylamide/8 M urea gel. The 860 base band was excised, eluted and hybridized to a nitrocellulose filter on which the 3.5 kb *Pst*I fragment digested with *Sac*I, *Sal*I, *Kpn*I and *Xho*I had been transferred.

(a)


PstI
XhoI
KpnI
SacI
SalI
PstI
630
430
340
860
1290

(b)

860 band

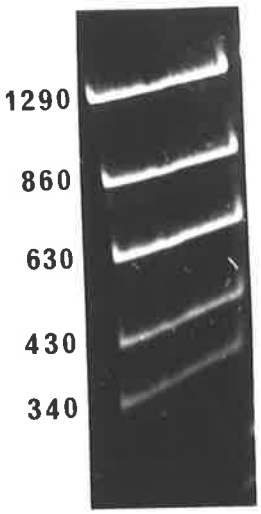
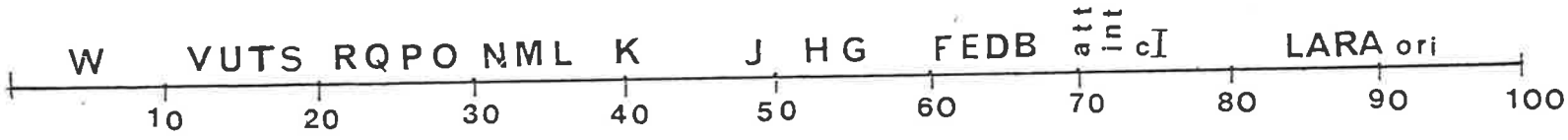

1290
860
630
430
340

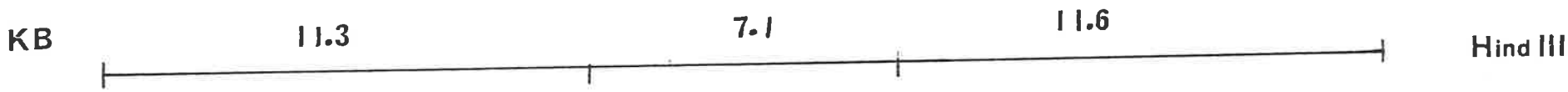
FIGURE 6.2

PHYSICAL MAP OF 186 SHOWING *Hin*dIII CLEAVAGE SITES

The sizes in kb of the fragments generated upon *Hin*dIII cleavage are indicated above the line.


W
VUTS
RQPO
NML
K
J
HG
FEDB
att
int
cI
LARA
ori
10
20
30
40
50
60
70
80
90
100

% 186


KB
11.3
7.1
11.6
Hind III

transcripts were preserved when *Hin*dIII digested 186 was transcribed as shown in Figure 6.3. A transcription pattern of 186 DNA digested with *Hin*dIII and *Pst*I is also shown in this figure. A prominent band at 860 bases was evident along with concomitant disappearance of Band 2. Thus Band 2 initiates approximately 860 base pairs to the left of the *Pst*I site at 77.3% and extends rightward, terminating about 590 base pairs to the right of 77.3%. Its promoter lies to the right of the Δ1 deletion. There are 1020 base pairs from the right boundary of the deletion to the *Pst*I site at 77.3% (B. Kalionis, pers. commun.), enough to accomodate a transcript of 860 bases and its accompanying promoter sequences.

2. Bands 1 and 3

It was determined, by hybridization to *Bam*HI digested 186 DNA, that Band 3 traversed the *Bam*HI site at 95.7%. So a *Bam*HI digestion of 186 DNA was the obvious choice for analysing run-off transcription of Band 3. No such information was available for Band 1, it was known only to originate to the right of 94.0%, within the 1.9 kb rightmost *Pst*I fragment. To facilitate the choice of an appropriate restriction enzyme for generating a run-off product of Band 1, this transcript was more precisely mapped within pEC200 by hybridization to subfragments of this clone. To confirm the earlier hybridization data, Band 3 was also hybridized to the subfragments of pEC200.

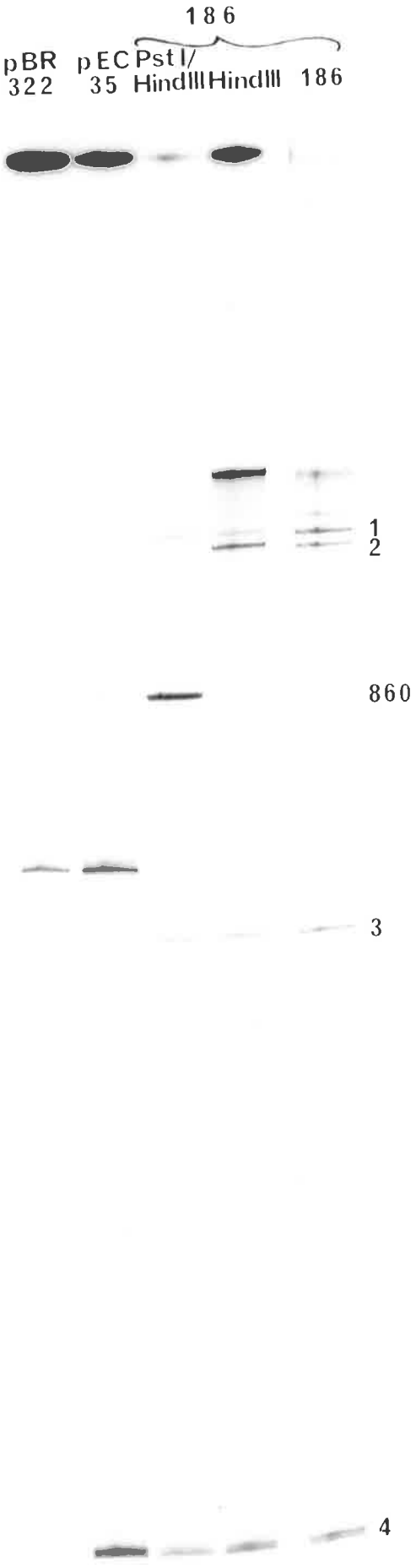
a) Hybridization of Bands 1 and 3 to pEC200 digested with *Eco*RI/*Pst*I/*Bam*HI

A nitrocellulose filter containing an *Eco*RI, *Pst*I,

FIGURE 6.3

RUN-OFF TRANSCRIPTION OF BAND 2

2.5 µg of 186 DNA either digested with *Hin*dIII or with *Hin*dIII and *Pst*I were transcribed *in vitro*. 2.5 µg each of uncut 186 DNA and *Pst*I digested pEC35 and pBR322 were transcribed and run in adjacent tracks for comparison. Electrophoresis was on a 20 x 40 x 0.025 cm 5% polyacrylamide gel containing TBE and 8 M urea.


186
pBR 322
pEC 35
PstI/ HindIII
HindIII
186
1
2
860
3
4

*Bam*HI triple digest of pEC200 was probed with 186 derived Bands 1 and 3. The results are presented in Figure 6.4. Bands 1 and 3 both hybridized to the adjacent 650 and 1390 base pair fragments. Thus, Bands 1 and 3 both traverse the *Bam*HI site at 95.7% indicating that these transcripts share sequences. However, it is not known whether these transcripts have the same 5'-ends.

Run-off transcription should locate the 5'-ends of Bands 1 and 3 and in so doing, reveal the arrangement of shared sequences.

b) Run-off transcription of Bands 1 and 3

To generate run-off transcripts for both Bands 1 and 3, 186 DNA was digested with *Bam*HI and transcribed *in vitro*. Figure 6.5 shows the result. The disappearance of Bands 1 and 3 was concomitant with the appearance of a single transcript approximately 390 bases in length. No other new transcript was detected, so it was likely that Bands 1 and 3 share a 5'-initiation site.

Figure 6.6 summarises the data obtained thus far, on the location of the four 186 *in vitro* transcripts.

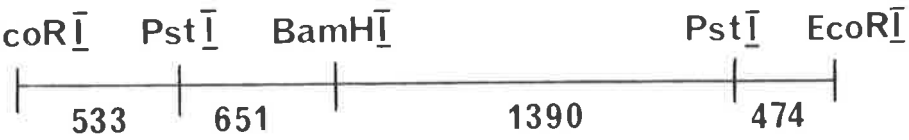
FIGURE 6.4

HYBRIDIZATION OF BANDS 1 AND 3 TO pEC200 DIGESTED WITH *Eco*RI, *Pst*I AND *Bam*HI

The diagram above the gel indicates the fragment sizes and positions of the *Bam*HI and *Pst*I sites on the *Eco*RI/*Eco*RI coehsive ends insert in pEC200.

2 100 µl *in vitro* transcription reactions of 186 were electrophoresed on a 20 x 40 x 0.05 cm preparative 5% polyacrylamide/8 M urea gel and Bands 1 and 3 excised, eluted and hybridized to a nitrocellulose filter onto which an *Eco*RI, *Bam*HI, *Pst*I digest of pEC200 had been transferred.

(a)


EcoRI
PstI
BamHI
PstI
EcoRI
533
651
1390
474

(b)

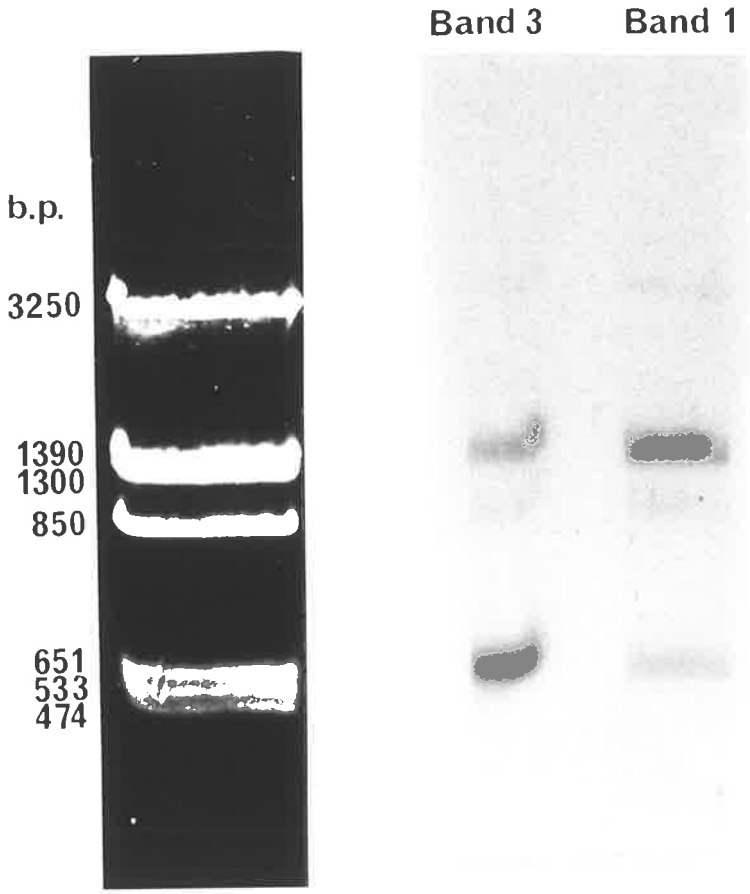

Band 3
Band 1
b.p.
3250
1390
1300
850
651
533
474

FIGURE 6.5

RUN-OFF TRANSCRIPTS OF BANDS 1 AND 3

2.5 µg of 186 DNA, undigested or digested with *Bam*HI was transcribed *in vitro*. Electrophoresis was as described in Figure 6.3. The tracks are indicated above the gel, and the transcripts at the side.

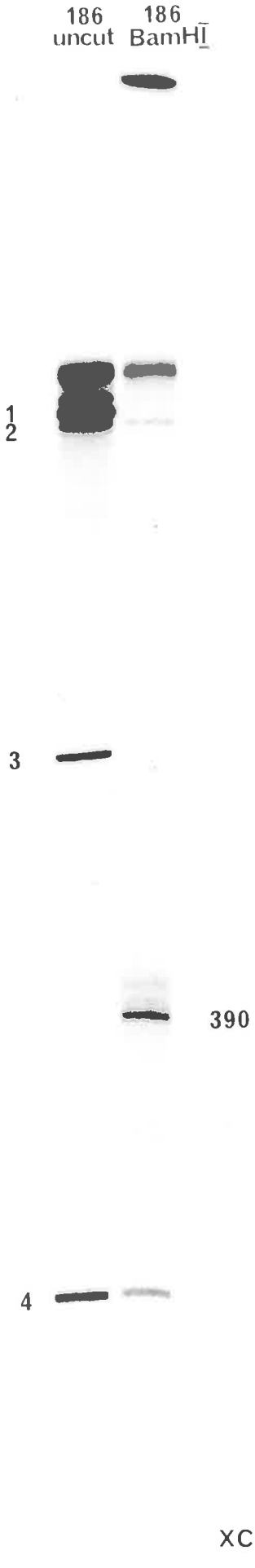

186
uncut
186
BamHI
1
2
3
390
4
XC

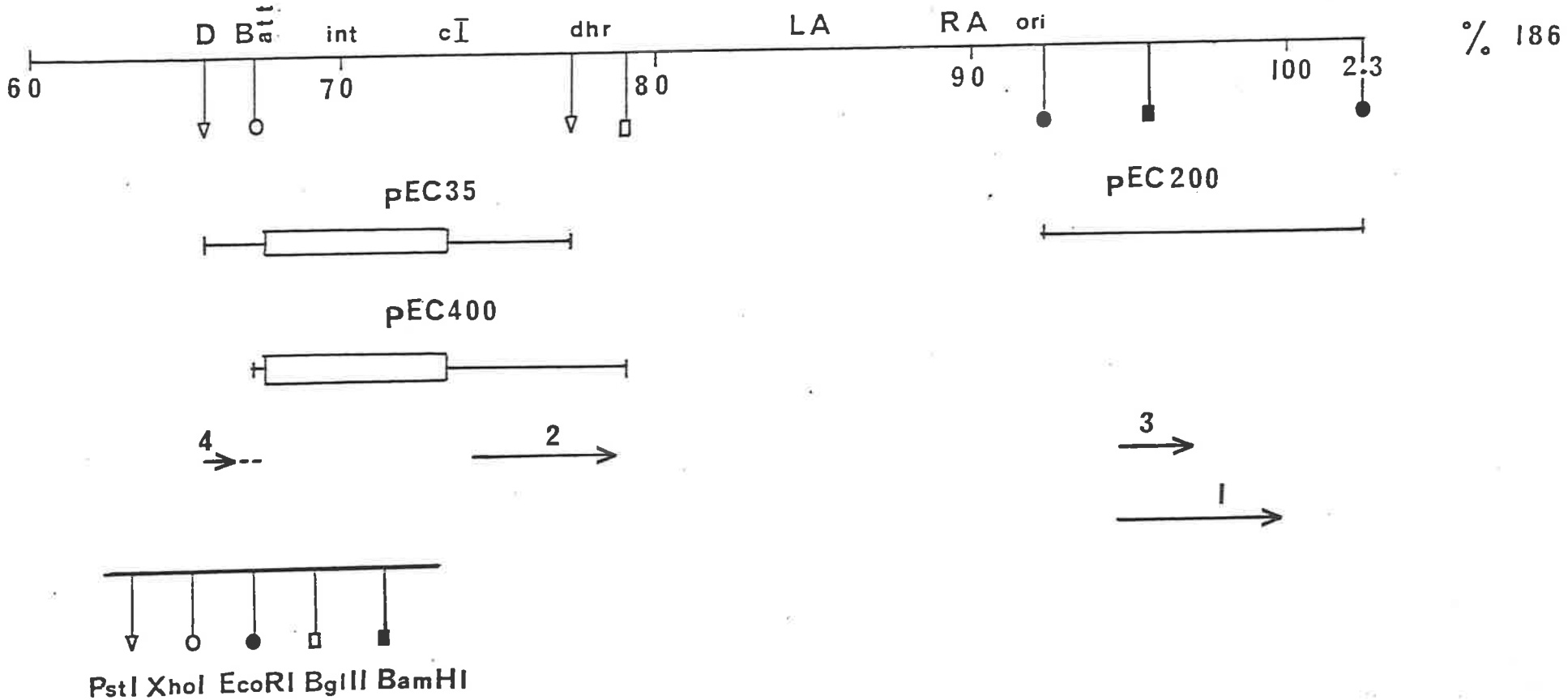
FIGURE 6.6

MAP SHOWING THE LOCATION OF 186 *IN VITRO* TRANSCRIPTS

The map is drawn from 60% to 2.3% of the 186 genome. It shows a summary of the transcription data from Chapters 3, 4, 5 and 6. The position of the Δ1 deletion is shown, indicated by a box in pEC35 and pEC400.

Band 4 is positioned within the 710 base pair fragment bounded on the left by a *Pst*I site and on the right by the left boundary of the deletion. The location of its 5'-end is unknown.

Bands 1 and 3 are drawn with the same 5'-end.


D
B
int
cI
dhr
LA
RA
ori
60
70
80
90
100
2.3
% 186
pEC35
pEC200
pEC400
4
2
3
1
PstI XhoI EcoRI BglII BamHI

CHAPTER 7

CHARACTERISATION OF TRANSCRIPTS:

SEQUENCE ANALYSIS OF *IN VITRO* TRANSCRIPTS

A. INTRODUCTION

The most definitive approach to the characterisation of RNA is by sequence analysis. The work presented in the previous chapters was undertaken prior to the availability of 186 DNA sequence information. 9.0 kb of the 186 genome has now been sequenced by cloning 186 fragments into M13 and then using the dideoxynucleotide chain termination technique of Sanger *et al.* (1977).

There is limited value in sequencing transcripts without the existence of the corresponding DNA sequence. Information about the control of transcription, for example, the location of promoters and repressor binding sites, is contained within the DNA sequence. However, there are two instances when the RNA sequence is valuable regardless of the existence of a DNA sequence with which to compare it. Firstly, once a transcript has been assigned to a genomic region by hybridization, DNA carrying nonsense mutations in the genes within this region, can be transcribed, and the mutant transcript sequenced. The identity of the gene encoding the transcript and hence the identity of the RNA would be determined by comparing the nucleotide sequence of the mutant and wild-type RNA's. Secondly, pertaining to this work, the sequence determination of Bands 1 and 3 would immediately corroborate or otherwise the run-off transcription data, indicating that Bands 1 and 3 have a common 5'-terminus.

The techniques available for the sequence determination of RNA rely on the presence of a fixed reference point within the RNA, e.g., a unique site for primed synthesis of

transcripts or a 5'- or 3'- radiolabelled terminus. The approach first envisaged for sequencing 186 RNA was to label the 5'-ends of the transcripts and then obtain their sequence by the enzymatic cleavage technique of Donis-Keller *et al.* (1977). The limiting factor in this approach was the efficiency of the 5'- labelling. Bands 3 and 4 were sequenced using this technique but the approach was unsuccessful for Bands 1 and 2.

Having DNA sequence data available for the regions to which Bands 1 and 2 had been assigned facilitated the choice of DNA fragments which were used as specific primers for sequencing by the dideoxynucleotide chain termination technique.

This chapter describes the sequencing of the 186 *in vitro* transcripts and the identification of Band 4 as the 186 B gene transcript.

B. RESULTS

[γ-^{32}P]-ribonucleotides may replace [α-^{32}P]-GTP as the labelled species in an *in vitro* transcription reaction, generating transcripts labelled at their 5'-terminii. [γ-^{32}P]-ATP and [γ-^{32}P]-GTP were generally available. D. Eckerman (BRESA) synthesised [γ-^{32}P]-UTP and [γ-^{32}P]-CTP especially. The four 186 *in vitro* transcripts were all visible when [γ-^{32}P]-ATP was included in the transcription reaction mix, signifying that they all initiated with an ATP. However, only Bands 3 and 4 were labelled sufficiently for sequencing.

1. Band 4

a) Sequence by enzymatic cleavage

3 100 µl *in vitro* transcription reactions were prepared using [γ-^{32}P]-ATP as the radiolabelled species. Band 4 was excised, eluted and sequenced by the enzymatic cleavage technique. Examples of 20% and 5% sequencing gels of Band 4 are presented in Figure 7.1. Band 4 was sequenced extensively with the view of comparing its sequence to Band 4 transcripts derived from B*amber* mutants (see below).

The enzymatic sequence of Band 4 persistently gave an ion front which obscured the initiating nucleotide. From [γ-^{32}P]-ATP labelling it was ascertained that Band 4 initiated with an ATP, therefore even though the first readable nucleotide was a G residue, the transcript must initiate with an ATP.

b) Identification of Band 4

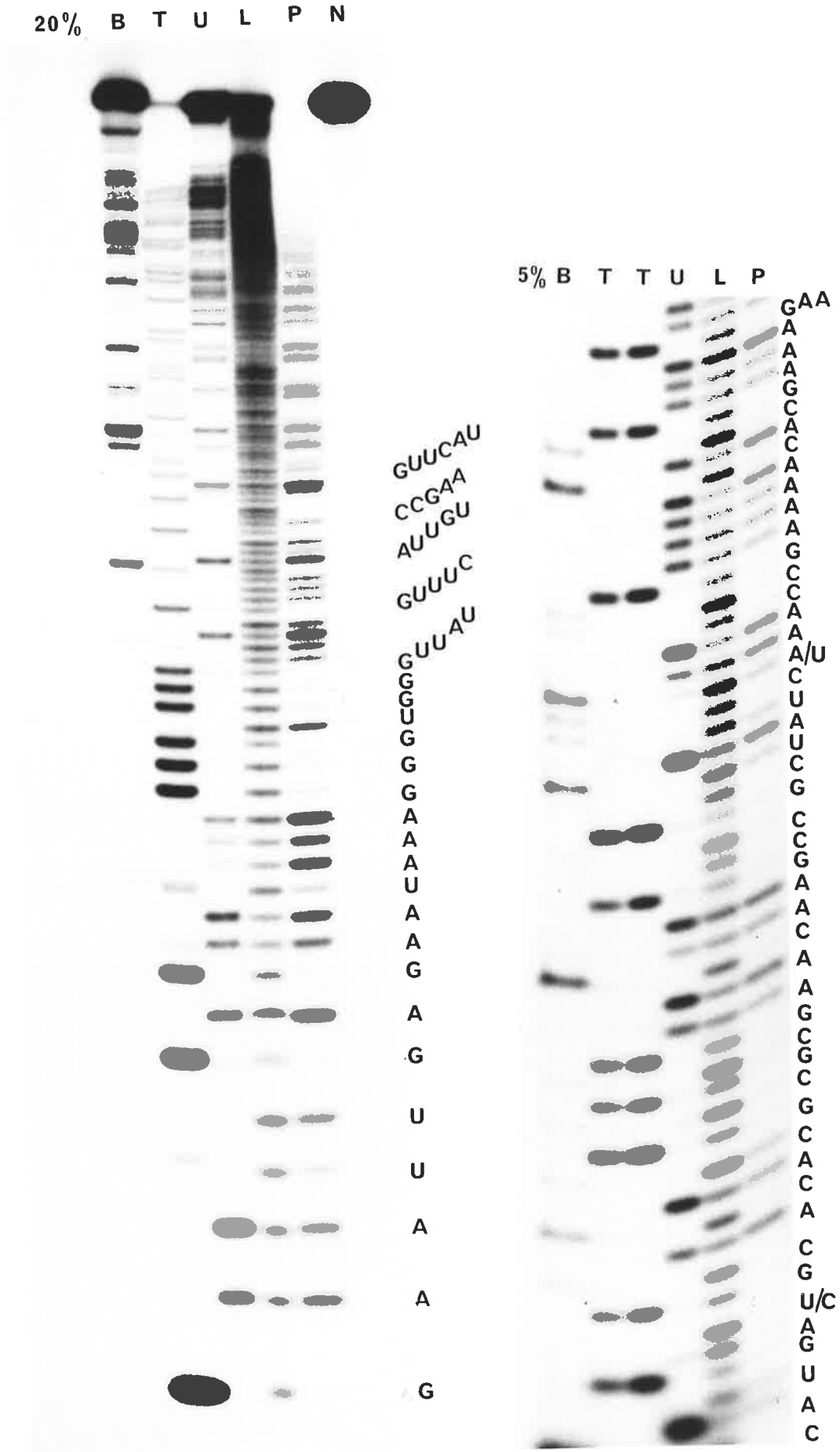
On comparison of the Band 4 RNA sequence with the 186 DNA sequence generated by B. Kalionis (pers. commun.), this transcript was found to originate from the region, defined from genetic studies, to encode the 186 late turn on gene, B.

Computer analysis of the DNA sequence predicted an open reading frame with an AUG initiation codon (see Chapter 9). Within this open reading frame, the positions of base pair changes which would generate amber codons were determined. The RNA sequences of Band 4 from amber mutants could then be scanned for these base changes. Band 4 was identified, firstly by confirming that it was the B gene transcript by sequencing the Band 4 RNA's generated upon transcription of 186 B*amber*

FIGURE 7.1

SEQUENCE BY ENZYMATIC CLEAVAGE OF BAND 4

[γ-^{32}P]-ATP labelled Band 4 was sequenced by the enzymatic cleavage technique of Donis-Keller *et al.* (1977) and the fragments separated on 20% or 5% polyacrylamide gels (20 x 40 x 0.05 cm or 20 x 80 x 0.05 cm) containing TBE and 8 M urea. The figure shows an example of each of these gels, the enzymes are written above the tracks and portions of the Band 4 sequence are written beside the sequence. Autoradiography was at -80°C with an intensifying screen.


20% B T U L P N
5% B T T U L P
GUUCAU
CCGAA
AUUGU
GUUUC
GUUAU
GGGUGGGAAAUAAGAGUUAAG
GAAAAAGCACAAAAAGCCAAAA/U
CUAUCGCCGAACAAAGCGCGCACACACGU/CAGUAC

mutants, and secondly, because interpretation of the RNA sequences of the B*amber* mutants was uncertain, by sequencing the DNA of B*amber* mutants.

i) RNA sequence of B*am*17 and B*am*57

DNA from B*am*17 and B*am*57 phage was transcribed *in vitro* with [γ-^{32}P]-ATP and the Band 4 transcripts isolated and sequenced enzymatically.

The nucleotide changes corresponding to the amber codons are indicated in Figure 7.2 next to the autoradiograms of the RNA sequencing gels. The amber codon for B*am*17 could not be read as such due to over digestion of the RNA by RNases U2 and *B. cereus*, but was recognized by the disappearance of a G residue comparable with the change in the B*am*17 DNA sequence (see below). B*am*57 showed an amber codon in phase with the computer predicted initiating AUG.

ii) DNA sequence of B*am*17 and B*am*57

Hocking and Egan (1981a) established the map order of 186 amber mutants. DNA from B*am*17 and B*am*57 phage was isolated and digested with *Pst*I and the 3.5 kb fragment (65.5%-77.3%) excised and extracted from low gelling temperature agarose. This fragment was further digested with *Xho*I and the 630 base pair fragment isolated from 2% low gelling temperature agarose and cloned into M13mp9 as described in Chapter 2, which had been doubly digested with *Pst*I and *Sal*I.

DNA from white plaque recombinants was propogated and sequenced as described in Materials and Methods. In Figure 7.3 is a sequencing ladder indicating the nucleotide changes responsible for the B*am*17 and B*am*57 mutations. The 186 amber

FIGURE 7.2

RNA SEQUENCE OF B*am*17 AND B*am*57

[γ-^{32}P]-ATP labelled transcripts from B*am*17 and B*am*57 DNA were sequenced enzymatically. The base change to give B*am*57 is indicated next to the gel. The tracks are indicated above the gels. The gels were both 5% polyacrylamide containing TBE and 8 M urea. The dimensions of the gel showing the B*am*57 mutation was 20 x 80 x 0.05 cm and the B*am*17 was 20 x 40 x 0.05 cm.

NOTE: Only the G residues are indicated on the B*am*17 gel. The number of nucleotides (N) between the G residues are shown as XN. Below the sequence obtained is aligned with the wild type sequence to emphasise the missing G residue.

400 410 420 430

AGTTGTACGTTTATGACAATGGAAACGATAGAGCG

$G_{2N}G_{3N}\ G\ _{5N}\ G\ _{5N}\ GG\ _{4N}\ G_{5N}\ G\ G$

B*am*17

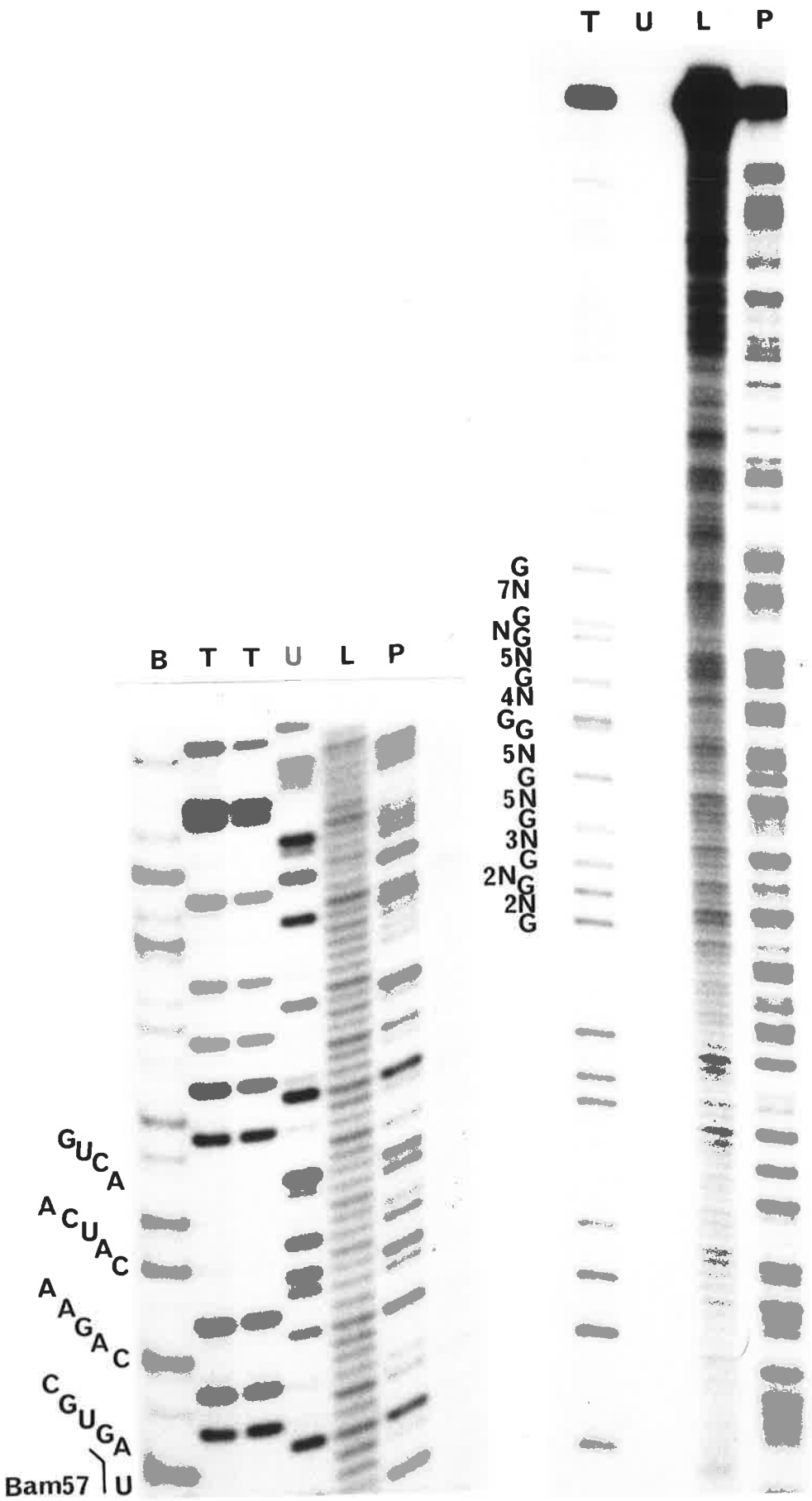
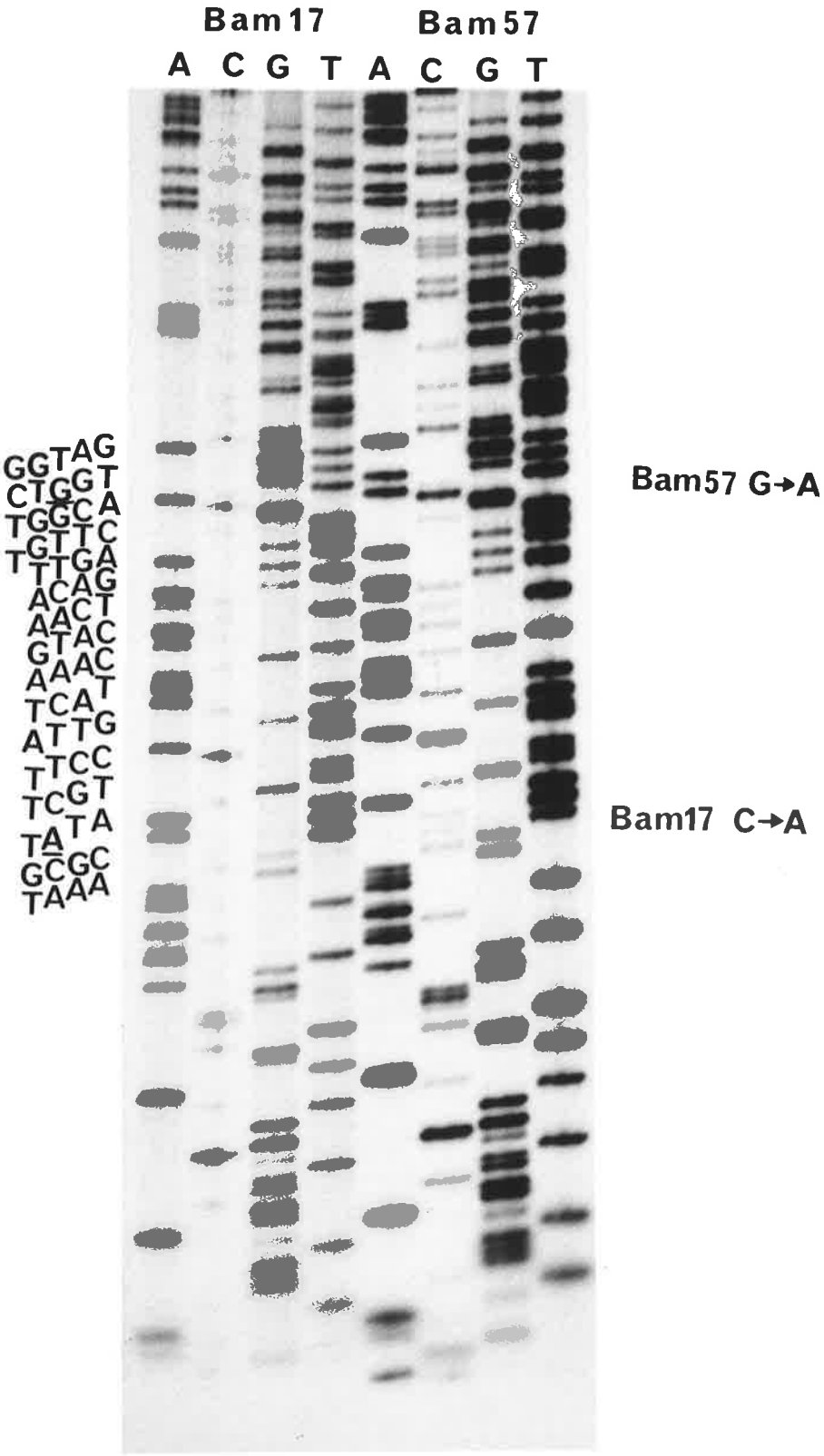

T U L P
B T T U L P
GUCA
ACUAC
AAGAC
CGUGA
Bam57 U
G
7N
G
N G
5N
G
4N
G G
5N
G
5N
G
3N
G
2N G
2N G

FIGURE 7.3

DNA SEQUENCE OF B*am*17 AND B*am*57

DNA sequencing ladders are shown depicting the base changes for the B*am*17 and B*am*57 mutations. The reactions were loaded onto 20 x 40 x 0.025 cm 5% polyacrylamide gels containing TBE and 8 M urea. The tracks are above the gel and the sequence is shown at the side. The incorporation of label in the C tracks was low, so the other tracks are over exposed.

The sequence is leftward from the *Xho*I site, hence is complementary to the RNA so the nucleotide changes which generate the amber codons in the RNA are seen here in the sense strand.


Bam17
Bam57
A C G T A C G T
Bam57 G→A
Bam17 C→A

DNA was sequenced leftward from the *Xho*I site at 67.6% towards the *Pst*I site at 65.5% and compared to the wild-type DNA sequence of the B gene (B. Kalionis, pers. commun.). B*am*57 was a $\begin{matrix} C & T \\ | & \rightarrow & | \\ G & & A \end{matrix}$ transition and B*am*17 was a $\begin{matrix} G & T \\ | & \rightarrow & | \\ C & & A \end{matrix}$ transversion.

2. Bands 1 and 3

a) Sequence by enzymatic cleavage of Band 3

Band 3 was labelled at its 5'-end with [γ-^{32}P]-ATP in an *in vitro* transcription reaction and then sequenced by the enzymatic cleavage technique. The result is shown in Figure 7.4. Due to underdigestion of the RNA by the RNase U2, there is some doubt in the assignment of A residues, however the sequence was sufficient to locate Band 3 precisely by comparison with the DNA sequence to the left of the *Bam*HI site at 95.7%.

The 5'-sequence of Band 3 was also determined by primer extension.

b) Sequence by primer extension of Bands 1 and 3

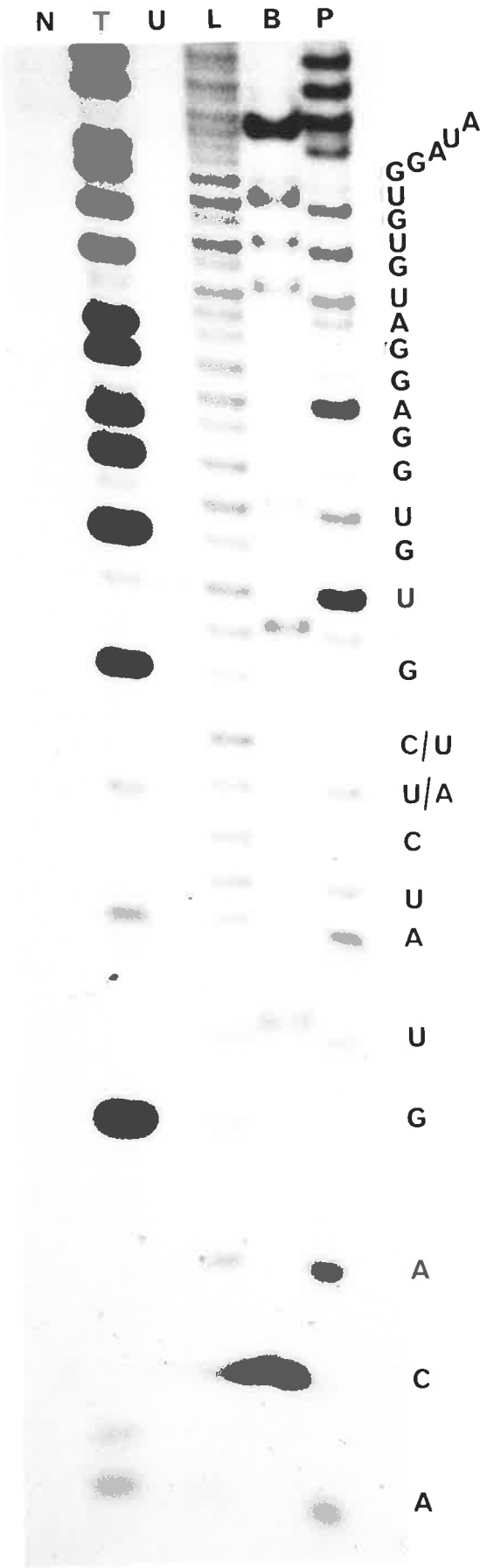
The sequence of 186 DNA has been determined to the left of the *Bam*HI site at 95.7% (A.V. Sivaprasad, pers. commun.). Data presented previously has shown that Bands 1 and 3 both cross the *Bam*HI site at 95.7%. A DNA primer was chosen to the left of the *Bam*HI site as illustrated in Figure 7.5. A single-stranded M13 clone of the *Pst*I to *Bam*HI (87.5%-95.7%) fragment of the same polarity as the RNA was supplied by A.V. Sivaprasad. A radio-labelled copy of the fragment was synthesised ~~and then the entire insert excised by digestion with *Eco*RI and *Hin*dIII which produced 2 fragments~~ the DNA digested with EcoRI and BamHI and then electrophoresed on LGTagarose s (see Figure 7.5). The 1.2 kb *Eco*RI-*Bam*HI (92.0%-95.7%) fragment was

FIGURE 7.4

SEQUENCE BY ENZYMATIC CLEAVAGE OF BAND 3

[γ-^{32}P]-ATP labelled Band 3 RNA was sequenced enzymatically and electrophoresed on 20% polyacrylamide gels (20 x 40 x 0.05 cm) containing TBE and 8 M urea. The tracks are marked as in Figure 7.1. Autoradiography was at -80°C with an intensifying screen.

The U2 RNase track was under digested.


N
T
U
L
B
P
GGAUA
GUGUGUAGGAGGUGUGUCU
C/U
U/A
C
U
A
U
G
A
C
A

excised from low gelling temperature agarose, extracted, redigested with *Hinf*I and *Hgi*AI and then loaded onto a 5% polyacrylamide non-denaturing gel. The 84 base pair primer was excised and eluted.

Unlabelled transcripts were synthesised in three 100 µl transcription reaction mixes by substituting the [α-^{32}P]-GTP with unlabelled GTP. The unlabelled transcripts were separated by electrophoresis. Bands 1 and 3 were excised using radioactive 186 transcripts run in adjacent tracks as markers. The RNA was eluted, then hybridized to the denatured 84 base primer and sequenced by extension of the primer. The sequencing ladders generated by Bands 1 and 3 are shown in Figure 7.5. Bands 1 and 3 have identical 5'-sequences, i.e., have the same initiation site.

Band 1 resulted as a readthrough at the Band 3 terminator. From the molar ratio of Bands 1 and 3, the termination efficiency was calculated to be 79%.

3. Band 2

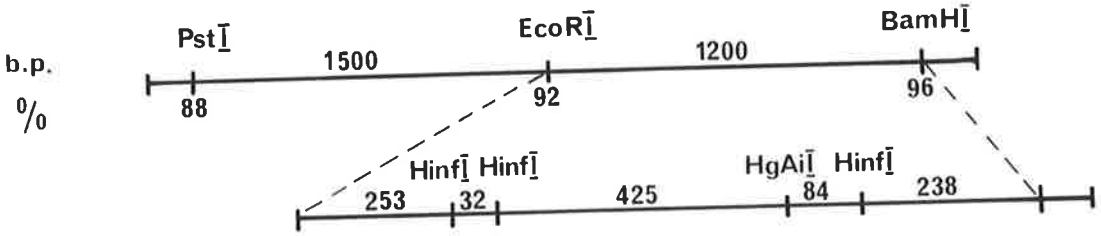
Band 2 has its 5'-end approximately 860 nucleotides to the left of the *Pst*I site at 77.3%. The DNA sequence in this region has been determined (B. Kalionis, pers. commun.). An appropriate primer was chosen by scanning the DNA sequence for restriction endonuclease recognition sites which would generate a primer of the desired length and position. Figure 7.6 shows a cleavage map of the region from which the primer was chosen. B. Kalionis supplied an M13mp7 *Hpa*II (73.7%-75.2%) clone from which a radioactive copy was synthesised. The 450 base pair *Hpa*II insert was excised by digestion with *Eco*RI and isolated from a 2% low gelling

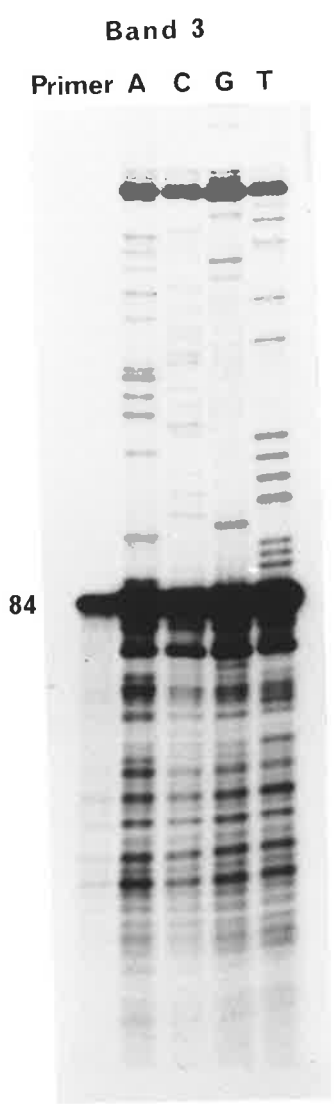
FIGURE 7.5

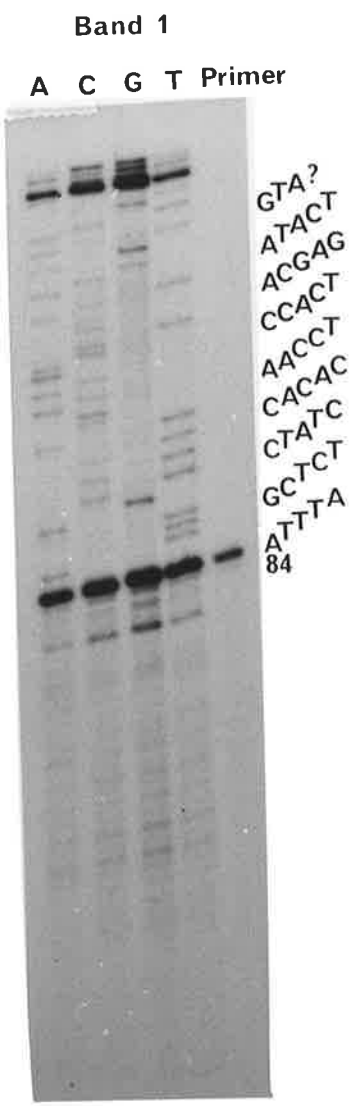
SEQUENCE BY PRIMER EXTENSION OF BANDS 1 AND 3.

A cleavage map of the cloned *Pst*I/*Bam*HI fragment from which the 84 base pair primer was chosen is shown above the sequencing gel. The gel was 5% polyacrylamide (20 x 40 x 0.025 cm) containing TBE and 8 M urea. The tracks are indicated above the gel. The size of the primer is indicated. Autoradiography was at room temperature.

NOTE: The initiating nucleotide of Bands 1 and 3 is obscured by an intense crossband which occurs as a consequence of the reverse transcriptase encountering the end of its template. Studies with [γ-^{32}P]-ATP indicate that both transcripts initiate with an ATP, therefore the crossband is interpreted as a T residue.


b.p.
%
PstI
EcoRI
BamHI
1500
1200
88
92
96
HinfI HinfI
HgAiI HinfI
253
32
425
84
238


Band 3
Primer A C G T
84


Band 1
A C G T Primer
84

temperature agarose gel. This fragment was further digested with *Hinf*I to produce the fragment sizes illustrated in Figure 7.6. The 118 base pair primer was hybridized to unlabelled Band 2 RNA and the primer extended using reverse transcriptase along with deoxy- and dideoxynucleotides. The sequence of the 5'-end of Band 2 is shown in Figure 7.6.

C. DISCUSSION

The reason for the low level of [γ-^{32}P]-ATP incorporation into Bands 1 and 2 cannot be completely explained by the low levels of these RNA's. The molar ratio of Band 2 is slightly higher than Band 3 and yet sufficient label was incorporated to enable the enzymatic sequencing of Band 3. Attempts to 5'-label Bands 1 and 2 by adding a labelled phosphate group to their 5'-terminii using T4 polynucleotide kinase were also unsuccessful, in that the specific activity of the RNA's was too low for any further manipulation.

I subsequently found that sequencing by extending a primer was a superior method as the RNases were susceptible to non-specific cleavage, particularly those of *B. cereus* and Phy M. Uncertainty in the interpretation of a sequence was disadvantageous especially when searching for amber codons.

In summary, all four of the 186 *in vitro* transcripts were located, only Band 4 was assigned to a particular gene. Bands 1, 2 and 3 originated from 186 regions of unknown gene content.

The transcripts are related to the DNA sequence and

FIGURE 7.6

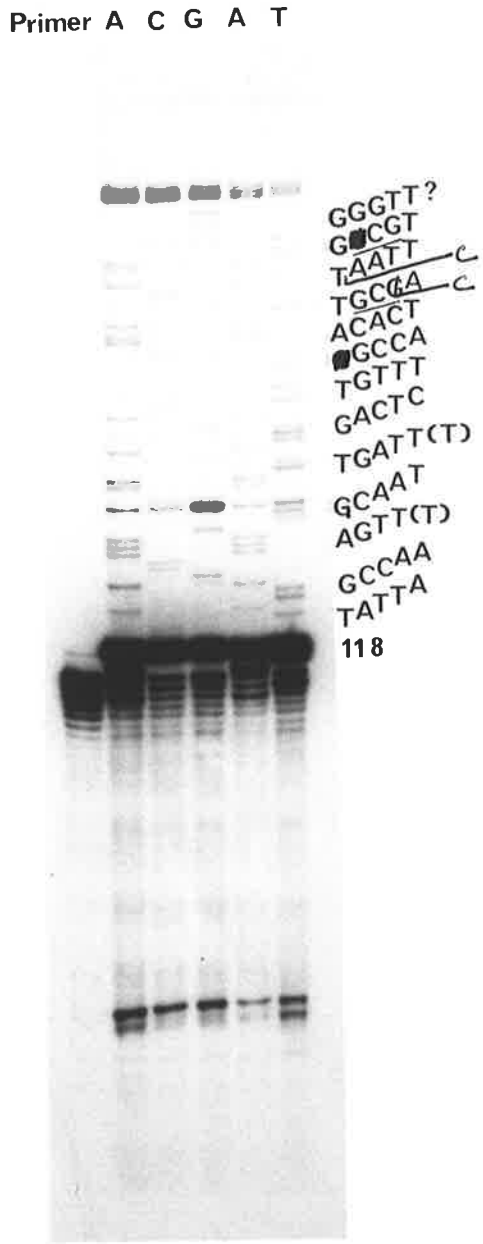
SEQUENCE BY PRIMER EXTENSION OF BAND 2

A cleavage map of the cloned *Hpa*II fragment from which the 118 base pair primer was derived is shown above the sequencing gel.

The gel was as described in Figure 7.5. Compressed regions are underlined and crossbanded nucleotides are in brackets.

The note in Figure 7.5 also applies to the interpretation of the Band 2 initiating nucleotide.


EcoRI HpaII
HinfI
HinfI HinfI
HpaII EcoRI
b.p.
198
98
33
118


Primer A C G A T
GGGTT?
GCCGT
TAATT
TGCCA
ACACT
GGCCA
TGTTT
GACTC
TGATT(T)
GCAAT
AGTT(T)
GCCAA
TATTA
118

discussed in context of the accompanying control features found in the DNA sequence in the general discussion (Chapter 9).

CHAPTER 8

COMPARISON OF 186 B AND P2 *OGR* GENES AT THE PRIMARY LEVEL

A. INTRODUCTION

The introduction to this thesis described the relationship between 186 and P2. The functional relationship of the 186 B gene and the P2 *ogr* gene was also discussed. Since ogr and B proteins are functionally similar and are interchangeable for the turn on of P2 late gene expression, it was the expectation that they would share amino acid homology.

E. Ljungquist (pers. commun.) had provided a preliminary sequence of the P2 early operon. The sequence was derived from a P2 phage carrying a deletion known as *del*6 which removes, by our estimation about 500 base pairs including the 3'-region of the P2 *int* gene. To the left of the deletionwas a rightward reading frame which stopped presumably within the deleted sequences. This open reading frame occupied the same relative map position as the 186 B gene, and although not designated as such by Ljungquist, was presumed to be the *ogr* gene. Figure 8.1 shows preliminary P2 sequence to the left of the *del*6 deletion as determined by Ljungquist. When this sequence was examined a rho-independent terminator was observed in the region 237 to 260. This terminator had a

```
G C C C C T C
| | | | | | |
C G G A G A G
```

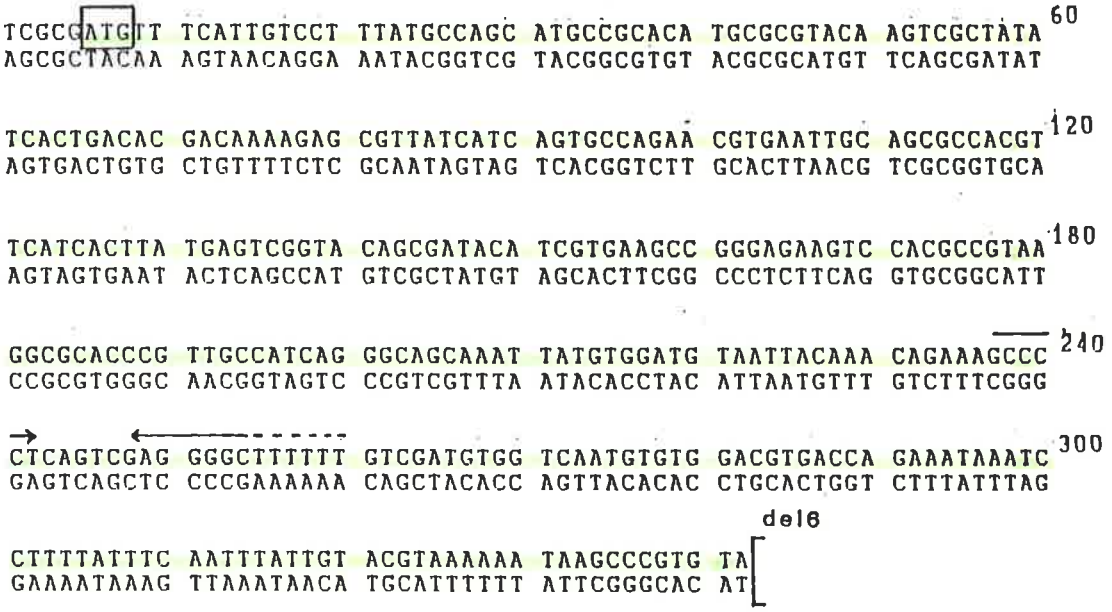
stem, and a 4 base loop followed by 6 T residues and was located in the same relative position as the 186 B gene terminator.

If the AUG at position 6 of the sequence was assigned as a protein initiation codon, the open reading frame continued beyond the rho-independent terminator. If this terminator was active, then this was somewhat of a dilemma. Since this region presumably encoded the *ogr* gene, it was reasoned that like the 186 B gene it may be present as an *in vitro* trans-

FIGURE 8.1

LJUNGQUIST'S PRELIMINARY P2 SEQUENCE

The preliminary P2 sequence to the left of the *del*6 deletion, provided by E. Ljungquist is shown. The stem of the rho-independent terminator is indicated by arrowed lines and the T residues are overlined with dashes. The ATG initiation codon assigned by B. Kalionis (this laboratory) is boxed and the open reading frame, continuing beyond the end of the sequence, is coloured.


TCGCGATGTT TCATTGTCCT TTATGCCAGC ATGCCGCACA TGCGCGTACA AGTCGCTATA 60
AGCGCTACAA AGTAACAGGA AATACGGTCG TACGGCGTGT ACGCGCATGT TCAGCGATAT
TCACTGACAC GACAAAAGAG CGTTATCATC AGTGCCAGAA CGTGAATTGC AGCGCCACGT 120
AGTGACTGTG CTGTTTTCTC GCAATAGTAG TCACGGTCTT GCACTTAACG TCGCGGTGCA
TCATCACTTA TGAGTCGGTA CAGCGATACA TCGTGAAGCC GGGAGAAGTC CACGCCGTAA 180
AGTAGTGAAT ACTCAGCCAT GTCGCTATGT AGCACTTCGG CCCTCTTCAG GTGCGGCATT
GGCGCACCCG TTGCCATCAG GGCAGCAAAT TATGTGGATG TAATTACAAA CAGAAAGCCC 240
CCGCGTGGGC AACGGTAGTC CCGTCGTTTA ATACACCTAC ATTAATGTTT GTCTTTCGGG
CTCAGTCGAG GGGCTTTTTT GTCGATGTGG TCAATGTGTG GACGTGACCA GAAATAAATC 300
GAGTCAGCTC CCCGAAAAAA CAGCTACACC AGTTACACAC CTGCACTGGT CTTTATTTAG
del6
CTTTTATTTC AATTTATTGT ACGTAAAAAA TAAGCCCGTG TA
GAAAATAAAG TTAAATAACA TGCATTTTTT ATTCGGGCAC AT

cript.

P2 DNA was transcribed *in vitro* and a transcript the same size as the 186 B gene transcript was observed. This transcript was sequenced and was found to map within Ljungquist's DNA sequence. The size of the *in vitro* transcript placed its 3'-end in the vicinity of the termination signal. Since the terminator was active, the P2 DNA sequence to the left of the *del*6 deletion was checked with the aim of discovering an error which would change the reading frame and allow the protein to terminate before the end of the transcript.

B. RESULTS

1. *In vitro* transcription of P2

P2 DNA was transcribed *in vitro* and the transcripts separated by electrophoresis. The result is shown in Figure 8.2 A P2 transcript the same size as 186 Band 4 was observed. This transcript was designated as the 290 base band. Other transcripts were evident but were not further characterised.

2. Sequence of the 290 base P2 transcript by enzymatic cleavage

In vitro transcription reactions were prepared in the absence of a radiolabelled substrate and the transcripts separated by electrophoresis. Radioactive markers were loaded in adjacent tracks to facilitate excision of the unlabelled 290 base transcript. The transcript was eluted, phosphatased and the 5'-terminii labelled using [γ-^{32}P]-ATP and T4 polynucleotide kinase. The labelled transcript was then subjected to enzymatic cleavage to generate a sequence. The sequencing gel is shown in Figure 8.3. The RNA sequence

FIGURE 8.2

IN VITRO TRANSCRIPTION OF P2

2.5 µg P2 and 186 DNA were transcribed in an *in vitro* transcription reaction. The transcripts were loaded onto a 5% (20 x 40 x 0.025 cm) polyacrylamide gel containing TBE and 8 M urea. The xylene cyanol was run to the bottom. The tracks are indicated above the gel, the 290 base P2 transcript is indicated beside the band.


P2 186
1
2
3
4
290

FIGURE 8.3

SEQUENCE BY ENZYMATIC CLEAVAGE OF THE P2 290 BASE TRANSCRIPT

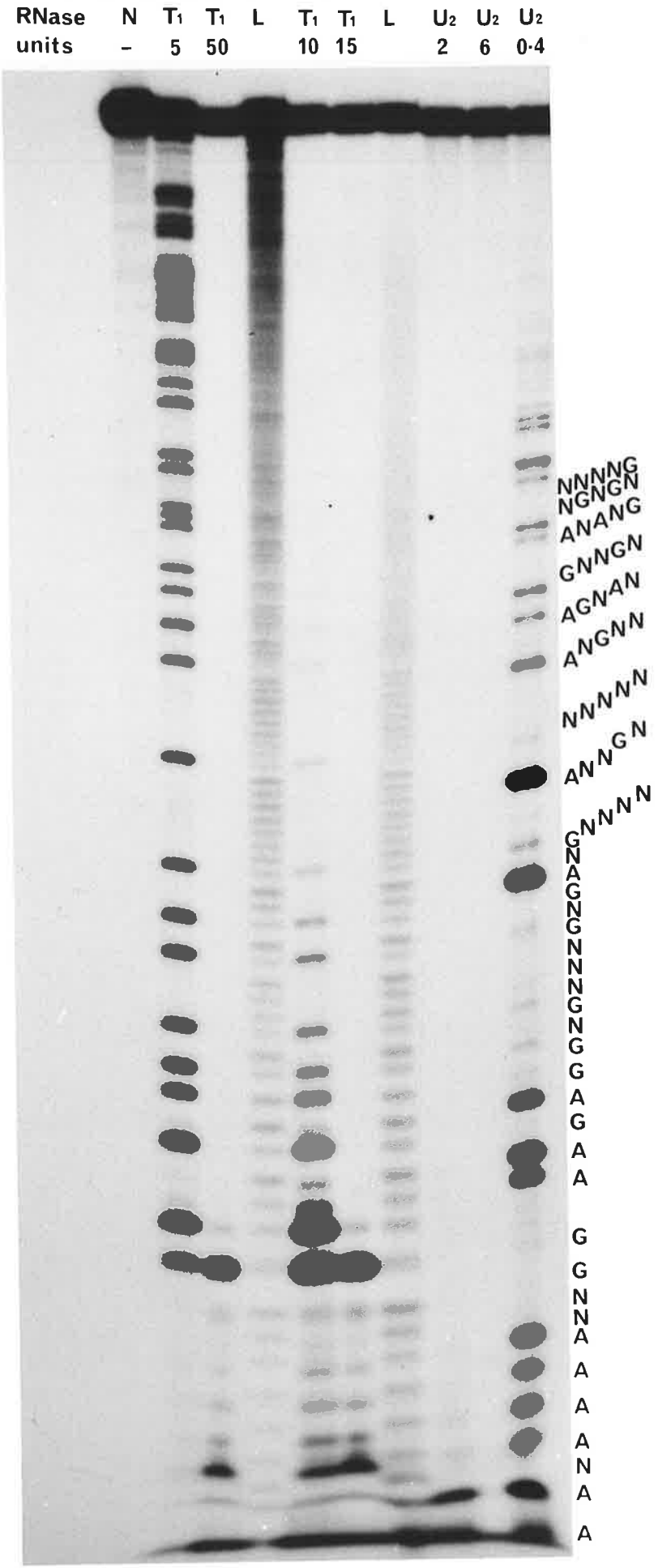
Unlabelled 290 base transcript was 5'-end labelled and subjected to enzymatic cleavage to generate a sequence. The gel is 20 x 40 x 0.05 cm 5% polyacrylamide containing TBE and 8 M urea. The tracks are indicated and the sequence written beside the ladder. RNase T1 and U2 were used to distinguish G and A residues. The positions of C or U are indicated by an N.

NOTE: Varying concentrations of U2 and T1 RNases were used as indicated.

Below is the RNA sequence obtained aligned with the sequence of Ljungquist which begins 5 residues before the AUG initiation codon.

```
                          10        20        30
                    TCGCGATGTTTCATTGTCCTTTATGCCAGCATGCCGCA
AANAAAANNGGAAGAGGNGNNNGNGANGNNNNANNGNNNNNNANGNNAGNANGNNGNA

 40        50
CATGCGCGTACAAG
NANGNGNGNNNNNG
```


RNase
units
N
–
T1
5
T1
50
L
T1
10
T1
15
L
U2
2
U2
6
U2
0·4
NNNNG
NGNGN
ANANG
GNNGN
AGNAN
ANGNN
NNNNN
ANNGN
NNNNN
GNAGNGNNNNGNGGAGAA
GGNNAAAAANAA

matches the DNA sequence of Ljungquist.

3. DNA sequence of a P2 *ogr* mutant

An *ogr* mutant was sequenced to establish the identity of the gene to the left of *del*6 and at the same time to check the sequence of Ljungquist.

The DNA to be cloned into M13 and sequenced was derived from a P2 *vir*22*ogr*52 phage. The *vir*22 phenotype has a 1.5 kb deletion, whose left boundary is believed to be the same as *del*6, and a 0.5% insertion, to the left of the *Bgl*II site at 77.2%. The P2 *vir*22*ogr*52 DNA was digested with *Pst*I and *Bgl*II and the 3.0 kb (62.6%-77.2%) fragment extracted from low gelling temperature agarose and ligated into M13mp9 digested with *Pst*I and B*am*HI. This fragment was sequenced in a leftward direction from the B*am*HI/*Bgl*II site to approximately 60 bases beyond the beginning of Ljungquists preliminary sequence. The DNA sequence to the left of Ljungquist's sequence was confirmed and extended by B. Kalionis (this laboratory) using primed synthesis on the DNA of the *Pst*I-*Bgl*II mp9 clone. Compared with Ljungquist's sequence two changes were evident, one was a $\overset{A}{|} \rightarrow \overset{G}{|}$ transition at position 130, the other was a $\overset{C}{|}$ insertion at position 182. Figure 8.4 shows an autoradigram of a sequencing gel depicting these changes.

4. 5'-sequence of P2 290 transcript by primer extension

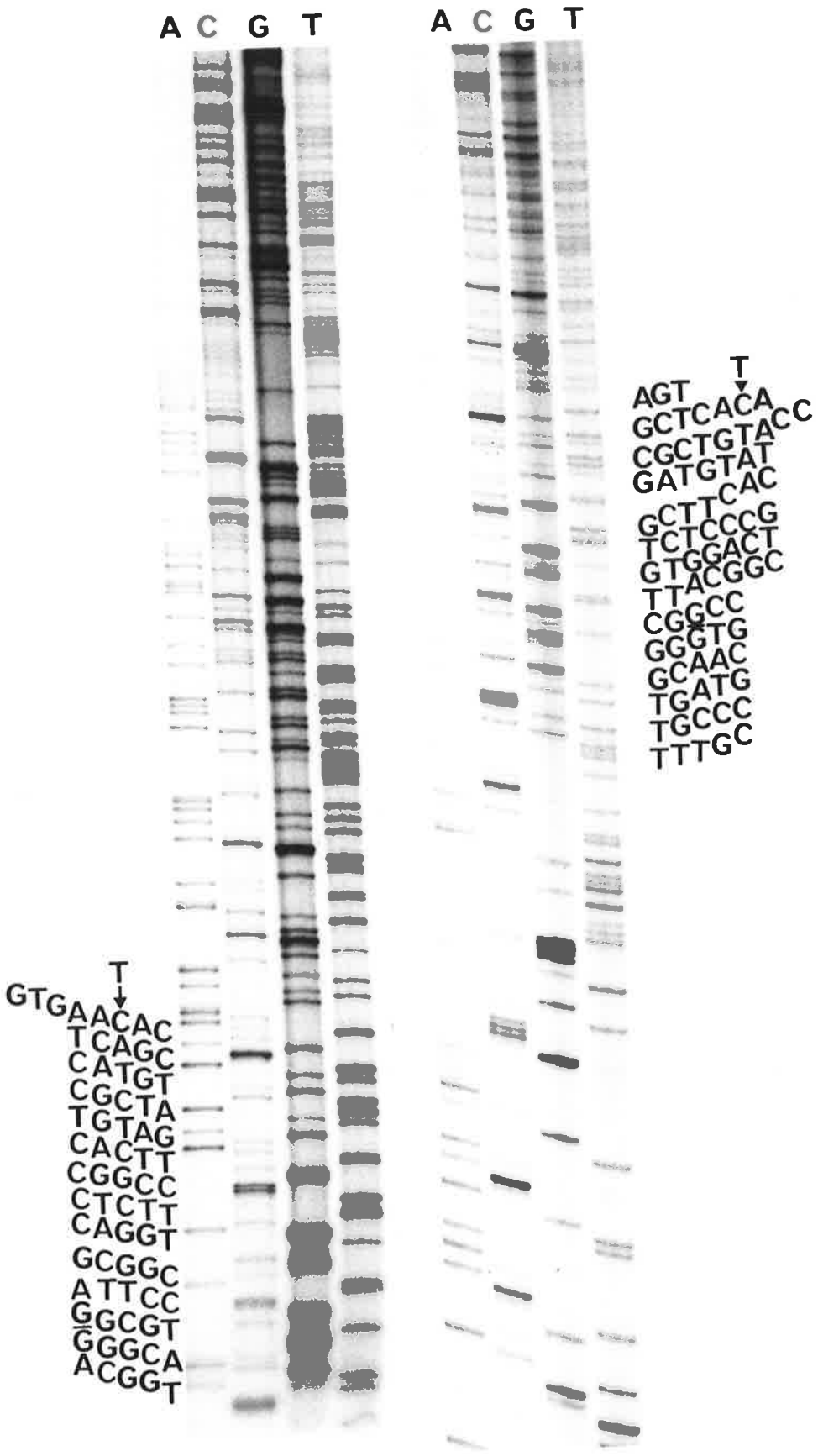
Once the 290 base transcript had been located within the region to the left of *del*6 and the DNA of this region cloned, an appropriate DNA fragment was chosen as a primer. A radioactive copy of the *Pst*I-*Bgl*II mp9 clone was synthesised and the primer, a 66 base pair *Hha*I fragment (see Figure 8.5)

FIGURE 8.4

DNA SEQUENCE OF P2*vir*22*ogr*52

Sequencing ladders of the P2 *ogr* gene are shown. The sequence is read leftward from the *Bgl*II site so is complementary to the 290 base transcript. The T → C transition is indicated by an arrow. This creates the TAT → TGT codon change in the 290 base transcript. The insertion is also indicated by underlining the inserted G.

The gel is 5% polyacrylamide (20 x 40 x 0.025 cm) containing TBE and 8 M urea.


A C G T
A C G T
T
GTGAACAC
TCAGC
CATGT
CGCTA
TGTAG
CACTT
CGGCC
CTCTT
CAGGT
GCGGC
ATTCC
GGCGT
GGGCA
ACGGT
T
AGT
GCTCACA
CGCTGTACC
GATGTAT
CAC
GCTTCCCG
TCTCCGACT
GTGGACGGC
TTACGCC
CGGGTG
GGAAC
GCATG
TGCCC
TGCCC
TTTGC

was excised and used to primer extend the unlabelled P2 290 base transcript. The sequence is presented in Figure 8.5. As is evident from the autoradiogram, the 5'-start site is not definitive, there are two T residues followed by a cross-band then an A and finally the characteristic intense cross-banding which signifies the end of the template. When the RNA sequence was matched to the DNA sequence, the nucleotide at which the intense cross-banding occurs is an adenine residue. This A has been designated as the initiating nucleotide for the 290 base transcript.

C. DISCUSSION

A 290 base *in vitro* P2 transcript was observed which when sequenced was found to originate from the P2 region presumed, from genetic studies to encode the *ogr* gene.

An apparent dilemma existed in that the transcript from this region terminated some 82 bases before the left boundary of the *del*6 deletion, whereas the translation of this region continued past the *del*6 boundary and presumably terminated within the region deleted by *del*6. A classical prokaryotic terminator with a stem-loop followed by a stretch of T residues, at position 237 to 260 of the P2 preliminary DNA sequence was found. The position of this terminator was compatible with the size of the 290 base transcript whose 5'-end had been located by primer extension (Figure 8.6).

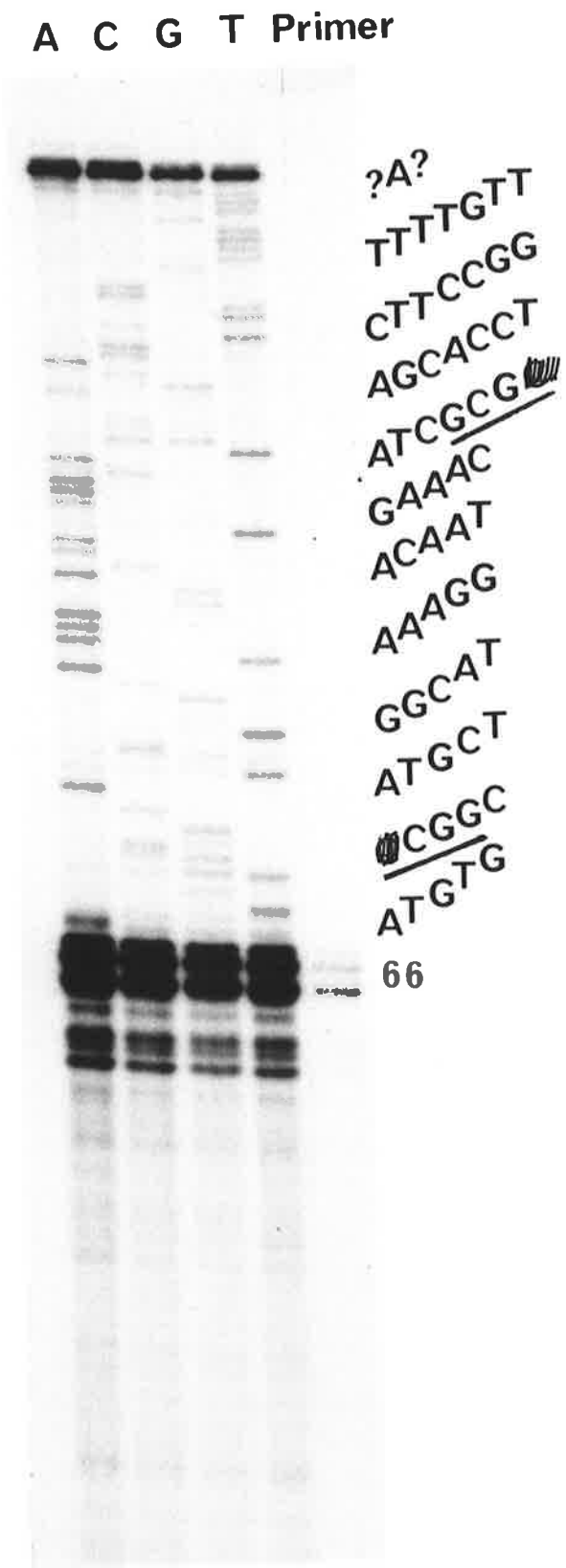
The reasons for sequencing the DNA of an *ogr* mutant were two-fold. Firstly, to confirm the identity of the *ogr* gene and hence the identity of the 290 base transcript and secondly, to look for a sequencing error which would enable

FIGURE 8.5

RNA SEQUENCE BY PRIMER EXTENSION OF THE 290 BASE TRANSCRIPT

A radiolabelled 66 base pair *Hha*I fragment was annealed to the 290 base transcript and primer extended. The sequencing reactions were loaded onto a 5% (20 x 40 x 0.025 cm) polyacrylamide gel containing TBE and 8 M urea.

The sequence is indicated alongside the ladder. Areas difficult to interprete due to compressions are underlined.


A C G T Primer
?A?
TTTTGTT
CTTCCGG
AGCACCT
ATCGCG
GAAAC
ACAAT
AAAGG
GGCAT
ATGCT
CGGC
ATGTG
66

the ogr protein to encounter a stop codon before the end of the transcript.

The revised sequence is presented in Figure 8.6. Two changes to Ljungquist's sequence were found. One was a transition $\begin{matrix}A\\|\\T\end{matrix} \rightarrow \begin{matrix}G\\|\\C\end{matrix}$ which upon translation of the RNA would substitute a cysteine for a tyrosine. The other was a $\begin{matrix}C\\|\\G\end{matrix}$ insertion which changed the reading frame of the protein, bringing the TAA stop codon at position 222 into phase, permitting the protein to terminate before the transcription termination signal.

Since the wild type P2 sequence has not been rechecked to support the existence of the $\begin{matrix}C\\|\\G\end{matrix}$ insertion, it is possible that this insertion is responsible for the *ogr* mutation by causing the premature termination of the *ogr* protein. However, the discovery of a second change in the sequence of the *ogr* mutant and the existence of the 290 base *in vitro* transcript along with a termination signal compatible with the size of the *in vitro* RNA, make this highly unlikely.

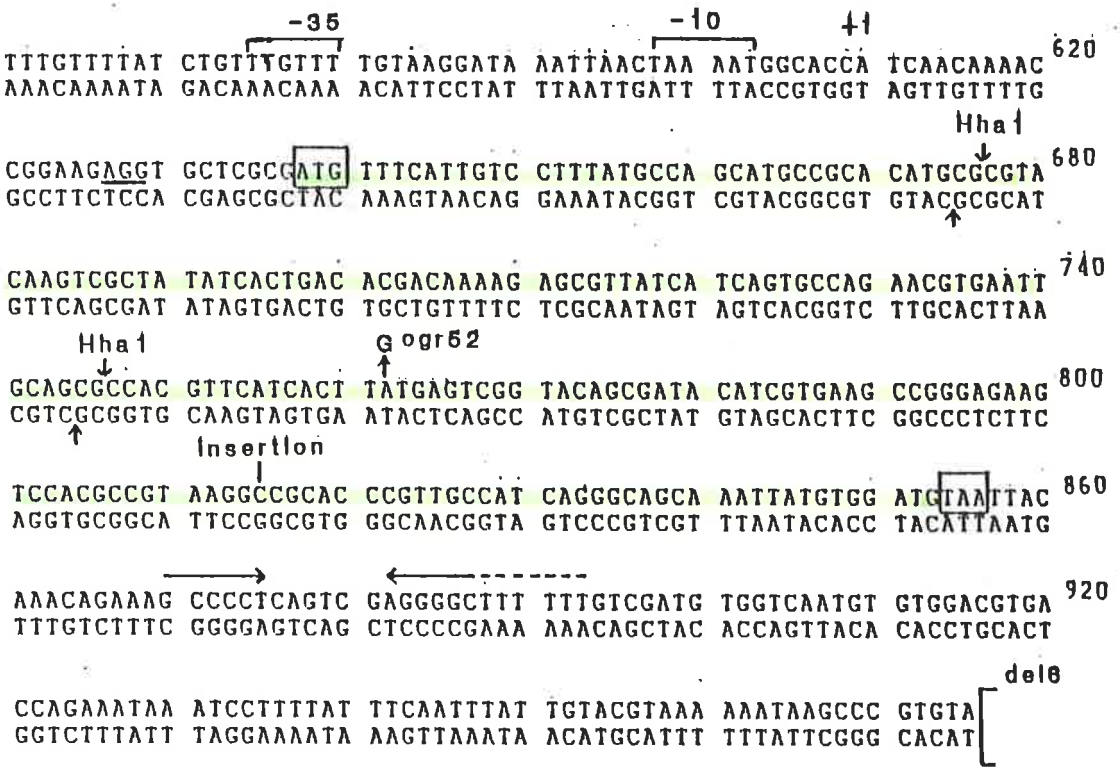
Thus the 186 B gene and the P2 *ogr* gene, whose protein products are functionally similar, encode the same size transcripts. Their translation products are also the same size, 72 amino acids.

A comparison of the translation products of the 186 B gene and the P2 *ogr* gene is presented in the Appendix.

FIGURE 8.6.

DNA SEQUENCE OF P2*ogr*52

The DNA sequence of P2*vir*22*ogr*52 is shown (this work and B. Kalionis, pers. commun.). The two changes from Ljungquist's sequence are depicted, one is the *ogr*52 mutation, the other is an insertion which changes the translation reading frame. The -10 and -35 regions of the *ogr* gene promoter are indicated. +1 represents the start of the *ogr* gene transcript. The ogr protein ribosome binding site is underlined and the initiation and termination codons are boxed. The open reading frame is coloured. The stem of the terminator structure is indicated by arrowed lines and the T residues are overlined with dashes. The *Hha*I fragment used as a primer is also indicated.


-35
-10
+1
620
Hha I
680
740
Hha I
ogr52
800
Insertion
860
920
del8

CHAPTER 9

GENERAL DISCUSSION

A. INTRODUCTION

The control of 186 gene expression is the ultimate interest of the group but before any controls can be elucidated a knowledge of the basic molecular biology of 186 is required. One facet of this study is to determine the first transcripts made by 186 upon entry of its DNA into a host cell. An *in vitro* transcription system was chosen because it contains no translation machinery and therefore, only the first transcripts made independently of phage protein synthesis would be generated. Thus an *in vitro* system should approximate the *in vivo* situation prior to phage protein synthesis.

186 DNA was transcribed *in vitro* by *E. coli* RNA polymerase and four major transcripts were observed on 5% polyacrylamide/urea gels. Preliminary mapping of the transcripts was accomplished by hybridizing them to restriction digests of 186 DNA Southern blotted onto nitrocellulose, and by analysis of the run-off transcription products generated upon *in vitro* transcription of 186 DNA digested with appropriate restriction enzymes. Sequencing the transcripts precisely located their 5'--ends.

Figure 9.1 presents a summary of the location of the four transcripts. Depicted on the figure are promoters, terminators and proteins associated with the *in vitro* transcripts.

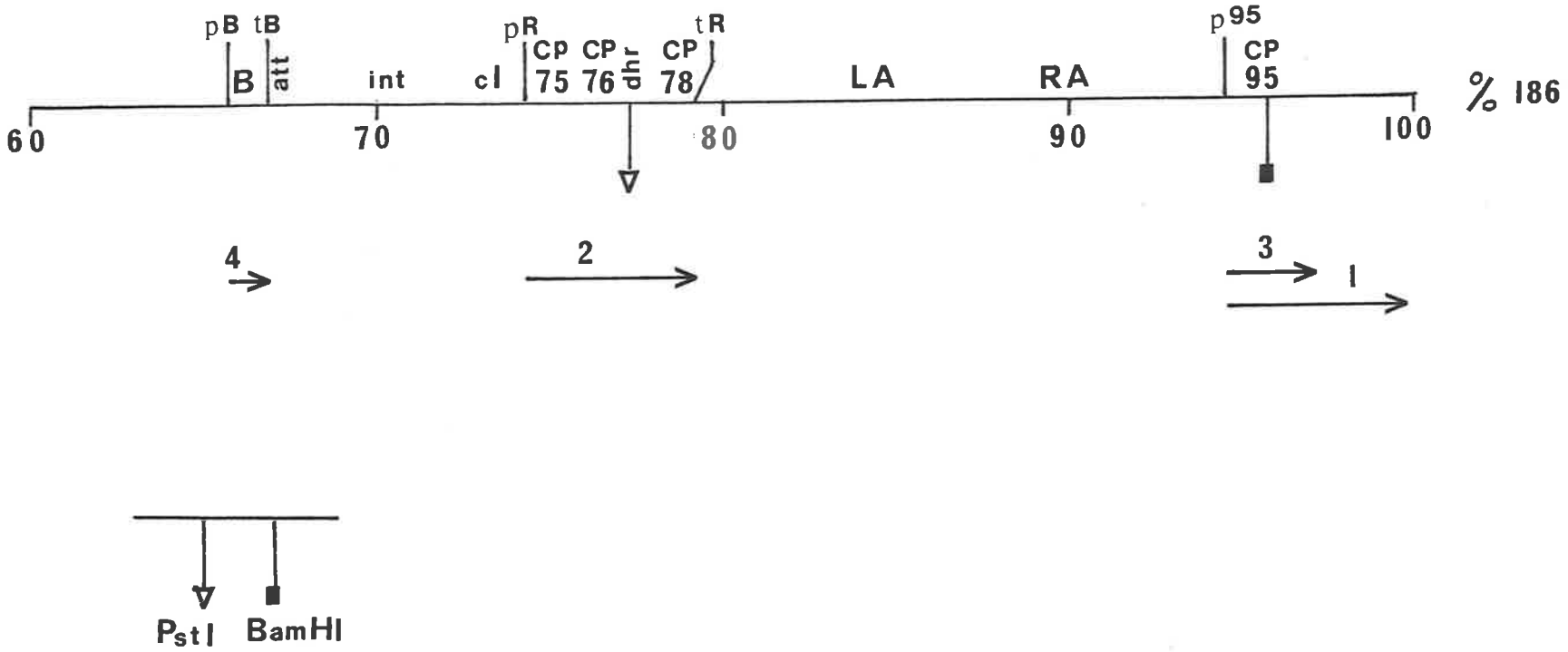
Promoters, -10 and -35 regions were predicted using a computer program designed by I. Dodd (this laboratory) similar to that of Staden (1984). Rho-independent terminators were identified by searching the DNA sequence for the char-

FIGURE 9.1

186 *IN VITRO* TRANSCRIPTION PATTERN

The map is drawn from 60% to 100%. The transcripts are drawn below the map corresponding to the positions of their promoters and terminators determined by sequencing the DNA. The terminators for Bands 1 and 3 have not been determined by sequence analysis and do not appear on the map.

The positions of the known genes and predicted open reading frames (computer proteins:CP) are shown.


pB
tB
B
att
int
cI
pR
CP 75
CP 76
dhr
CP 78
tR
LA
RA
p 95
CP 95
% 186
60
70
80
90
100
4
2
3
1
PstI
BamHI

acteristics described by Rosenberg and Court (1979), these being a G-C rich stem-loop followed by a stretch of T residues. Open reading frames were designated as encoding proteins by three criteria:

1) the presence of an initiation codon
2) the presence of a ribosome binding site using the program of Stormo *et al.* (1982)
3) codon usage, based on the program of Staden and McLachlan (1982)

The aim of this chapter is to relate the 186 *in vitro* transcripts to the DNA sequence and to other studies previously and concurrently undertaken in this laboratory. The current studies referred to are those of B. Kalionis, A.V. Sivaprasad and H. Richardson. A comparison will be made with the relevant aspects of P2 transcriptional control. In a later section, the *in vitro* transcription pattern of 186 will be compared with that of λ.

1. Band 4

Band 4 was sized on methyl mercury gels at 290 bases. The 5'-end of this transcript was sequenced and compared with the DNA of the region between 65.5% and 67.6% (B. Kalionis, pers. commun). Figure 9.2, shows the DNA sequence of the region from which Band 4 originated. The transcript initiates with an ATP and associated with the transcript is:

1) a promoter, with a -10 region, ~~TACTAT~~ TATCAT and a -35 region, ~~ACACTT~~ TTCACA.
2) a terminator with a CCCGC / GGGCG stem, a 5 base loop, followed by TTTTTTTAT.

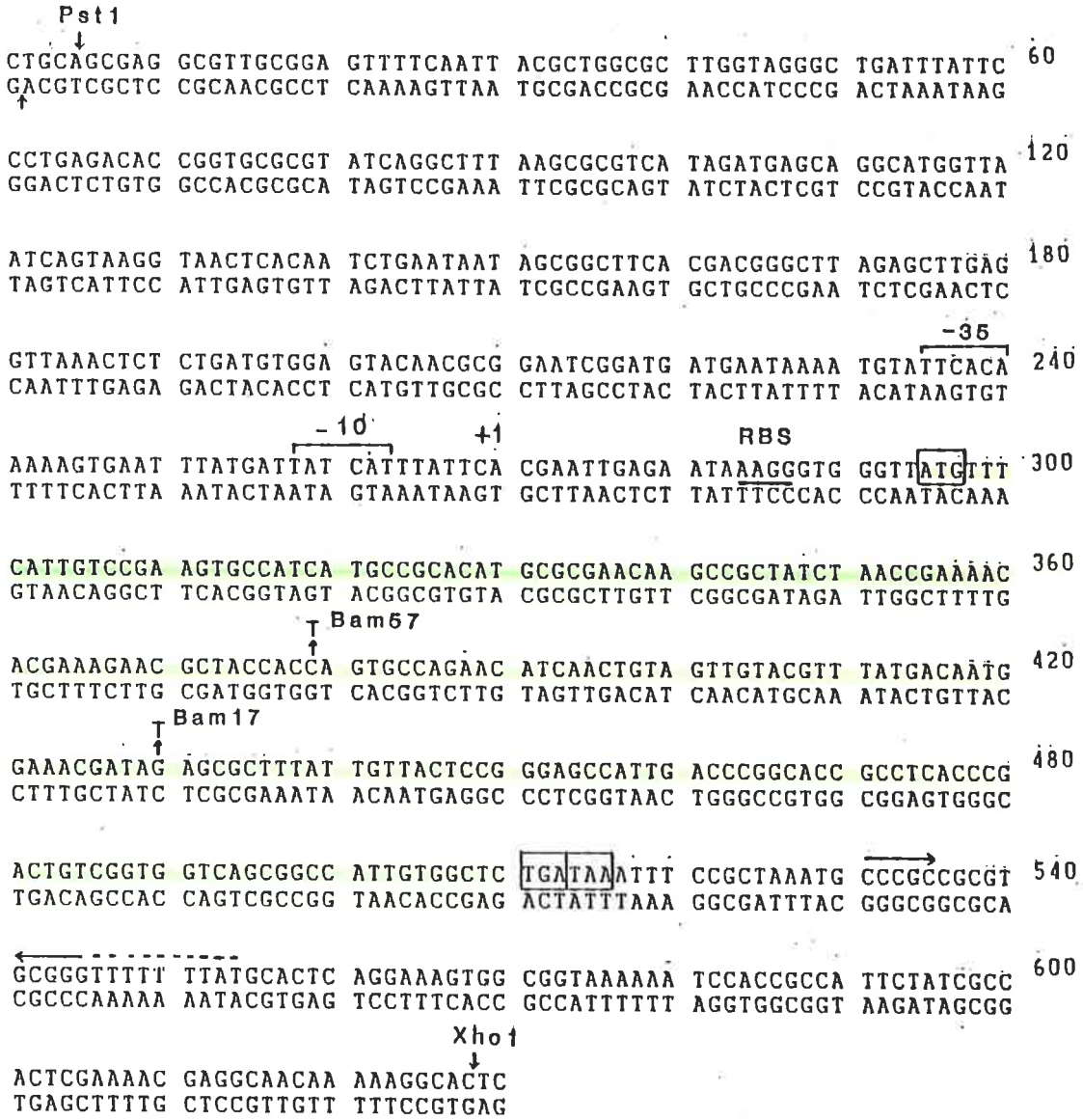
From the DNA sequence Band 4 is 284 bases in length.

FIGURE 9.2

THE DNA SEQUENCE ASSOCIATED WITH BAND 4

The DNA sequence from *Pst*I to *Xho*I (65.5% to 67.6%) is shown (B. Kälionis, pers. commun.). The -35 and -10 region of the Band 4 promoter is indicated. +1 signifies the start of Band 4. The ribosome binding site is underlined and the protein initiation and termination codons are boxed. The open reading frame is coloured. The stem of the terminator is indicated by arrowed lines and the T residues are overlined with dashes.

The nucleotide changes for B*am*17 and B*am*57 are indicated.


Pst1
-35
-10
+1
RBS
Bam57
Bam17
Xho1

Associated with the methionine initiation codon of the open reading frame is a ribosome binding site.

Genetic studies had located the B gene to this region of the 186 genome (Hocking and Egan, 1982a). Band 4 RNA obtained from transcribing two B*amber* mutants was sequenced as was the DNA from the *Pst*I-*Xho*I (65.5%-67.2%) region of the same B*amber* mutants, and Band 4 was identified as the B gene transcript. The base pair changes responsible for the B*am*17 and B*am*57 mutations are depicted on the sequence. A translation of the B gene sequence is shown in Figure 9.3. The B gene transcript encodes a protein of 72 amino acids. Amber codons substitute the GAG (glutamic acid) at position 430 in B*am*17 and the CAG (glutamine) at position 379 in B*am*57. These mutations exist in the map order predicted by marker rescue frequencies (Hocking and Egan, 1982a).

The 186 B gene is not required until late in the infection cycle of 186, when it acts to turn on expression of the late genes (Finnegan and Egan, 1981). Finnegan and Egan (1981) had also concluded that the function of the 186 A gene was required, either directly or indirectly, for B gene transcription. Therefore the ability of an unmodified *E. coli* RNA polymerase to transcribed the B gene *in vitro* without the need of any phage encoded product was an unexpected surprise. Nevertheless, that being the case, the same capacity was anticipated for the host RNA polymerase *in vivo* and, since there is no evidence of *c*I repressor control to the left of the *att* site, it was predicted that B promoter would be active in a lysogen. B. Kalionis who is studying the *in vivo* transcription pattern of 186 has indeed shown the presence of the B gene transcript

FIGURE 9.3

TRANSLATION OF THE B GENE

The figure shows the anti-sense strand sequence of the B gene, from 221 nucleotides to the right of the *Pst*I site at 65.5% to just past the B gene terminator, along with the translation. The positions of the B*am*17 and B*am*57 mutations are shown.

```
TAAAATGTATTCACAAAAAGTGAATTTATGATTATCATTTATTCACGAATTGAGAATAAA

          met phe his cys pro lys cys his his ala ala his
GGGTGGGTT ATG TTT CAT TGT CCG AAG TGC CAT CAT GCC GCA CAT

ala arg thr ser arg tyr leu thr glu asn thr lys glu arg tyr
GCG CGA ACA AGC CGC TAT CTA ACC GAA AAC ACG AAA GAA CGC TAC

his gln cys gln asn ile asn cys ser cys thr phe met thr met
CAC CAG TGC CAG AAC ATC AAC TGT AGT TGT ACG TTT ATG ACA ATG
             ↓
             T
           Bam57

glu thr ile glu arg phe ile val thr pro gly ala ile asp pro
GAA ACG ATA GAG CGC TTT ATT GTT ACT CCG GGA GCC ATT GAC CCG
             ↓
             T
             Bam17

ala pro pro his pro thr val gly gly gln arg pro leu trp leu
GCA CCG CCT CAC CCG ACT GTC GGT GGT CAG CGG CCA TTG TGG CTC

*** ***
TGA TAA  ATTTCCGCTAAATGCCCGCCGCGTGCGGGTTTTTTTAT
```

in a lysogen. Furthermore, the B protein is autoregulatory, since it has been shown *in vivo* that the level of the B gene transcript is increased when the B protein carries an amber mutation (Kalionis *et al.*, manuscript in preparation).

The B transcript is present in a lysogen and yet does not act to turn on late genes. Either the B protein is not present, or else the B protein cannot act to turn on late gene expression. The latter is most likely, since Hocking and Egan (1982a) found that even if B protein is provided by a super-infecting 186*vir* phage, no late functions can be rescued from the prophage by complementation. We predict that B protein requires a replicating template on which to act. Such a prediction, together with autogenous control, is consistent with two other observations. Firstly, Finnegan and Egan (1981) concluded that B gene transcription was dependent upon a functional 186 A gene, since hybridization of 186 messenger RNA after thermal induction of a 186*c*I*ts*A*am* prophage to pEC35, a clone presumed to be diagnostic for B gene transcription, showed very little transcriptional activity from this region. Secondly, Hocking and Egan (1982a) found that B protein activity could be rescued by complementation from a 186*c*I*ts* prophage by a super-infecting 186B*amvir* phage. The super-infecting B*amvir* phage provides a replicating template. Based on the number of polarity groups found by Hocking and Egan (1982b), we expect at least four sites on each replicon to which B protein may bind to activate late gene expression. This increase in the number of B binding sites titrates the finite pool of B protein originally present in the lysogen and thereby opens up B

gene transcription. The A gene is not required for B gene transcription *per se* as suggested by Finnegan and Egan (1981) but A gene function is required for the increase in B gene transcription required for late turn on.

2. P2 *ogr*

An *in vitro* transcript the same size as the 186 B gene transcript was observed upon transcription of P2. This transcript was sequenced and located on the DNA sequence. It was identified as the *ogr* gene transcript by sequencing the appropriate region of the DNA from a P2*ogr*52 mutant.

Sauer *et al.* (1982) concluded that *ogr* gene expression depends on a functional A gene by coinfecting a host *gro* mutant, which is unable to support a P2*ogr*$^+$ infection, with P2A*am*129*ogr*1 and P2*vir*1. Since P2A*am* mutants cannot be complemented (Lindahl, 1970), P2A*am*129 will not replicate, but if the mutant *ogr* gene can be expressed from P2A*am*129*ogr*1 to produce mutant ogr protein, it should complement P2*vir*1*ogr*$^+$ to produce a burst of progeny. No such complementation was observed. However, Sauer *et al.*, did not directly determine *ogr* gene transcription in their experiment and since, the *ogr* gene is transcribed *in vitro* by an unmodified RNA polymerase, like the 186 B gene, we predict the *ogr* gene transcript will be present in a P2 lysogen, ie., exist in the absence of gene A function. Negative complementation may explain the failure of Sauer and his colleagues to detect *ogr*1 complementation rather than the dependence of *ogr* gene transcription on a functional A gene. The P2*vir*1*ogr*$^+$ phage would replicate so the amount of ogr$^+$ protein would be much greater than the amount of mutant ogr protein provided by a non-replicating P2A*am*129*ogr*1 phage.

This excess of ogr$^+$ protein, although defective in a *gro* host could result in a situation of negative complementation. To resolve the apparent paradox, further experiments on the *in vivo* transcripts of P2 are required.

The 186 B and P2 ogr proteins are functionally interchangeable for late gene turn on but a difference does exist between the two proteins. The turn on of P2 late genes by ogr is mediated through the α-subunit of the host RNA polymerase (Fujiki *et al.*, 1976). 186 is able to grow on host *gro* mutants which carry a mutation in the α-subunit of the host RNA polymerase (Fukiki *et al.*, 1976; B. Egan, pers. commun.), a mutation which prevents the growth of P2*ogr*$^+$.

3. Band 2

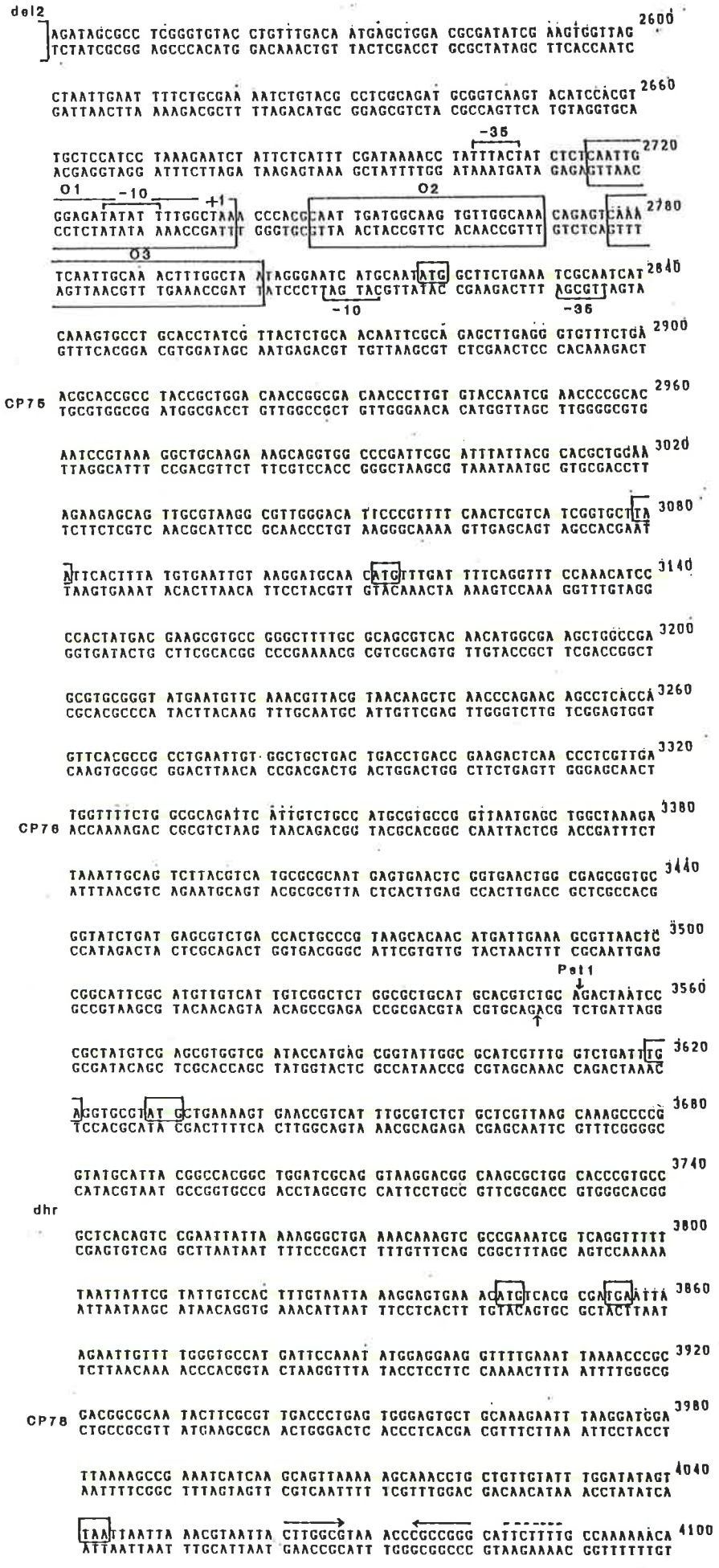
The 5'-end of Band 2 was sequenced and was found to initiate with an ATP 809 bases to the left of the *Pst*I site at 77.3%. The DNA sequence from which Band 2 originates is shown in Figure 9.4. The sequence to the left of 77.3% is from B. Kalionis (pers. commun.) and to the right of this site is from H. Richardson (pers. commun.).

Scanning the DNA sequence in this region located three related sequences, the putative operators which presumably bind repressor. The promoter associated with Band 2 has a -10 region, TATATT and a -35 region, TTTACT. The -10 region is located within O1, which places this promoter under the direct control of the *c*I repressor, which when bound to O1 would block the initiation of Band 2 RNA. Band 2 has been designated as the *p*R transcript, analogous to the λ rightward transcript.

FIGURE 9.4

THE DNA SEQUENCE ASSOCIATED WITH BAND 2

The DNA sequence between the right boundary of the Δ2 deletion and just past the presumptive Band 2 terminator is shown (B. Kalionis, H. Richardson, pers. commun.). The *Pst*I site at 77.3% is indicated. The -35 and -10 regions of the Band 2 promoter and the presumptive leftward promoter are indicated. +1 represents the start of Band 2. The putative operators are boxed. The protein initiation and termination codons for CP75, CP76, *dhr* and CP78 are boxed and the open reading frames coloured. The stem of the presumptive terminator is indicated by arrowed lines and the T residues are overlined with dashes.


del2
2600
2660
-35
2720
O1
-10
+1
O2
2780
O3
2840
-10
-35
2900
CP75
2960
3020
3080
3140
3200
3260
3320
3380
CP76
3440
3500
Pst1
3560
3620
3680
3740
dhr
3800
3860
3920
CP78
3980
4040
4100

The other interesting feature near the 5'-end of *p*R is a computer predicted -10 and -35 region of a leftward promoter. This leftward transcript is most likley that of the *c*I gene located to this region by heteroduplex studies (Younghusband *et al.*, 1975). Studies are currently underway to confirm the *in vivo* activity of this leftward promoter. The location of the leftward promoter is to the right of 03 so transcription from *p*R and *p*C may be overlapping divergent. Attempts to identify the *c*I transcript *in vitro* were negative. This is not unexpected, since if the promoter for *p*R is strong compared with the leftward promoter, *p*R may preferentially initiate transcription *in vitro*, whereas environmental factors may contribute to the rate of *c*I transcription *in vivo*. Thus the lysis/lysogeny decision of 186 may be controlled by the modulation of overlapping divergent transcripts.

The *c*I/operator region has a similar arrangement in P2 (Ljungquist *et al.*, 1984, in press). The operator region located by sequencing the DNA of virulent mutants, indicates two possible repressor binding sites. The repressor gene C, identified by sequencing the DNA of C*amber* mutants is transcribed leftward from a putative promoter located to the right of 02. The first gene in the early operon, presumably regulated by repressor, has its putative promoter overlapped by 01. No transcription studies are available to confirm the activity of these putative promoters but if they do promote C and *cox* transcription then, as postulated for 186, P2 transcription is overlapping divergent.

The size estimate for Band 2 from methyl mercury gels

was 1450 bases, but from the DNA sequence there is no obvious terminator in this region. There is a rho-independent terminator located 1323 base pairs downstream of the Band 2 initiation site (see Figure 9.4). This terminator has a stem C T T G G C G / G G G C C G C with a 6 base loop, followed by CATTCTTTT. Confirmation of the activity of this terminator will come from the precise location of the 3'-end of of Band 2.

Band 2, or the *p*R transcript, spans four open reading frames. The first two, designated CP (computer protein) 75 and CP76 are as yet unidentified genes. The third open reading frame has been identified as that of the *dhr* gene (formerly *dho*) postulated by Finnegan and Egan (1981), (H. Richardson, pers. commun.). The fourth open reading frame, designated CP78, is also unidentified.

Finnegan and Egan (1981) also postulated the existence of a gene X, whose transcription was a prerequisite for any further phage transcription (see Fig. 1.5). Gene X was postulated to explain the result obtained by hybridization of 186 *in vivo* transcripts after heat induction of a prophage, derived from chloramphenicol treated cells, to 186 restriction fragments cloned into pBR322. pEC35 exhibited a high transcriptional activity in the presence of chloramphenicol. This activity was presumed to originate from the 0.4 kb fragment (now known to be 1.0 kb) to the right of the Δ2 deletion, since it was thought that in the absence of the A gene product, the B gene was not transcribed. Given now that the B gene is transcribed in a lysogen, some of this transcriptional activity would have been contributed by the B gene transcript.

Although Finnegan and Egan did not consider this contribution, their prediction that the high transcriptional activity originated from the right of the Δ2 deletion remains valid for the following reason. B gene transcription is under autogenous control and if the B protein is stable under chloramphenicol treatment, then B transcription should be repressed and so not contribute appreciably to the transcriptional activity of pEC35. This is supported by comparing hybridization of RNA after thermal induction of a 186*cIts* prophage from chloramphenicol treated cells, with the RNA after heat induction of an A*amber* mutant. The hybridization to pEC35 in both cases was the same (Finnegan and Egan, 1981) consistent with the following explanation. In an A*amber* mutant, B transcription would be negatively regulated by the B protein. If the B protein was unstable, then upon chloramphenicol treatment of a 186*cIts* lysogen, an increase in the transcriptional activity of pEC35 should occur compared with the 186A*am* prophage due to the removal of B gene repression and hence increased transcription of the B gene. No such increase was evident, therefore B transcription is repressed upon thermal induction of a 186*cIts* chloramphenicol treated lysogen.

Given that Band 2 extends some 543 bases into the adjacent *Pst*I fragment (77.3%-84.6%), the same fragment studied by Finnegan and Egan (1981) as a pBR322 clone called pEC17.2, the hybridization results rationalised by Finnegan and Egan as ratios of c.p.m./kb can be refined as follows. Finnegan and Egan found 292 c.p.m. associated with the 0.4 kb fragment in pEC35 and 136 c.p.m. associated with the 2.3 kb

fragment from pECl7.2 which generated a ratio of 730:59 c.p.m./kb for the two intervals. Now, the 0.4 kb interval has been found to be 1.0 kb with 809 base pairs contributing to the Band 2 transcript, and instead of considering the entire 2.3 kb interval of pECl7.2, only 543 base pairs need be considered. The ratio of c.p.m./kb now becomes 365.6:250.5. Assuming Band 2 only is contributing to the 250.5 c.p.m./kb found associated with pECl7.2, then the transcription contributed by pEC35 to the right of Δ2 is 250.5 c.p.m./kb. Therefore, the B gene is contributing 115.1 c.p.m./kb to the transcriptional activity of pEC35, less than half the amount contributed by Band 2. This result is consistent with B gene repression being maintained in the presence of chloramphenicol, as *in vitro* the B gene transcript is 3.4 times more abundant than Band 2.

The postulate of a gene X, as a prerequisite for further phage transcription remains valid. 186 must somehow transcribe its replication genes, so a mechanism must exist for extending rightward transcription. There are two possible mechanisms. Firstly, antitermination of Band 2 to facilitate further rightward transcription, or alternatively, reinitiation at downstream promoters to extend transcription across the replication genes. 186 may encode a product, analogous to the λ N gene product, which would act as an antiterminator. Or a product may be encoded which would activate transcription from downstream promoters. Whatever, the mechanism, it is probable that this antiterminator or activator protein will be encoded by Band 2. CP76 and CP78 are candidates for gene X. CP75 is most likely the counterpart of the P2 *cox*

gene product, as when translated these proteins share some amino acid homology (B. Kalionis, pers. commun.). There is preliminary evidence to suggest that extension of rightward transcription is via reinitiation rather than antitermination (B. Kalionis, pers. commun.).

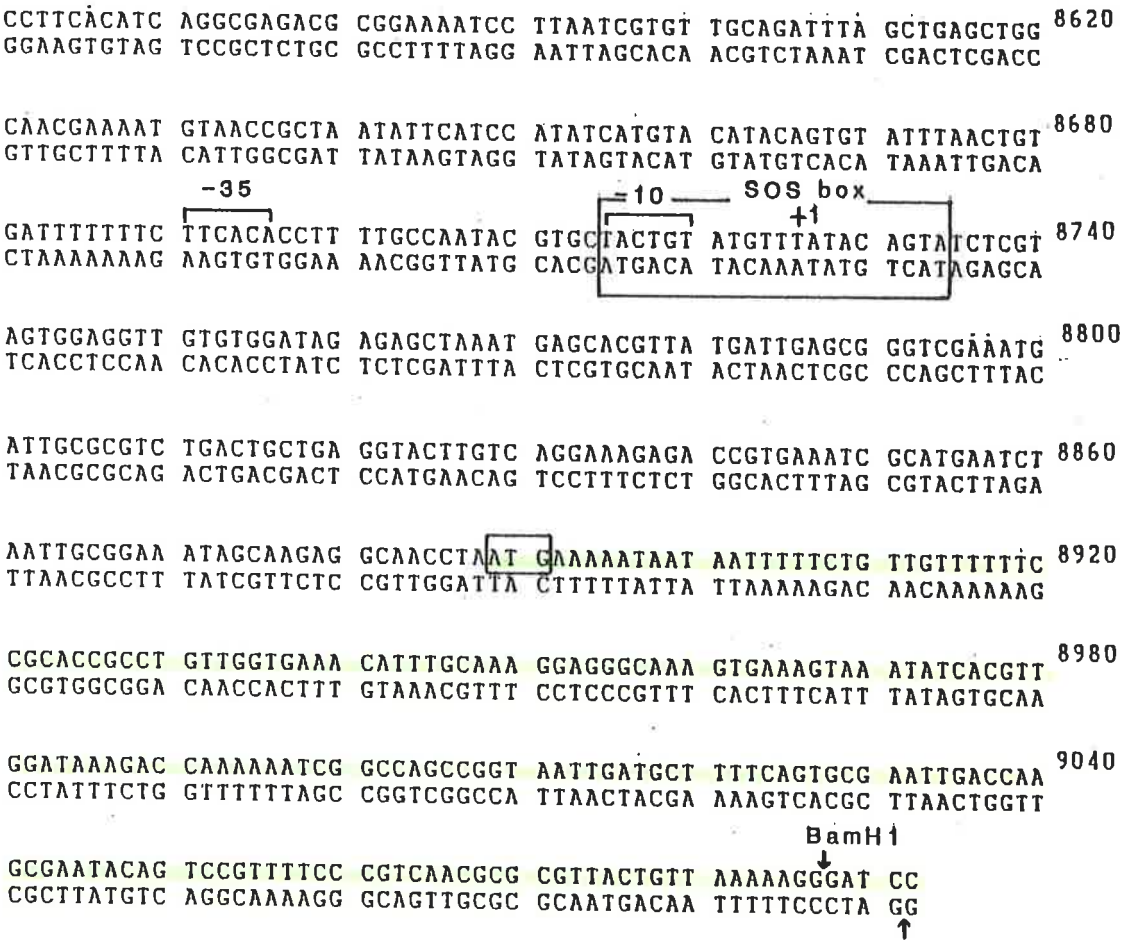
4. Bands 1 and 3

Bands 1 and 3 were positioned by run-off transcription to initiate approximately 390 bases to the left of the B*am*HI site at 95.7%. The 5'-ends of Bands 1 and 3 were sequenced and found to be identical, confirming that these transcripts shared an initiation site. The sequence of Bands 1 and 3 were compared with the DNA sequence of the region to the left of 95.7% provided by A.V. Sivaprasad (pers. commun.). This DNA sequence is presented in Figure 9.5. Bands 1 and 3 initiate with an ATP and are associated with a promoter with a -10 region, TACTGT and a -35 region, TTCACA. There is an open reading frame with a ribosome binding site and a methionine initiation codon associated with these transcripts but these features are located 161 bases from the 5'-end of the transcripts, which is rather further downstream than the ribosome binding sites and protein initiation codons associated with Bands 2 and 4 (78 bases and 24 bases respectively). This open reading frame extends across the B*am*HI site at 95.7%. Sequence information beyond 95.7% is unavailable so nothing is known about the 3'-ends of Bands 1 and 3 except that their size precludes them from traversing the cohesive ends, and Band 1 is a readthrough of Band 3. The termination efficiency of Band 3 is 79%. There are no known genes in the region of Bands 1 and 3 so

FIGURE 9.5

THE DNA SEQUENCE ASSOCIATED WITH THE 5'-ENDS OF BANDS 1 AND 3

The DNA sequence from *Pst*I to *Bam*HI (94.0% to 95.7%) is shown (A.V. Sivaprasad). The -35 and -10 regions of the promoter for Bands 1 and 3 are indicated. +1 signifies the common start of the Band 1 and Band 3 transcripts. The SOS box is boxed. The protein initiation codon for CP95 is boxed and the open reading frame which continues beyond the *Bam*HI site is coloured.


-35
-10
SOS box
+1
BamH1

their function is speculative, based upon the features in the DNA sequence.

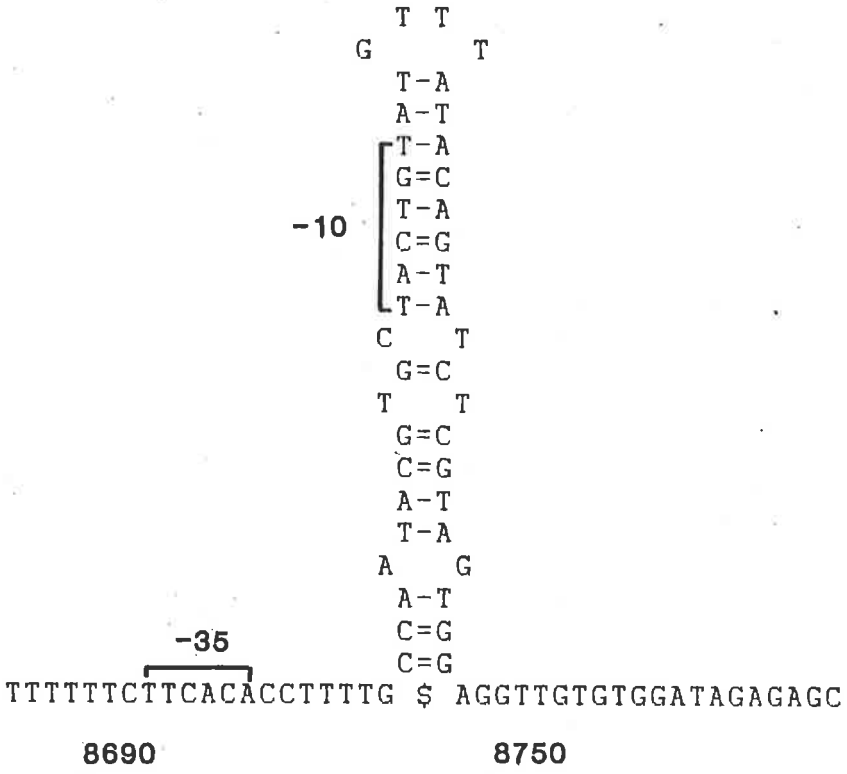
A pallindromic sequence exists around the promoter for Bands 1 and 3. If the DNA near this region is negatively supercoiled, then cruciform structures are possible (Hsich and Wang, 1975; Benham, 1982; Gellert *et al.*, 1983) and the -10 region of the promoter may be sequestered into the stem of a DNA stem-loop of theoretical ΔG=-13.75 kcal/mole, as shown in Figure 9.6. The -10 region would be rendered unavailable for binding RNA polymerase, until a nick occurred in the DNA, say, at the onset of replication. Hence the control of expression of Bands 1 and 3 may be at the level of DNA structure. Linearised plasmid or 186 phage DNA which would not exist *in vitro* in a supercoiled state, were used as templates for *in vitro* transcription experiments, so Bands 1 and 3 would always be present.

More recently, an interesting feature was discovered in the DNA sequence in the region of the promoter for Bands 1 and 3. By scanning 186 sequences, B. Kalionis (pers. commun.) located an SOS box spanning the -10 region of the promoter (see Figure 9.5). LexA is a repressor for many *E. coli* SOS functions. To effect repression, the LexA protein binds to a DNA sequence around the promoter regions of the genes which it controls. This DNA sequence is known as the SOS box and has been found in the promoter-operator regions of *lex*A and *rec*A (Little *et al.*, 1982) and *uvr*B (G.B. Sancar *et al.*, 1982). The SOS box also exists in λ (Sprichitsky and Kopylov, 1983). It has been found to span the promoter of the *oop* transcript. These workers proposed that LexA would bind to

FIGURE 9.6

CRUCIFORM STRUCTURE NEAR *p*95

The possible stem-loop structure near *p*95, the promoter for Bands 1 and 3, is shown. The -10 region is sequestered into the stem of the cruciform. The ΔG = -13.75 Kcal/mole.


T T
G T
T-A
A-T
T-A
G=C
T-A
C=G
A-T
T-A
C T
G=C
T T
G=C
C=G
A-T
T-A
A G
A-T
C=G
C=G
-10
-35
TTTTTTCTTCACACCTTTTG $ AGGTTGTGTGGATAGAGAGC
8690
8750

the SOS box and repress *oop* transcription. Upon irradiation of a λ prophage by ultra voilet light, the LexA protein would be cleaved and the repression of *oop* transcription lifted. The leftward *oop* transcript would then compete with the *p*R transcript preventing the synthesis of *c*II protein and forcing λ to undergo lytic development rather than remain in a potentially damaged cell.

The role of the SOS box in 186 has yet to be elucidated but it is possible that it may have a role in controlling the expression of some phage function required late in infection, whose expression is opened up after replication by titration of the LexA protein.

Located to the right of 92% on the P2 chromosome is a constitutive function defined by a gene *old* (Bertani, 1975). The *old* gene is active in a P2 lysogen as evidenced by the interference of λ by a P2 prophage. When the prophage has a mutation in the *old* gene this interference of λ is abolished (Lindahl *et al.*, 1970). Since the *old* gene is expressed in a P2 lysogen, it is possible that the *old* transcript may be generated by an unmodified RNA polymerase in an *in vitro* transcription system, and although this expected *old* transcript originates from the region analogous to Bands 1 and 3 of 186, their functions are not analogous as λ forms plaques on a 186 lysogen.

5. Comparison of 186 and λ

The genes controlling the sequential appearance of λ proteins and the transcripts made, are described in the general introduction (Chapter 1) along with a diagramatic representation (Figure 1.3).

Superficially λ and 186 transcription follows a similar pattern of sequential gene expression, with early genes activating delayed middle transcription which in turn controls DNA replication and late gene expression. Within this framework the details of control mechanisms differ. 186 has one repressor controlled promoter for essential genes, with the operator mapping to the right of *c*I; in comparison λ has two essential promoters with the associated operators on either side of the *c*I gene.

λ *o*R binds repressor to modulate the transcription from two promoters, *p*R and *prm*, whose sequences overlap the operators (see Figure 1.4) and the transcription is divergent. The 186 operator may also modulate transcription from two promoters, *p*C (presumptive *c*I transcript) and *p*R, but the transcription is overlapping divergent. In λ the repressor binding site negatively controlling transcription from *p*R is closest to the gene it regulates (*cro*) whilst in 186, the repressor binding site controlling *p*R is furtherest from the gene it controls.

The major deviation in 186 from the sequentialia exhibited by λ lies in the control of the late turn on gene, B. In λ, transcription of the late turn on gene, Q, depends upon antitermination of *p*R at *t*R2 by the λ N protein. The late turn on gene of 186, the B gene, is transcribed independently of phage protein synthesis. λQ acts as an antiterminator to extend transcription into the head and tail genes and its activity is retained *in vitro* (Grayhack and Roberts, 1982) indicating that its action is not dependent upon a replicating template. All evidence suggests that the B protein acts at the late promoters as an activator of transcription and is

only active on a replicating template.

6. The future

In vitro transcription studies have laid the groundwork for numerous other pursuits.

Band 4, the B gene transcript, is being studied *in vivo* to determine its control. It has been postulated that its control is autoregulatory, but this needs to be tested directly. Isolated B protein should inhibit transcription of Band 4 when added to an *in vitro* transcription reaction.

Once the 5'-end of the *c*I transcript has been located, expression of the overlapping transcripts can be studied with regard to their influence on the lysis/lysogeny decision of 186. These studies have already been initiated. Isolated repressor protein can be used *in vitro* to directly block *p*R transcription to confirm that its promoter is negatively controlled by the *c*I gene product. The mechanism whereby 186 extends transcription to the right of Band 2 must be determined.

Some light may be shed upon the role of Bands 1 and 3 by firstly testing the theories suggested in Section 9.4. Negative supercoils may be introduced into a plasmid, for example, pEC200, carrying the appropriate region of the 186 genome, by DNA gyrase. If cruciform structures are formed rendering the promoter for Bands 1 and 3 unavailable, then Bands 1 and 3 should not be generated upon *in vitro* transcription of such a plasmid. Upon introducing a single-strand nick into the DNA to release the supercoiling, Bands 1 and 3 should reappear.

The presence of the SOS box may be confirmed by transcribing 186 DNA *in vitro* in the presence of LexA protein.

If LexA acts to repress initiation of Bands 1 and 3, the transcripts should not be generated.

The information gained from the study of 186 *in vitro* transcripts has increased the knowledge of 186 molecular biology and although much remains unknown, the realm of the control of 186 gene expression has been opened up, thereby justifying the study of the initial transcripts of coliphage 186.

APPENDIX

COMPARISON OF THE LATE CONTROL PROTEINS OF COLIPHAGES P2, P4 AND 186

The product of the P2 gene *ogr* controls late gene transcription. The 186 B gene and P2 *ogr* gene products are functionally interchangeable as evidenced by the existence of viable P2/186 hybrids that possess P2 late genes and 186 gene B for their activation. The δ gene of the satellite phage P4 can substitute for 186 B and P2 *ogr* in late gene turn on. It was therefore expected that the products of the P2 *ogr*, 186 B and P4 δ genes would exhibit amino acid homology.

Figure A.1 shows a translation of the 186 B gene DNA sequence compared to translations of the P2 *ogr* gene sequence gained from the preliminary sequence data of Ljungquist and the P4 δ gene sequence gained from the preliminary sequence of Lin. The δ protein is larger than its 186 or P2 counterparts, containing 154 amino acids compared to 72 amino acids for B and ogr. There is a striking homology between B and ogr, with 32 of 39 amino terminal residues identical. The homology is less with ogr and δ, with 19 identical amino acids among the 39 amino terminal residues.

FIGURE A.1

TRANSLATION OF 186 B, P2 *ogr* AND P4 δ

The figure shows the sequences and the translation products of the 186 B gene, the P2 *ogr* gene and the P4 δ gene. They are compared with P2 *ogr*.

```
        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
                                                                                    60

        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      C.T GGC ..C ..G ATA ATG AT. GCA ... ... .G. ..T A.T T.. ... TCT ..G ..A ..T GAA
                                                                                   120

        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 ile his phe arg gln thr glu
P4      G.G .G. ..A AGC T.T TTC TCT G.. .G. AAA G.. T.A ... AT. CAT TT. AGG CA. .CA GA.
                                                                                   180

        his pro thr val gly gly gln arg pro leu trp leu *** ***
186     ... ... ACT GTC GGT ..T ... .GG CCA T.. ... C.G .G. TAA

        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

        ile pro arg phe thr asp val ser leu leu arg tyr glu thr pro glu lys asn aln cys
P4      ATA ... CGA TTC A.T .AC GTC AGC C.C C.. C.. TAT G.. ACG CCC GAA AAG AAT GCT TGT
                                                                                   240

        tyr arg arg tyr cys cys arg arg cys lys gly leu ala leu ser arg thr ser arg arg
P4      TAC CGG CGG TAT TGT TGT CGG AGA TGT AAA GGC CTT GCA CTG TCA AGA ACA TCG CGG CGT
                                                                                   300

        leu ser gln glu val thr glu arg phe tyr val cys thr asp pro gly/cys gly leu val
P4      CTG TCT CAG GAA GTC ACC GAG CGT TTT TAT GTG TGC ACG GAT CCG GGC TGT GGT CTG GTG
                                                                                   360

        phe lys thr leu gln thr ile asn arg phe ile val arg pro val thr pro ala asn trp
P4      TTT AAA ACG CTT CAG ACC ATC AAC CGC TTC ATT.GTC CGC CCG GTC ACG CCG GCG AAC TGG
                                                                                   420

        gln asn ala cys met lys asn arg asn cys arg gln tyr gly ***
P4      CAG AAC GCC TGC ATG AAA AAC AGG AAC TGC CGC CAG TAC GGT TAA
                                                                  465
```

REFERENCES

ADHYA, S., GOTTESMAN, M., de CROMBRUGGHE, B. (1974) *Proc. Natl. Acad. Sci. U.S.A.* 71: 2534-2538

APPLEYARD, R.K. (1954) *Genetics* 39: 440-452

BAIL, J.M., DAVIDSON, N. (1976) *Anal. Biochem.* 70: 75-85

BALDWIN, R.L., BARRAND, P., FRITSH, A., GOLDWAIT, D.A., JACOB, F. (1966) *J. Molec. Biol.* 17: 343-357

BENHAM, C.J. (1982) *Biopolmers* 21: 679-696

BERTANI, G. (1968) In: *Molecular Genetics* (Eds. H. Wittman & N. Schauster) pp. 180-186

BERTANI, G. (1975) *Mol. Gen. Genetics* 136: 107-137

BERTANI, L.E. (1968) *Virology* 36: 87-103

BERTANI, L.E. (1970) *Proc. Natl. Acad. Sci. (Wash).* 65: 331-336

BIRMBOIN, H.C., DOLY, J. (1979) *Nucleic Acids Res.* 7: 1513-1523

BRENT, R., PTASHNE, M. (1981) *Proc. Natl. Acad. Sci. U.S.A.* 78: 4204-4208

BOLIVAR, F., RODRIQUEZ, R.L., GREENE, P.J., BETLACH, M.C., HEYNEKER, H.L., BOYER, H.L. (1977) *Gene* 2: 95-113

BRADLEY, C., ONG, P.L., EGAN, J.B. (1975) *Mol. Gen. Genet.* 140: 123-135

BRUENING, G., GOULD, A.R., MURPHY, P.J., SYMONS R.H. (1982) *F.E.B.S. Letters* 148: 71-78

CAMPBELL, A. (1965) *Virology* 27: 329-339

CHACONAS, G., VAN DE SAUDE, J.H. (1980) *Methods Enzymol.* 65: 75-85

CHATTORAJ, D.K., INMAN, R.B. (1973) *Proc. Natl. Acad. Sci. U.S.A.* 70: 1768-1771

DHARMARAJAH, V.K. (1975) Honours Thesis, University of

Adelaide

DUGAICZYK, A., BOYER, H.W., GOODMAN, H.M. (1975) *J. Molec. Biol.* 96: 174-184

DIANA, C., DEHO, G., GEISSELSODER, J., TINELLI, L., GOLDSTEIN, R. (1978) *J. Molec. Biol.* 126: 433-445

DONIS-KELLER, H., MAXAM, A.M., GILBERT, W. (1977) *Nucleic Acids Res.* 4: 2527-2539

DUNN, J.J., STUDIER, F.W. (1973) *Proc. Natl. Acad. Sci. U.S.A.* 70: 1599-1563

ECHOLS, H. (1971) In: *The Bacteriophage Lambda* (Ed. A.D. Hershey) pp. 247-270

ECHOLS, H., GREEN, L., OPPENHEIM, A.B., OPPENHEIM, A., HONIGMAN, A. (1973) *J.Molec.Biol.* 80: 203-216

FINNEGAN, E.J., EGAN, J.B. (1979) *Mol. Gen. Genet.* 172: 287-293

FINNEGAN, E.J., EGAN, J.B. (1981) *J. Virol.* 38: 987-995

FOLKMANIS, A., MALTZMAN, W., MELLAN, P., SKALKA, A., ECHOLS, H. (1977) *Virology* 81: 352-362

FRANKLIN, N.C. (1974) *J.Molec.Biol.* 89: 33-48

FRIEDMAN, D.I., BARON, L.S. (1974) *Virology* 58: 141-148

FRIEDMAN, D.I., BAUMANN, M., BARON, L.S. (1976) *Virology* 73: 119-127

FRIEDMAN, D.I., JOLLY, C.A. MURAL, R.J., PONCE-CAMPOS, R., BAUMANN, M.F. (1976) *Virology* 71: 61-73

FRIEDMAN, D.I., SCHAUER, A.T., BAUMANN, M.F., BARON, L.S., ADHYA, S.L. (1981) *Proc. Natl. Acad. Sci. U.S.A.* 78: 1115-1118

FUJIKI, H., PALM, P., ZILLIG, W., CALENDAR, R., SUNSHINE, M. (1976) *Mol. Gen. Genet.* 149: 19-22

FUNNELL, B.E., INMAN, R.B. (1983) *J. Molec. Biol.* 167: 311-334

GEISSELSODER, J. (1976) *J. Molec. Biol.* 100: 13-22

GEISSELSODER, J., CHIDAMBARAM, M., GOLDSTEIN, R. (1978) *J. Molec. Biol.* 126: 447-456

GEISSELSODER, J., MANDEL, M., CALENDER, R., CHATTORAJ, D.K. (1973) *J. Molec. Biol.* 77: 405-415

GEISSELSODER, J., YOUDARIAN, P., DEKO, G., CHIDAMBARAM, M., GOLDSTEIN, R., LJUNGQUIST, E. (1981) *J. Molec. Biol.* 148: 1-19

GELLERT, M., O'DEA, M.H., MIZUUCHI, K. (1983) *Proc. Natl. Acad. Sci. U.S.A.* 80: 5545-5549

GOLDSTEIN, R., SEDIVY, J., LJUNGQUIST, E. (1982) *Proc. Natl. Acad. Sci. U.S.A.* 79: 515-519

GOLOMB, M., CHAMBERLIN, M. (1974) *J. Biol. Chem.* 249: 2858-2863

GONZALEZ, N., WIGGS, J., CHAMBERLIN, M. (1977) *Arch. Biochem. Biophys.* 182: 404-408

GRAYHACK, E.J., ROBERTS, J.W. (1982) *Cell* 30: 637-648

GREENBLATT, J., LI, J., ADHYA, S., FRIEDMAN, D., BARON, L.S., REDFIELD, B., KING, H.F., WEISBACH, H. (1980) *Proc. Natl. Acad. Sci. U.S.A.* 77: 1991-1994

GUSSIN, G., JOHNSON, A., PABO, C., SAUER, R. (1983) In: *Lambda II* (Ed. R.W. Hendrix, J.W. Roberts, F.W. Stahl, and R.A. Weisberg) pp. 93-123

HARRIS, J.D., CALENDAR, R. (1978) *Virology* 85: 343-358

HERSKOWITZ, I., HAGEN, D. (1980) *Ann. Rev. Genet.* 14: 399-445

HOCKING, S.M., EGAN, J.B. (1982a) *Mol. Gen. Genet.* 187: 87-95

HOCKING, S.M., EGAN, J.B. (1982b) *J. Virol.* 44: 1056-1067

HOCKING, S.M., EGAN, J.B. (1982c) *J. Virol.* 44: 1068-1071

HOCKING, S.M., EGAN, J.B. (1982d) *Mol. Gen. Genet.* 187: 174-176

HOOPER, I., WOODS, W.H., EGAN, J.B. (1981) *J. Virol.* 40: 341-349

HSICH, T.S., WANG, J.C. (1975) *Biochem.* 14: 527-535

JOHNSON, A., MEYER, B.J., PTASHNE, M. (1978) *Proc. Natl. Acad. Sci. U.S.A.* 75: 1783-1787

JOHNSON, A.D., MEYER, B.J., PTASHNE, M. (1979) *Proc. Natl. Acad. Sci. U.S.A.* 76: 5061-5065

KAFATOS, F.C., WELDON JONES, C., EFSTRATIADIS, A. (1979) *Nucleic Acids Res.* 7: 1541-1552

KALIONIS, B., EGAN, J.B. (1981) *Gene* 15: 95-98

KOURILSKY, P., BOURGUIGNON, M.F, GROS, F. (1971) IN: *The Bacteriophage Lambda* (Ed. A.D. Hershey) pp. 647-666

LEHRACH, H., DIAMOND, D., WOZNEY, J.M., BOEDTKER, H. (1977) *Biochemistry* 16: 4743-4751

LINDAHL, G. (1970) *Virology* 42: 1522-1533

LINDAHL, G. (1971) *Virology* 46: 520-633

LINDAHL, G., SIRONI, G., BIALY, A., CALENDAR, R. (1970) *Proc. Natl. Acad. Sci. U.S.A.* 66: 587-594

LINDGVIST, B.H., SIX, E.W. (1971) *Virology* 43: 1-7

LITTLE, J.W., MOUNT, D.W., YANISCH-PERRON, C.R. (1981) *Proc. Natl. Acad. Sci. U.S.A.* 78: 4199-4203

LJUNGQUIST, E., KOCKUM, K., BERTANI, L.E. (1984) *Proc. Natl. Acad. Sci. U.S.A.* in press

MANIATIS, T., PTASHNE, M., BACHMAN, K., KLEID, D., FLASHMAN, S., JEFFREY, A., MAURER, R. (1975) *Cell* 5: 109-113

MAURER, R., MEYER, B.J., PTASHNE, M. (1980) *J. Molec. Biol.* 139: 147-161

MESSING, J. (1981b) The Cleveland Symposium on Macromolecules; Recombinant DNA (Ed. A. Walton, Elsevier) Amsterdam pp. 143-153

MESSING, J., CREA, R., SEEBURG, P.H. (1981a) *Nucleic Acids Res.* 9: 309-323

MEYER, B.J., KLEID, D.G., PTASHNE, M. (1975) *Proc. Natl. Acad. Sci. U.S.A.* 72: 4785-4789

MURRAY, K., MURRAY, N.E. (1973) *Nature* 243: 134-139

PADMANABHAN, R., WU, R. (1972) *J. Molec. Biol.* 65: 447-467

PORTIER, M.M., MARCAUD, L., COHEN, A., GROS, G. (1972) *Mol. Gen. Genet.* 117: 72-81

PTASHNE, M., BACHMAN, K., HUMAYUN, M.Z., JEFFREY, A., MAURER, R., MEYER, B., SAUER, R.T. (1976) *Science* 194: 156-161

RAO, R.N., ROGERS, S.G. (1979) *Gene* 7: 79-82

REICHARDT, L.F. (1975) *J.Molec.Biol.* 93: 267-288

ROBERTS, J.W. (1970) *Cold Spring Harbour Symp. Quant. Biol.* 35: 121-126

ROSENBERG, M., COURT, D. (1979) *Ann. Rev. Genetics* 13: 319-355

RUPP, W.D., WILDE, C.E., RENO, D.L., HOWARD-FLANDERS, P. (1971) *J. Molec. Biol.* 61: 25-44

SALSTROM, J.S., SZYBALSKI, W. (1978) *J. Molec. Biol.* 124: 195-221

SANCAR, A., SANCAR, G.B., RUPP, W.D., LITTLE, J.W., MOUNT, D.W. (1982) *Nature* 298: 96-98

SANCAR, G.B., SANCAR, A., LITTLE, J.W., RUPP, W.D. (1982) *Cell* 28: 523-530

SANGER, F., NICKLEN, S., COULSON, A.R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74: 5463-5467

SAUER, B., CALENDAR, R., LJUNGQUIST, E., SIX, E.W., SUNSHINE, M. (1982) *Virology* 116: 523-534

SHORE, D., DEHO, G., TSIPIS, J., GOLDSTEIN, R. (1978) *Proc. Natl. Acad. Sci. U.S.A.* 75: 400-404

SILBERKLANG, M., GILLUM, A., RAJBHANDARY, U.L. (1979) *Methods Enzymol.* 59: 58-109

SIMON, L.D., GOTTESMAN, M., TOMEZAK, K., GOTTESMAN, S. (1979) *Proc. Natl. Acad. Sci. U.S.A.* 76: 1623-1627

SIX, E.W. (1975) *Virology* 67: 249-263

SIX, E.W, LINDQVIST, B.H. (1978) *Virology* 87: 217-230

SIX, E.W., KLUG, C. (1973) *Virology* 51: 327-344

SKALKA, A., HANSON, P. (1972) *J. Virol.* 9: 583-593

SOUTHERN, E.M. (1975) *J. Molec. Biol.* 98: 503-517

SOUZA, L., CALENDAR, R., SIX, E.W., LINDQVIST, B.H. (1977) *Virology* 81: 81-90

STADEN, R. (1984) *Nucleic Acids Res.* 12: 505-519

STADEN, R., McLACHLAN, A.D. (1982) *Nucleic Acids Res.* 10: 141-156

STAHL, S.J., CHAMBERLIN, M.J. (1977) *J.Molec. Biol.* 112: 577-601

STORMO, G.D., SCHNEIDER, T.D., GOLD, L.M. (1982) *Nucleic Acids Res.* 10: 2971-2995

SUNSHINE, M.G., SAUER, B. (1975) *Proc. Natl. Acad. Sci. U.S.A.* 72: 2770-2774

SUNSHINE, M.G., THORN, M., GIBBS, W., CALENDAR, R., KELLY, G. (1971) *Virology* 46: 691-702

SZYBALSKI, W., KUBINSKI, H., HRADECNA, Z., SUMMERS, W.C. (1971) *Methods in Enzymol.* 21: 383-413

TAKADA, Y. (1979) *J. Molec. Biol.* 127: 177-189

WAHL, G.M., STERN, M., STARK, G.R. (1979) *Proc. Natl. Acad. Sci.* 76: 3683-3687

WALZ, A., PIRROTTA, V., INEICHEN, K. (1976) *Nature* 262: 665-669

WANG, J.C. (1967) *J. Molec. Biol.* 28: 403-411

WOODS, W.H., EGAN, J.B. (1974) *J. Virol.* 14: 1349-1356

YOUNGHUSBAND, H.B., EGAN, J.B., INMAN, R.B. (1975) *Mol. Gen. Genetics* 140:101-110

YOUNGHUSBAND: H.B., INMAN, R.B. (1974) *Virology* 62: 530-538