



**CLONING AND EXPRESSION OF THE GENES ENCODING  
BACTERIOPHAGE T7 & SP6 RNA POLYMERASE**

A thesis submitted to the University of Adelaide  
for the degree of Doctor of Philosophy

by

Rhett Swanson B.Sc. (Hons)

*Accepted for publication 10/1/91*

Department of Biochemistry  
The University of Adelaide  
Adelaide, South Australia

July, 1990

To Janette

## STATEMENT

This thesis contains no material which has been previously submitted for an academic record at this or any other University and is the original work of the author, except where due reference is made in the text. I consent to this thesis being made available for photocopying and loan.

Rhett Swanson

## ACKNOWLEDGEMENTS

I wish to thank Professors W.H. Elliott and G.E. Rogers for permission to undertake these studies in the Department of Biochemistry and my supervisor Professor Bob Symons for his help, advice and encouragement.

Life in Lab 104 has been enhanced by the very good company of many people and in particular: Dr. Chris Davies for his supply of beer and lies in the face of experimental disasters; Drs. Iain Borthwick, Jason Loveridge and Steve Wilton for showing me how to *really* eat pizza; Kym Bennett and Syu for their wonderful humour, diversions and transport; Dr. Joel Bornstein for academic advice and submarine adventures; Alx for pretending to be a space monster; David and Henk for company in the 'Hills'; Colin for stories while fencing; Amal and Chris for extraordinary food and friendship; Sharon Harrison for smiling during enzyme assays; Gem for guidance on dharma and Dr. Derek Skingle for friendship and showing me Nepal.

There is also the ever flowing stream of humanity which constitute the Symons group at any given time. In particular, Anke Warner, Alex Jeffries, Candice Sheldon and Andrew Rakowski have been ideal comrades.

I wish to thank Jenny Cassady, Tammy Greatrex, Jan Gunter and Barbara Engel for technical assistance, Lesley Crocker for advice on photography, Brian Denton for his engineering skills and forgetting the job-order and to Jane Sommerville for easing a few sticky situations with the front office.

I especially wish to thank my friends and Jack, Hope, Janet and Janette for their love.

## DEFINITIONS

The bacteriophage T7 genome is 39,936 bp of linear DNA. The various diagrams in this work show numbers above restriction sites etc. The numbers refer to the number of base pairs from the left end of the genome as defined by Dunn and Studier (1982), in which early genes are on the left and late genes are on the right. T7 DNA in plasmids show their T7 coordinates, not the local plasmid coordinates.

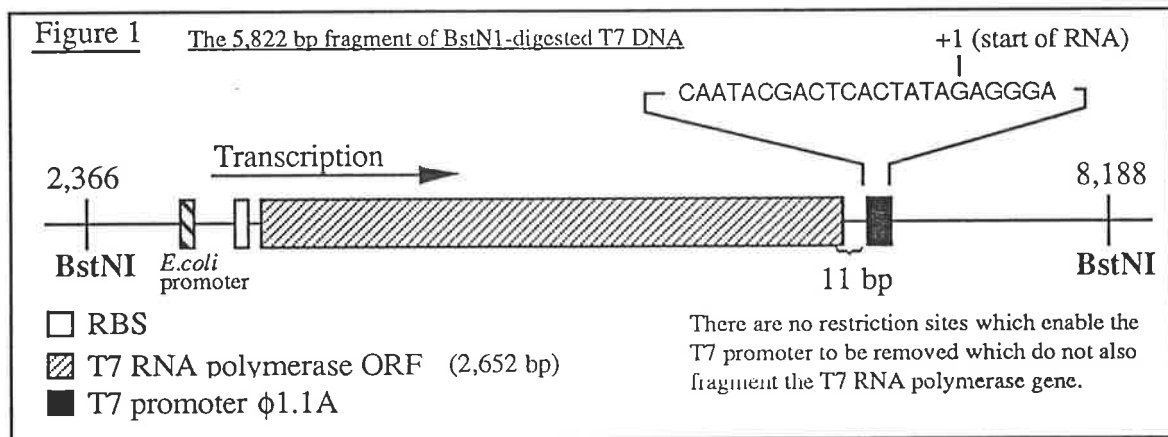
## ABBREVIATIONS

A <sub>600</sub>	absorbance at 600 nm
BSA	bovine serum albumin
CIP	calf intestinal alkaline phosphatase
dNTP	deoxynucleoside 5'-triphosphate
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
GF/A	glass fibre disk
HPLC	high performance liquid chromatography
IPTG	isopropyl- $\beta$ -D-thiogalactopyranoside
kbp	kilobase pair
NP40	Nonidet P40 detergent
NTP	nucleoside 5'-triphosphate
PCR	polymerase chain reaction
PEG	polyethylene glycol
PNK	T4 polynucleotide kinase
RBS	ribosome binding site
RF	replicative form
SDS	sodium dodecyl sulphate
SP6	bacteriophage SP6
T7	bacteriophage T7
TAE	Tris acetate EDTA buffer
TBE	Tris borate EDTA buffer
TEMED	N,N,N',N',-tetramethylethylenediamine
TMAC	tetramethylammonium chloride
Tris	tris(hydroxymethyl)aminomethane

## SUMMARY

The general aim of this project was to develop approaches for rapidly cloning and expressing genes which may be useful to Molecular Biology and which may have commercial value.

The initial focus of this project was to clone the gene for bacteriophage SP6 RNA polymerase, which at that time had not been cloned and for which the gene locus was unknown. Prior to the cloning of this gene by Kotani et al. (1987), the genetic organisation around the SP6 RNA polymerase gene was thought to be analogous to that of the T7 RNA polymerase gene. The latter gene lies upstream to a promoter element,  $\phi 1.1A$  (see figure 1).



A DNA fragment encoding both the T7 RNA polymerase gene and promoter  $\phi 1.1A$  was known to be lethal to the host bacterium. Cloning this gene required the isolation of a DNA fragment encoding the gene but lacking T7 promoter  $\phi 1.1A$  (Davanloo et al. 1984).

The T7 RNA polymerase gene thus presents a well characterised target on which to develop strategies for cloning genes without relying on restriction sites. Toward this end, two experimental strategies were tested using the cloning of the T7 RNA polymerase gene as a model system:

### 1. Bacteriophage M13 oligonucleotide mutagenesis.

Site-directed oligonucleotide mutagenesis using bacteriophage M13mp18 (Zoller and Smith 1983) was used in conjunction with Bal 31 nuclease to isolate the T7 RNA polymerase gene. The strategy was to intentionally over-digest, with Bal 31 nuclease, a restriction fragment of T7 genomic DNA that contained the T7 RNA polymerase gene and the potentially lethal promoter  $\phi 1.1A$ , size the fragments and then clone appropriately sized fragments into *E.coli*. The resulting bacterial

colonies will contain a recombinant vector in which there is a deletion of the T7 promoter  $\phi 1.1A$ , the 5' coding region or deletions to both termini of the gene. A Bal 31 nuclease digested T7 DNA fragment bearing a partial length T7 RNA polymerase gene was cloned into bacteriophage M13mp18 and site-directed oligonucleotide mutagenesis was then used to repair the DNA fragment back to a full length T7 RNA polymerase gene.

The gene was expressed under control of the leftward promoter ( $P_L$ ) from bacteriophage lambda using the expression vector pPLC236 (Remaut et al. 1981). The *E.coli* host for this vector produces a temperature sensitive *ci857* repressor allowing induction of the T7 RNA polymerase gene by raising the temperature of the bacterial culture from 30°C to 42°C. The RNA polymerase was expressed to a very high level but was enzymically inactive.

To ensure the enzymic inactivity of the recombinant RNA polymerase was not due to mutations in the gene, the nucleotide sequence of the gene was determined by dideoxysequencing and found to match the sequence published by Moffatt et al. (1984). In addition, the recombinant T7 RNA polymerase was purified and subjected to N-terminal amino acid sequencing, revealing that the first 12 amino acids of the protein matched the first 12 amino acids of T7 RNA polymerase.

The enzymically inactive recombinant T7 RNA polymerase was thus initiated from the first AUG codon 3' to the ribosome binding site and is the product of a gene lacking mutations. The lack of demonstrable enzyme activity therefore appears to be due to some factor other than the primary structure of the protein.

It has been reported that phage T7 gene 19, when cloned downstream of the lambda  $P_L$  promoter and expressed in *E.coli* at 42°C, yielded protein which was enzymically inactive. However, attenuating the expression level of T7 gene 19 from the  $P_L$  promoter resulted in the recovery of T7 gene 19 protein which showed enzymic activity (White and Richardson 1988).

Whether attenuating the expression level of the T7 RNA polymerase gene would also facilitate the recovery of enzymically active T7 RNA polymerase was then determined.

## **2. Attenuated expression of T7 RNA polymerase in *E.coli***

The approach taken to reduce the expression level of T7 RNA polymerase in *E.coli* was to re-clone



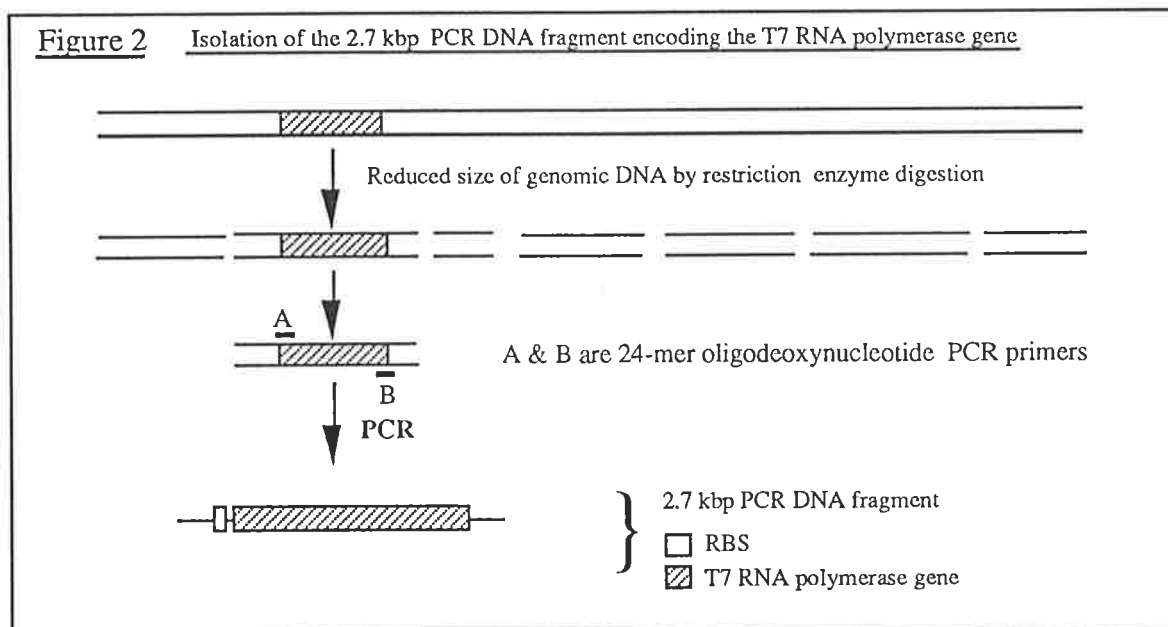
the gene into another expression vector (pEDV-1) downstream of a controllable promoter ( $P_{lac}$ ), which exhibits weaker promoter strength *in vivo* than the  $P_L$  promoter (Deuschle et al. 1986). Expression of the T7 RNA polymerase gene in *E.coli* was then induced at 37°C by adding IPTG to the bacterial culture.

Prior to cloning the T7 RNA polymerase gene into pEDV-1, the T7 RNA polymerase gene was isolated from T7 genomic DNA by exploiting the polymerase chain reaction. This allowed a direct comparison to the previous approach of repairing Bal 31 nuclease-digested DNA with site-directed mutagenesis, as a means of isolating a specific DNA fragment from genomic DNA, independently of restriction sites.

### 3. Isolating the T7 RNA polymerase gene by the Polymerase Chain

#### Reaction (PCR) and its expression in *E.coli* from promoter $P_{lac}$

The polymerase chain reaction (PCR) was tested to see if a large DNA fragment (the T7 RNA polymerase gene is 2,652 base pairs) could be isolated without PCR-induced mutations. The T7 RNA polymerase gene isolated by this approach was extracted from T7 genomic DNA with two 24-mer PCR primers (see figure 2).



The 2.7 kbp PCR DNA fragment encodes the T7 RNA polymerase gene and the ribosome binding site 5' to the gene but does not include the promoter  $\phi 1.1A$ . This

DNA fragment was ligated into an expression vector (pEDV-1) that allows the gene to be expressed under the control of the lac operon promoter from *E. coli* ( $P_{lac}$ ) and cloned into *E. coli* JM101.

Induction of the gene at 37°C with IPTG yielded enzymically active T7 RNA polymerase, demonstrating that the expression of this protein in a form that shows enzyme activity depends on the controllable promoter in the expression vector and, possibly, on the temperature during gene induction and the genotype of the host bacterium.

#### **4. Isolating and cloning the RNA polymerase gene from bacteriophage**

##### **SP6 and its expression in *E. coli*.**

The technique of extracting a large DNA fragment from genomic DNA, by PCR, prior to cloning and expressing the gene in *E. coli*, was also tested by isolating the gene encoding the RNA polymerase from bacteriophage SP6. The PCR used two 24-mer oligonucleotide primers designed to extract the SP6 RNA polymerase gene and its ribosome binding site, and were designed using the SP6 DNA sequence data of Kotani et al. (1987). The same protocol used in extracting the T7 gene was used for extracting the RNA polymerase gene from phage SP6 genomic DNA. A PCR-derived DNA fragment bearing the SP6 RNA polymerase gene was ligated into the expression vector pEDV-1 and was expressed in *E. coli* JM101. The SP6 RNA polymerase isolated from the recombinant *E. coli* bearing this PCR-derived gene was purified to homogeneity and was found to be biochemically identical to SP6 RNA polymerase purified from phage SP6-infected *Salmonella typhimurium*.

#### **5. Summary of PCR protocol**

The method used for successfully isolating the *E. coli* clones harbouring the T7 RNA polymerase gene and the SP6 RNA polymerase gene involved:

1. Extraction of an RNA polymerase gene and its ribosome binding site from phage genomic DNA using the polymerase chain reaction.
2. Purification of the PCR products by agarose gel electrophoresis and isolation of appropriately sized DNA fragments.

3. Ligation of the purified DNA fragments into expression vector pEDV-1 and cloning into *E.coli* JM101.
4. The *E.coli* transformants were individually cultured in liquid media with antibiotic selection, induced during mid-log growth phase by the addition of IPTG, and ruptured by sonication 2-3 hours after induction. After centrifugation, the supernatant from the lysed cultures were assayed for either SP6 RNA polymerase or T7 RNA polymerase activity.

Among the *E.coli* transformants (SP6 clones and T7 clones) were bacteria which showed partial enzyme activity. Consequently the technique may be limited by the need to validate the PCR-derived DNA, due to potential mutations produced by the PCR amplification. The technique is therefore very useful when the structural integrity of a protein can be verified, in these cases by its highly specific enzymic assay.

## TABLE OF CONTENTS

<b>CHAPTER 1</b> .....	1
<b>INTRODUCTION</b> .....	1
<b>CHAPTER 2</b> .....	5
<b>MATERIALS and METHODS</b> .....	5
2.1 MATERIALS.....	6
2.1.1 Reagents .....	6
2.1.2 Enzymes .....	7
2.1.3 Radioisotopes .....	7
2.1.4 Synthetic oligonucleotides .....	7
2.1.5 Vector DNA .....	8
2.1.6 Bacterial strains .....	8
2.1.7 Media and Solutions.....	9
2.1.8 SDS polyacrylamide (protein) slab gels.....	9
2.1.9 Buffers .....	10
2.1.10 DNA molecular weight markers .....	10
<b>2.2 METHODS</b> .....	11
2.2.1 Preparation of single strand template DNA for sequencing.....	11
2.2.2 Preparation of plasmid DNA .....	11
2.2.3 Agarose gel electrophoresis and extraction of DNA .....	12
2.2.4 Polyacrylamide gel electrophoresis and purification of DNA.....	13
2.2.5 Polyacrylamide gel purification of proteins.....	14
2.2.6 DNA vector preparation.....	14
2.2.7 DNA ligations .....	14
2.2.8 End-filling DNA fragments.....	15
2.2.9 Preparation of single-stranded DNA markers .....	15
2.2.10 Growth of bacterial strains .....	15

2.2.11 Transformation of <i>E.coli</i> .....	16
2.2.12 Purification of synthetic oligonucleotides .....	16
2.2.13 DNA sequencing .....	17
2.2.14 Isolation and purification of bacteriophage DNA .....	17
2.2.15 Ethanol precipitation of DNA .....	18
<b>CHAPTER 3</b> .....	19
<b>BACTERIOPHAGE T7 RNA POLYMERASE: -ISOLATING AND CLONING THE PHAGE T7 RNA POLYMERASE GENE AND EXPRESSION OF RECOMBINANT T7 RNA POLYMERASE IN <i>E.COLI</i></b>	
3.1 INTRODUCTION .....	20
3.2 ISOLATING AND CLONING THE T7 RNA POLYMERASE GENE .....	22
3.2.1 INTRODUCTION.....	22
3.2.2 Experimental strategy for isolating the T7 RNA polymerase gene.....	24
3.3 METHODS and RESULTS.....	25
3.3.1 Preparation of starting material.....	25
3.3.2 Bal 31 digestion to remove promoter $\phi$ 1.1A .....	26
3.3.3 Cloning the Bal 31 nuclease-digested DNA fragments.....	27
3.3.4 DNA sequencing the T7 DNA component of pR75 .....	27
3.3.5 Shotgun cloning of Bal 31 nuclease-digested T7 DNA fragments.....	28
3.3.6 Cloning the 2,600 to 3,400 bp T7 DNA fragments into <i>E.coli</i> .....	29
3.3.7 Identification of bacteria harbouring the T7 RNA polymerase gene .....	30
3.3.7.1 Preparation of probe for the T7 RNA polymerase gene .....	30
3.3.7.2 Preparation of DNA-bound nitrocellulose filters and hybridization with the 2.3 kbp partial length T7 RNA polymerase gene probe .....	30
3.3.8 DNA sequencing of the T7 DNA component of pR80 .....	31
3.3.9 Engineering the T7 DNA component of pR80 .....	32
3.3.9.1 Introduction .....	32
3.3.9.2 Deletion of promoter $\phi$ 1.1A.....	35

3.3.9.2.1 Mutagenesis Reaction .....	36
3.3.9.3 Detection of mutant M13/6 DNA .....	36
3.3.9.4 Repair of the 5'-deletion in M13/6H .....	37
3.3.10 Expression of T7 RNA polymerase in <i>E.coli</i> E832 .....	38
3.3.11 Assays of T7 RNA polymerase expression .....	38
3.3.11.1 Detection of protein.....	38
3.3.11.2 Assay of enzyme activity .....	39
3.3.11.3 Is the Mr 98,000 protein from <i>E.coli</i> RS101 soluble or insoluble?.....	40
3.3.11.4 DNA sequencing strategy.....	40
3.3.11.5 Protein sequencing .....	41
3.4 DISCUSSION.....	42
<b>CHAPTER 4</b> .....	<b>45</b>
<b>BACTERIOPHAGE T7 RNA POLYMERASE:</b>	
<b>-ISOLATING THE PHAGE T7 RNA POLYMERASE GENE BY THE</b>	
<b>POLYMERASE CHAIN REACTION</b>	
<b>-EXPRESSION OF THE PCR-DERIVED T7 RNA POLYMERASE</b>	
<b>GENE IN <i>E.COLI</i></b>	
4.1 INTRODUCTION .....	46
4.2 METHODS and RESULTS.....	48
4.2.1 Construction of expression vector pEDV-1 .....	48
4.2.2 Isolation of the T7 RNA polymerase gene from phage T7 DNA by PCR .....	49
4.2.2.1 Design of the PCR oligonucleotide primers.....	49
4.2.2.2 PCR amplification of phage T7 DNA.....	49
4.2.2.3 Cloning the PCR-derived DNA into <i>E.coli</i> .....	50
4.2.3 Assay for enzymically active T7 RNA polymerase <i>E.coli</i> clones .....	51
4.2.4 Validation of expression vector pEDV-1.....	51
4.2.5 T7 RNA polymerase expression in <i>E.coli</i> RS51.....	52
4.2.5.1 Assay of enzyme activity .....	52
4.2.5.2 Examination of <i>E.coli</i> RS51 by SDS-PAGE .....	53
4.3 DISCUSSION.....	54

<b>CHAPTER 5</b> .....	56
<b>BACTERIOPHAGE SP6 RNA POLYMERASE:</b>	
<b>-ISOLATING THE PHAGE SP6 RNA POLYMERASE GENE BY THE POLYMERASE CHAIN REACTION</b>	
<b>-EXPRESSION OF THE PCR-DERIVED SP6 RNA POLYMERASE GENE IN <i>E.COLI</i></b>	
5.1 INTRODUCTION .....	57
5.2 METHODS and RESULTS .....	58
5.2.1 Isolation of the SP6 RNA polymerase gene by PCR.....	58
5.2.1.1 Design of the PCR oligonucleotide primers.....	58
5.2.1.2 PCR amplification of phage SP6 DNA .....	58
5.2.2 Cloning the PCR-derived DNA into <i>E.coli</i> .....	59
5.2.3 Assay for enzymically active SP6 RNA polymerase <i>E.coli</i> clones .....	59
5.2.4 Protein sequencing the SP6 RNA polymerase from <i>E.coli</i> RS49.....	60
5.2.5 Qualitative SP6 RNA polymerase transcription assay.....	61
5.3 DISCUSSION.....	62
<b>GENERAL DISCUSSION</b> .....	63
<b>REFERENCES</b> .....	66

## **CHAPTER 1**

### **INTRODUCTION**





The work reported in this thesis was directed towards developing strategies for optimising the processes of cloning and expressing genes in bacteria.

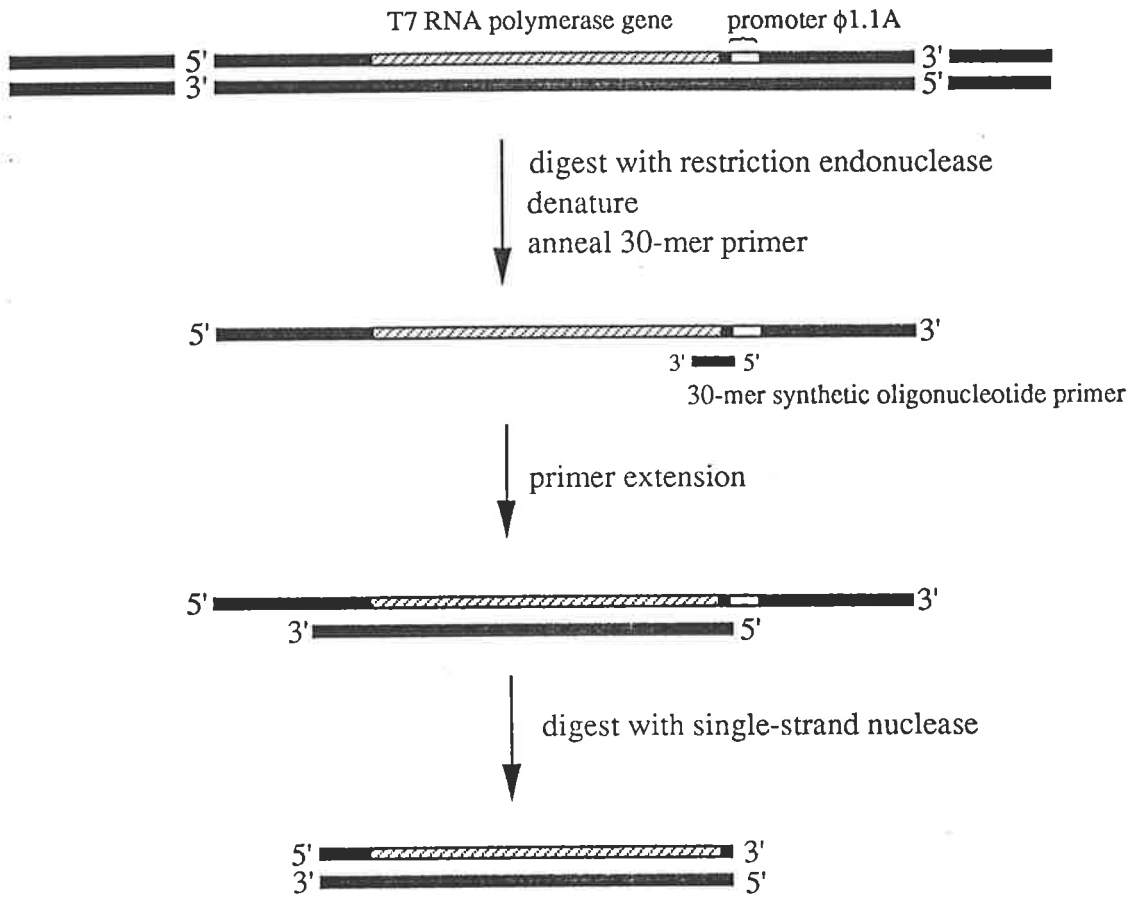
One interest in this laboratory is to be able to rapidly clone genes coding for proteins which are potentially useful for Molecular Biology. The isolation of a gene which encodes a given protein, and its subsequent cloning and DNA sequencing, should potentially facilitate researchers in other laboratories to reproduce the work. However, even in cases where the DNA sequence for a gene and its flanking regions is known, re-isolating and cloning the gene from genomic DNA is potentially difficult. The reasons why re-cloning a fully characterised gene should be problematical can be understood in terms of the techniques generally used in laboratories for manipulating DNA. The problem is simply that techniques for digesting and mutating DNA are either highly specific or random. The specific means of digesting DNA is with restriction endonucleases. The naturally occurring location of restriction sites within genomic DNA may be to an experimenter's advantage, or perhaps not. Alternate strategies of fragmenting DNA such as sonication or digestion with DNase I lack the precision required for reliably manipulating DNA. Hence, there is need for strategies which are inherently repeatable and which do not require the convenient placement of restriction sites in genomic DNA.

One interesting example of a gene which took considerable effort to clone is the gene encoding the RNA polymerase from bacteriophage T7. The cloning of this gene into *E.coli* serves as a good example of how powerful, though limited, techniques for manipulating DNA can make the task of reproducing a gene-cloning experiment very difficult.

Essentially, the problem to be overcome in cloning the T7 RNA polymerase gene, in a form that expressed enzymically active T7 RNA polymerase (Davanloo et al. 1984) was due to the genomic organisation around the gene. Situated 11 bp downstream of the T7 RNA polymerase gene is a promoter element,  $\phi 1.1A$ . Cloning a T7 DNA fragment encoding both the T7 RNA polymerase gene and the T7 promoter  $\phi 1.1A$  was found to be lethal to *E.coli* (Davanloo et al. 1984). However, removal of the promoter from the gene enabled the gene to be cloned. The fundamental problem to be overcome was that of removing the promoter  $\phi 1.1A$  away from the T7 RNA polymerase gene. The problem of isolating the gene away from the promoter was exacerbated by there being no restriction sites between the gene and promoter or

Figure 1.1

Initial strategy for isolating the T7 RNA polymerase gene away from promoter  $\phi$ 1.1A



*Product:* T7 DNA fragment encoding a full length T7 RNA polymerase gene without promoter  $\phi$ 1.1A.

within the promoter which would not also digest the gene into fragments. Cloning of the T7 RNA polymerase gene by Davanloo et al.(1974) and later by Tabor and Richardson (1985) involved the manipulation of T7 deletion mutant DNAs to get a T7 DNA fragment encoding the RNA polymerase gene but lacking promoter  $\phi$ 1.1A. The fact that bacteriophage T7 had been genetically characterized and its 40 kbp genome completely sequenced was probably of some help toward cloning the T7 RNA polymerase gene.

Our aim in the sorting of precise techniques for gene manipulation and cloning was to use the problems inherent in cloning the T7 RNA polymerase gene as a model system. Specifically, our aim was to isolate and then clone the T7 RNA polymerase gene into *E.coli* using biochemical techniques, which could be applied to the rapid re-cloning of other genes.

When this research project began, the most promising methods for extracting specific DNA sequences from genomic DNA involved a variety of site-directed mutagenesis techniques (Wallace et al.1981, Zoller and Smith 1982). Site-directed mutagenesis using phage M13 enables a target DNA sequence to be modified precisely, but first requires the target DNA to be cloned into the M13 phage.

Our initial strategy for cloning the T7 RNA polymerase gene involved the use of a synthetic oligonucleotide (see figure 1.1). A 30-mer oligonucleotide primer was designed to be complimentary to the T7 RNA polymerase gene upper strand, with the 5'-end of the primer binding at the start of the T7 promoter  $\phi$ 1.1A. Extension of the primer with Klenow polymerase would thus be from the start of the T7 promoter towards the start of the T7 RNA polymerase gene. If primer extension of about 2.6 kb to the start of the T7 RNA polymerase gene had been successful, the protruding termini would then have been digested with nucleases, leaving a DNA fragment encoding the T7 RNA polymerase gene but not promoter  $\phi$ 1.1A. Unfortunately, the strategy failed to yield the desired DNA fragment. The major difficulty with this approach was incomplete primer extension, yielding DNA fragments too short to encode a full-length T7 RNA polymerase gene.

Our first successful attempt at cloning the T7 RNA polymerase gene used a combination of random and highly specific techniques. Essentially, nuclease Bal 31 was used to produce a partial-length T7 RNA polymerase gene which was cloned into *E.coli*. The partial-

length gene was subsequently reconstructed by site-directed mutagenesis. Although this strategy yielded an *E.coli* clone bearing a full-length T7 RNA polymerase gene, the T7 RNA polymerase expressed by the bacteria was enzymically inactive. The expression of enzymically inactive T7 RNA polymerase appeared to be a function of the expression system. This problem was overcome by expressing the T7 RNA polymerase gene from a different expression vector in a different strain of *E.coli*.

The value of the polymerase chain reaction in extracting the RNA polymerase genes from T7 and SP6 genomic DNA was then examined. The genes for both T7 and SP6 RNA polymerase were successfully isolated with this approach. Subsequent cloning of these genes into *E.coli* resulted in the expression of enzymically active T7 and SP6 RNA polymerase.

**CHAPTER 2**  
**MATERIALS and METHODS**

## 2.1 MATERIALS

### 2.1.1 Reagents

General laboratory reagents were of analytical grade.

Acrylamide: Sigma Chemical Co.

Agarose: Sigma Chemical Co.

Amberlite MB-1 ion exchange resin: Sigma chemical Co.

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

Ammonium persulphate: May and Baker.

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

Bovine serum albumin: Bresatec (Adelaide). Acetylated before use to remove nucleases according to the procedure of Gonzalez et al.(1977) and kept as a 10 mg/ml solution in water at -20°C.

CsCl: Bethesda Research Laboratories.

dNTPs and NTPs: Sigma Chemical Co. Stock solutions were prepared by neutralizing the compounds in aqueous solution with a molar excess of triethylamine which has been redissolved from ninhydrin, followed by lyophilization and redissolving the compounds in water and storing at -20°C.

Dithiothreitol: Sigma Chemical Co. Stored as a 100 mM solution in water at -20°C.

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

EDTA: Disodium salt. Sigma Chemical Co.

Formamide: B.D.H. Laboratories. Aust. De-ionised and stored in the dark at -20°C.

IPTG: Sigma Chemical Co. Stored as a 100 mM solution in water at -20°C.

β-Mercaptoethanol: Sigma Chemical Co.

N,N'-methylene-bis-acrylamide: Bio-rad Laboratories.

Nonidet P40: B.D.H. Laboratories. Aust.

Phenol: B.D.H. Laboratories. Aust. Redistilled and stored in the dark at -20°C.

PEG 8000: Sigma Chemical Co.

SDS: Sigma Chemical Co.

TEMED: Tokyo Kasei.

Tetramethylammonium chloride: Aldrich Chemical Co.

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

Toluidine blue: Aldrich Chemical Co.

Urea: Sigma Chemical Co.

Polyvinylidene difluoride membrane (immobilon) was from Millipore. Diethylaminoethyl cellulose (NA45 membrane) and Nitrocellulose membrane (BA85 0.45 $\mu$ ) were from Schleicher and Schuell. GeneClean kits for DNA purification (Vogelstein and Gillespie 1979) from Bio-101 (La Jolla) were purchased from Bresatec.

### **2.1.2 Enzymes**

Bovine pancreatic DNase I, Klenow fragment of DNA polymerase I, T4 DNA ligase, T7 RNA polymerase and T4 polynucleotide kinase were supplied by Bresatec. Bal 31 nuclease and calf intestinal phosphatase were from Boehringer Mannheim. Lysozyme and RNase A were from the Sigma Chemical company. Restriction endonucleases were from New England Biolabs. DNase I was from Worthington Biochemical.

### **2.1.3 Radioisotopes**

$\alpha$ -<sup>32</sup>P-dATP,  $\alpha$ -<sup>35</sup>S-dATP,  $\alpha$ -<sup>32</sup>P-dCTP,  $\alpha$ -<sup>32</sup>P-UTP and  $\gamma$ -<sup>32</sup>P-ATP (3,000 Ci/mmmole) were from Bresatec.

### **2.1.4 Synthetic oligonucleotides**

Synthetic oligonucleotides were supplied by Bresatec (Adelaide) in a crude form.

Mutagenesis primers for Exps. 3.3.9.2.1 and 3.3.9.4

SR-462 5'-dGTCGTTCTTAGCGATCTTAATCGTGTTTCATATGTTACCTCCGGGT  
ACCGAGCTCGA-3'

SR-463 5'-dAGTGAGCTGTATAAGCTTGGCGTTACGCGA-3'

PCR primers for Exp. 4.2.2.2

RRS-887 5'-dGATTACTAACTGGAAGAGGCACT-3'

RRS-879 5'-dGAGTCGTATTGATTTGGCGTTACG-3'

#### PCR primers for Exp. 5.2.1.2

RRS-782 5'-dGTACCGCATGAGGATACAAGATGC-3'

RRS-783 5'-dTATTAGGCAAATACGTATTCAGAA-3'

### **2.1.5 Vector DNA**

Plasmids pUC18 and pUC19 (Yanisch-Peron et al. 1985) were supplied by Bresatec. The pUC plasmids are high copy number, general purpose cloning vectors which can also be used as IPTG-inducible, lacZ-fusion expression vectors. The vectors contain a polylinker encoding numerous unique restriction sites.

Plasmid pSP72 (Krieg and Melton 1987) was supplied by Bresatec. This vector contains bacteriophage SP6 and T7 promoters for transcriptions *in vitro* of DNA cloned into the multiple cloning site polylinker.

Plasmid pPLC236 (Remaut et al. 1981) was a gift from Dr. Barry Egan, Biochemistry Department, University of Adelaide. The vector contains the rightward promoter from bacteriophage lambda. In the appropriate host (*E. coli* E832) a gene cloned into the few cloning sites can be induced by heat induction at 42°C.

Plasmid pEV601 (Brumby 1986) was a gift from Dr. Barry Egan, Biochemistry Department, University of Adelaide. This derivative of pKT52 (Talmadge and Gilbert 1980) contains a strong IPTG-inducible synthetic hybrid promoter, P<sub>trc</sub> (DeBoer et al. 1983) and ribosome binding site.

Bacteriophage M13mp18 and M13mp19 RF DNA (Yanisch-Peron et al. 1985) were supplied by Bresatec.

### **2.1.6 Bacterial strains**

*E. coli* JM101: Δ(lac, pro) supE44 thi F' traD36 proAB lacI<sup>q</sup> ZΔM15 (Messing 1979).

*E. coli* MC1061: araD139 Δ(ara, leu)7697 ΔlacX74 gal U<sup>-</sup> gal K<sup>-</sup> hsr<sup>-</sup> hsm<sup>+</sup>strA (Casadaban and Cohen 1980).



*E.coli* E832: M72 lacZam Sm<sup>R</sup> Δbio-uvrB ΔtrpEA2 (λ Nam 7 Nam53 cI857ΔH1)

ΔH1 deletes λ genes cro-R-A-J-b2 (Bernard et al. 1979).

*E.coli* B.

*S.typhimurium* LT2.

Bacteriophage T7 and Bactriophage SP6: Generously provided by Dr. Peter Reeves (Sydney University).

### 2.1.7 Media and Solutions

L (Luria) broth: 1% (w/v) bacto-tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl, pH 7.0.

2 x YT broth: 1.6% (w/v) Bacto-tryptone, 1% (w/v) yeast extract, 0.5% (w/v) NaCl, pH 7.0

ML-broth: 1% (w/v) Bacto-tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl, 0.3% (w/v) KH<sub>2</sub>PO<sub>4</sub>, 0.6% (w/v) Na<sub>2</sub>HPO<sub>4</sub>.

M13 minimal medium: 1.05% (w/v) K<sub>2</sub>HPO<sub>4</sub>, 0.45% (w/v) KH<sub>2</sub>PO<sub>4</sub>, 0.1% (w/v)

(NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.05% (w/v) Na<sub>3</sub>citrate.2H<sub>2</sub>O, made to a volume of one litre. This solution was autoclaved, cooled to 45°C and the following added from separately prepared solutions; 10 ml of 20% (w/v) glucose, 0.8 ml of 1 M MgSO<sub>4</sub>, 0.5 ml of 1% (w/v) thiamine-HCl.

100 x Denhardt's solution: 2% (w/v) Ficoll 400, 2% (w/v) BSA, 2% (w/v) polyvinyl pyrrolidone.

### 2.1.8 SDS polyacrylamide (protein) slab gels

*Electrode buffer* (10 x). 250 mM Tris-glycine, pH 8.3, 0.2% (w/v) SDS.

10 x stock is prepared by adding glycine (approx. 144 g) to 30.25 g of Tris in water until the solution reaches a pH of 8.3. 10 g of SDS is then added and the solution adjusted to a volume of 1 litre.

*Separating gel buffer* (2x). 750 mM Tris-HCl, pH 8.8, 0.2% (w/v) SDS.

*Stacking gel buffer* (2x). 250 mM Tris-HCl, pH 6.8, 0.2% (w/v) SDS.

*Loading buffer* (2x). 125 mM Tris-HCl, pH 6.8, 4% (w/v) SDS, 20% (v/v) glycerol, 0.004% (w/v) bromophenol blue.

*Separating gel acrylamide solution.* 40% (w/v) acrylamide, 0.32% (w/v) N,N'-methylene-bis-acrylamide.

*Stacking gel acrylamide solution.* 30% (w/v) acrylamide, 0.8% (w/v) N,N'-methylene-bis-acrylamide.

The acrylamide solutions are dionized with amberlite MB-1 ion exchange resin (2 g/100 ml) for at least 1 hour, with gentle stirring, followed by filtration through Whatman 541 paper. Stocks are kept at room temperature.

### **2.1.9 Buffers**

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

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

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

10 x CIP: 0.5 M Tris-HCl, pH 7.6, 10 mM spermidine, 1 mM ZnCl<sub>2</sub>.

10 x End-fill buffer: 0.5 M Tris-HCl, pH 7.6, 0.5 M NaCl, 0.1 M MgCl<sub>2</sub>.

10 x PNK: 600 mM Tris-HCl, pH 7.6, 90 mM MgCl<sub>2</sub>.

20 x SSC: 3 M NaCl, 0.3 M Na<sub>3</sub>citrate, pH 7.4.

Formamide loading solution: 95% formamide, 10 mM EDTA, 0.02% (w/v) xylene cyanol).

10 x Sucrose loading buffer: 40% (w/v) sucrose, 0.05% (w/v) xylene cyanol, 0.05% (w/v) bromophenol blue.

10 x Glycerol loading buffer: 50% (v/v) glycerol, 0.5 mM EDTA, 0.02% (w/v) bromophenol blue, 0.02% (w/v) xylene cyanol.

### **2.1.10 DNA molecular weight markers**

Phage SPP-1 DNA digested with EcoR1. Obtained from Bresatec at 0.5 µg/µl in water.

Fragment sizes in bp: 7840, 6960, 5860, 4690, 3370, 2680, 1890, 1800, 1450, 1330, 1090, 880, 660, 480 and 380.

pUC19 DNA digested with HpaII. Obtained from Bresatec at 0.5 µg/µl in water.

Fragment sizes in bp: 501, 489, 404, 331, 242, 190, 147, 111, 110, 67, 2 x 34 and 26.

## 2.2 METHODS

### 2.2.1 Preparation of single strand template DNA for sequencing M13 clones

Single-stranded M13 DNA was prepared by the following protocol, based on the method of Winter and Fields (1980).

An overnight culture of *E.coli* JM101, grown in minimal media at 37°C, was inoculated (1:100 dilution) into 2 x YT broth with phage toothpicked from single plaques, and incubated in 10 ml plastic tubes in a roller drum at 37°C for 6 hours. A 1.5 ml sample of the culture was then spun in an Eppendorf centrifuge for 10 minutes at room temperature and 1 ml of the supernatant was transferred to a 1.5 ml tube containing 270 µl of 20% (w/v) PEG 8000 and 2.5 M NaCl. After 15 minutes at room temperature, the tube was spun in an Eppendorf centrifuge for 10 minutes and the phage pellet resuspended in 200 µl of 50 mM Tris-HCl, pH 7.6, 5 mM EDTA and 0.5% (w/v) SDS. To this was added 100 µl of Tris-saturated phenol, with vortexing. The aqueous and phenol phases were separated by spinning in an Eppendorf centrifuge for 1 minute. 170 µl of the upper (aqueous) phase was transferred to a clean 1.5 ml tube with 17 µl of 3 M sodium acetate, pH 5.2 and 375 µl of ethanol. After 15 minutes at -20°C the DNA was pelleted by spinning in an Eppendorf centrifuge for 10 minutes at 4 °C, washed with 2 x 2 ml of ethanol, dried *in vacuo* and redissolved in 30 µl of TE buffer or water.

### 2.2.2 Preparation of plasmid DNA

Small scale preparations of plasmid DNA were done by a variation of the method described by Birnboim and Doly (1979) or by the method described by Serghini et al.(1989).

A single bacterial colony from a fresh plate was grown overnight at 37°C in the appropriate media with antibiotic selection. A 1.5 ml sample of the culture was spun in an Eppendorf centrifuge for 5 minutes and the pelleted cells were resuspended in 100 µl of 30 mM Tris-HCl, pH 8.0, 15% (w/v) sucrose, 10 mM EDTA, lysozyme at 5 mg/ml and left for 5 minutes at room temperature. 200 µl of 0.2 M NaOH, 0.1% (w/v) SDS was then added and the solution was left for 10 minutes at 4°C. 300 µl of 3 M sodium acetate, pH 5.2 was then

added and again the solution was left for 10 minutes at 4°C. Following this, the solution was spun in an Eppendorf centrifuge for 10 minutes and 0.5 ml of the supernatant was transferred to a clean 1.5 ml tube containing 1 ml of ethanol and left for 5 minutes at room temperature. The DNA was pelleted by spinning in an Eppendorf centrifuge for 10 minutes, washed with 2 x 2 ml of ethanol, dried *in vacuo* and redissolved in 40 µl of TE buffer or water.

Despite the short incubation times and lack of phenol extraction, this protocol almost invariably yields DNA which can be digested to completion with restriction endonucleases. Failure of the plasmid DNA to digest cleanly can be overcome by precipitating the DNA with spermine, followed by ethanol precipitation: The plasmid DNA solution was diluted with water to a volume of 80 µl, 1.6 ml of 0.1 M spermine was added, and the DNA left for 15 minutes on ice. The DNA was spun in an Eppendorf centrifuge for 15 minutes at 4°C and then 200 µl 75% (v/v) ethanol, 0.3 M sodium acetate pH 5.2, 10 mM magnesium acetate was added. After 60 minutes on ice the DNA was pelleted by spinning in an Eppendorf centrifuge for 10 minutes, washed with 2 x 2 ml of ethanol, dried *in vacuo* and redissolved in 40 µl of TE buffer or water.

### **2.2.3 Agarose gel electrophoresis and extraction of DNA**

Samples of DNA were run in TAE buffered horizontal (submarine) agarose gels, the concentration of which depended on the size of the fragments to be resolved (Maniatis et al. 1982). Samples were prepared for loading by the addition of 0.1 volumes of sucrose gel loading buffer or glycerol gel loading buffer (2.1.8). Gels made with low gelling temperature agarose (LGT) were run at 4°C. Bands were stained with 10 µg/ml ethidium bromide, and after rinsing the gel with water, the bands were visualized under ultraviolet light.

DNA was extracted from agarose gel by electroelution (Maniatis et al. 1982), transferred to DEAE-cellulose membrane (Schleicher and Schuell NA 45) *in situ*, or by use of a GeneClean kit under conditions described by the manufacturer.

Recovery of DNA from agarose using DEAE-membranes was by a modification of the procedure described by Holland and Wangh (1983). After the DNA had run the required distance on the gel, the DNA was stained with 10 µg/ml ethidium bromide and visualized with

ultraviolet light. Strips of NA 45 paper were inserted into the gel just below the bands and the gel re-electrophoresed until the bands had run onto the NA 45 strips (The NA 45 was pre-treated by washing in 10 mM EDTA for 10 minutes, then in 0.5 M NaOH for 5 minutes followed by 5 x 2 minute washes in water). DNA was eluted from the NA 45 after washing twice in 150  $\mu$ l of 20 mM Tris-HCl, pH 8.0, 0.1 mM EDTA, 150 mM NaCl for 3 minutes at room temperature. The DNA was eluted by incubating the strips twice in 150  $\mu$ l (total of 300  $\mu$ l) of 20 mM Tris-HCl, pH 8.0, 0.1 mM EDTA, 1 M NaCl for 30 minutes at 68°C. The DNA was then phenol extracted and ethanol precipitated.

DNA was isolated from LGT agarose by the following method. A maximum of 140  $\mu$ l of gel slice was melted in a 1.5 ml tube at 65°C, two volumes of preheated extraction buffer (50 mM Tris-HCl, pH 8.0, 0.5 mM EDTA) were added, followed by extraction with 140  $\mu$ l of Tris-saturated phenol. After brief centrifugation the supernatant was removed, re-extracted with phenol and the DNA precipitated at room temperature for 30 minutes with two volumes isopropanol, 0.4 volumes 5 M ammonium acetate. The DNA was pelleted by spinning in an Eppendorf centrifuge for 15 minutes, washed with 2 x 2 ml of ethanol, dried *in vacuo* and redissolved in 40  $\mu$ l of TE buffer or water.

#### **2.2.4 Polyacrylamide gel electrophoresis and purification of DNA**

DNA to be purified was added to an equal volume of formamide loading solution (2.1.8), heated for 30 seconds at 80°C and snap cooled on ice before loading onto a 35 x 20 x 0.05 cm polyacrylamide gel. The gel percentage was usually 5% but varied according to the size of the bands to be resolved (acrylamide: bisacrylamide = 25:1 (w/w)). Gels were run in TBE (2.1.8). Non-radioactive bands were visualized by staining in 0.05% (w/v) toluidine blue followed by destaining in water. DNA was eluted from the gel slices in 10 mM Tris-HCl, pH 7.6, 0.1% (w/v) SDS, 1 mM EDTA, for 6-16 hours at 37°C and then precipitated by the addition of 3 M sodium acetate (pH 5.2) to 0.3 M and 2.5 volumes of cold ethanol. After storage on ice for 10-30 minutes the DNA was pelleted by spinning in an Eppendorf centrifuge for 15-30 minutes at 4°C, washed with 2 x 2 ml of ethanol, dried *in vacuo* and redissolved in 40  $\mu$ l of TE buffer or water. The poor recovery of small molecules can sometimes be rectified by

substituting the 2.5 volumes of ethanol with two volumes of ethanol and one volume of acetone.

### **2.2.5 Polyacrylamide gel purification of proteins**

Proteins were resolved on SDS-polyacrylamide slab gels Laemmli (1970). The separating gel concentration was 10% (w/v) acrylamide and the stacking gel concentration was 3% (w/v) acrylamide (2.1.8). Samples to be electrophoresed were diluted 1:1 with 2 x loading buffer and  $\beta$ -mercaptoethanol was added to a final concentration of 5% (v/v) and the sample heated for 5 minutes at 90°C. 140 x 150 x 1.5 mm gels were run at 120 V without pre-electrophoresis.

### **2.2.6 DNA vector preparation**

Digestion with restriction endonucleases of double stranded plasmid and M13 DNAs for use as vectors was carried out under the conditions specified in the New England Biolabs catalogue. Phosphatasing was carried out using calf intestinal alkaline phosphatase (CIP) essentially as described by Maniatis et al. (1982). Vector DNA was phosphatased using 1 unit of CIP. The DNA was then purified on an agarose-TAE gel and the DNA extracted, usually with Bio-101 GeneClean. Following estimation of DNA concentration on an agarose gel the, DNA was diluted to a final concentration of 20 or 40 ng/ $\mu$ l in 0.1 mM EDTA or water prior to use. When the linearized, dephosphorylated vector DNA was subjected to a ligation reaction under the conditions described in 2.2.7 and then used to transform competent bacteria, no transformants were found. Vector DNA was only used if it produced a background of zero colonies.

### **2.2.7 DNA ligations**

In general, ligation of DNA fragments into vector DNA was carried out as described below. 20 ng of vector DNA was incubated with the desired insert fragment, at a molar ratio of approximately 1:3 (vector:insert) in 50 mM Tris-HCl, pH 7.6, 10 mM MgCl<sub>2</sub>, 10 mM DTT and 1 mM ATP with T4 DNA ligase. 0.02 units/ $\mu$ l of T4 DNA ligase was used for sticky-ended fragments and 0.1 units/ $\mu$ l for blunt fragments. To increase the ligation efficiency of blunt-ended fragments, 4% PEG 8000 was added when required. Ligation reactions were

carried out at 15°C or at room temperature, usually overnight.

### **2.2.8 End-filling DNA fragments**

DNA with 5'-overhanging ends was blunted using the Klenow fragment of DNA polymerase I as described by Maniatis et al. (1982). One or two nucleotide 3'-overhanging ends were blunted by digestion with the Klenow enzyme in the absence of dNTPs. DNA with longer 3'-overhanging ends were blunted with T4 DNA polymerase in a 30 µl reaction containing 1-3 pmol of 3'ends, 70 mM Tris-HCl, pH 7.4, 10 mM MgCl<sub>2</sub>, 5 mM DTT, 0.5 mM dNTPs and 5 units of T4 DNA polymerase, for 20 minutes at 37°C. The reaction was stopped by heating for 10 minutes at 70°C.

### **2.2.9 Preparation of single-stranded DNA markers**

pUC19 plasmid DNA was digested with Hpa II and end-labelled with  $\alpha$ -<sup>32</sup>P-dCTP using the Klenow fragment of DNA polymerase I as described by Maniatis et al.(1982). Prior to loading, the markers were denatured by heating for 3 minutes at 100°C with an equal volume of formamide loading solution (2.1.8) containing 50 mM NaOH, and snap cooled on ice.

### **2.2.10 Growth of bacterial strains**

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

Log phase cultures and indicator bacteria were prepared by diluting a fresh stationary phase culture one hundred fold into sterile broth and incubating with aeration in a gyrotary water bath at the appropriate temperature, until the required cell density was reached (usually  $3 \times 10^8$  cfu/ml, which occurs at  $A_{600}=0.8$  in L-broth). Cell density was measured by observing the  $A_{600}$  using a Gilford 300 T-1 spectrophotometer. Indicator bacteria were chilled and kept on ice until required.

### **2.2.11 Transformation of *E.coli***

In general, when only moderate efficiencies of transformation were required, a shortened

method was used, based on the technique described by Hanahan (1985).

A scrape of culture from a fresh plate was grown in 50 ml of liquid media to an  $A_{600}$  of 0.4, and 40 ml was transferred to a clear polycarbonate oak ridge tube and centrifuged for 5 minutes at 5°C-25°C in a swing-out rotor (Sorval HB-4) at 4,000 rpm. The pelleted cells were resuspended in 2 ml of cold 50 mM  $CaCl_2$  and used immediately.

For higher efficiency transformation, the protocol was as follows. A scrape of culture from a fresh plate was grown in 50 ml of liquid media to an  $A_{600}$  of 0.4, and 40 ml was transferred to a clear polycarbonate oak ridge tube and centrifuged for 5 minutes at 5°C in a swing-out rotor (HB-4) at 4,000 rpm. The pelleted cells were resuspended in 5 ml of cold 50 mM  $CaCl_2$  and left for 30 minutes on ice. The cells were again centrifuged for 5 minutes at 5°C in a swing-out rotor at 4,000 rpm and the pelleted cells were resuspended in 2 ml of cold 50 mM  $CaCl_2$ , and left on on ice for 40-60 minutes prior to use.

#### **2.2.12 Purification of synthetic oligonucleotides**

Synthetic oligonucleotides for site-directed mutagenesis and as primers for PCR amplification of DNA were purified as follows.

Approximately 20-40  $\mu$ g of the crude oligonucleotide mixture was lyophilized, dissolved in 20  $\mu$ l of formamide loading buffer, heated at 80°C for 1 minute and snap cooled on ice prior to loading onto a 4 cm wide well of a 15% polyacrylamide, 7 M urea, TBE gel (2.2.4). Bands were visualized by staining with 0.05% (w/v) toluidine blue followed by washing with sterile water. The slowest band (presumed to be the full-length product of synthesis) was excised and eluted overnight at 37°C in 400  $\mu$ l of 10 mM Tris-HCl, pH 7.5, 0.1% (w/v) SDS, 1 mM EDTA. The DNA was precipitated by adding 40  $\mu$ l of 3 M sodium acetate, pH 5.2 and 880  $\mu$ l of cold ethanol, and left for at least 15 minutes at -20°C. The DNA was pelleted by spinning in an Eppendorf centrifuge for 15-30 minutes at 4°C, washed with 1ml of ethanol, dried *in vacuo* and redissolved in TE buffer or water.

#### **2.2.13 DNA sequencing**

The dideoxy chain termination sequencing technique (Sanger et al.1977, 1980) was used to



determine DNA sequence. DNA sequencing was performed using Bresatec Dideoxy Sequencing kits with either  $\alpha$ -<sup>32</sup>P-dATP or  $\alpha$ -<sup>35</sup>S-dATP.

#### Annealing.

6  $\mu$ l of single-stranded DNA template (2.2.1) was mixed in an Eppendorf tube with 1  $\mu$ l of 10 x annealing buffer (100 mM Tris-HCl, pH 8.0, 100 mM MgCl<sub>2</sub>), 1  $\mu$ l (2.5 ng) M13 sequencing primer and 1  $\mu$ l of water. The tube was transferred to an aluminium block at 80°C for 1 minute, and the block was then transferred to an aluminium plate, allowing the DNA solution to cool slowly to < 37°C.

#### Sequencing.

DNA sequencing using  $\alpha$ -<sup>32</sup>P-dATP was carried out in Eppendorf tubes as recommended by the Bresatec protocol. The 0.25 mm thick sequencing gel (Sanger and Coulson 1978) was run at 1200-1500 V. The DNA was fixed in the gel by washing with 10% (v/v) acetic acid until the dyes were no longer visible and then washed with 2 litres of 20% (v/v) ethanol over a period of 30 minutes. The gel was dried in a 110°C exhaust-oven for 30 minutes.

DNA sequencing using  $\alpha$ -<sup>35</sup>S-dATP was done on a plastic petri dish which was floating in a 37°C water bath, as described by Ner et al.(1988). After running the sequencing gel, a sheet of Whatman number 4 paper was laid onto the gel for a few seconds and the paper and gel peeled off the glass plate. The gel was immediately dried at 50°C under vacuum on a gel-dryer.

### **2.2.14 Isolation and purification of bacteriophage DNA**

#### Phage assays.

Phage lysates were assayed for phage by mixing 100  $\mu$ l of the phage diluted in phage storage buffer (PSB: 10 mM Tris-HCl, pH 7.0, 10 mM MgSO<sub>4</sub>), 200  $\mu$ l of log phase indicator bacteria (2.2.10) and 3 ml of melted 0.7% (w/v) soft agar overlay and pouring the mixture onto the appropriate plates. The agar was allowed to solidify and the plates were inverted and incubated overnight 37°C. Plaques were counted and scored as plaque forming units per ml (pfu/ml).

### Isolation of phage.

A stationary phase culture of bacteria was inoculated into 800 ml of the appropriate broth and grown in an air shaker at 37°C until the  $A_{600}$  reached 0.6. Phage were then added at a multiplicity of infection of 0.05 and the incubation continued until cell lysis occurred. DNase I and RNase A were each added to a concentration of 1  $\mu\text{g/ml}$  and the solution left for 30 minutes at room temperature. NaCl was then added to a concentration of 0.5 M and the solution left for 30 minutes at 4°C with gentle stirring. The solution was clarified by centrifugation for 20 minutes at 9,000 rpm in a Sorval GS-3 rotor at 4°C. PEG 8000 was added to the supernatant to a concentration of 6% (w/v) and left overnight at 4°C with gentle stirring. The phage were pelleted by centrifugation for 20 minutes at 10,500 rpm in a Sorval GS-A rotor at 4°C. The phage were resuspended in 5 ml of PSB supplemented with 0.5 M NaCl. 2 ml of phage were then added to 2 ml of PSB containing 2 g CsCl and the solution layered onto a CsCl step gradient. The CsCl solutions were 1.4 g/ml and 1.6 g/ml. The phage were centrifuged for 2-3 hours at 40,000 rpm in a Beckman Ti 50 rotor at 10°C, and the phage band removed and dialysed overnight against 1 litre of PSB, 0.5 m NaCl.

### Isolation of phage DNA.

The phage suspension was adjusted to 20 mM EDTA, 0.5% (w/v) SDS, 50  $\mu\text{g/ml}$  proteinase K and incubated for 70 minutes at 65°C. The solution was twice phenol extracted and then ethanol precipitated.

#### **2.2.15 Ethanol precipitation of DNA**

To a given volume of DNA was added 0.1 volumes of 3 M sodium acetate, pH 5.2 and 2.2 volumes of ethanol.

## **CHAPTER 3**

### **BACTERIOPHAGE T7 RNA POLYMERASE**

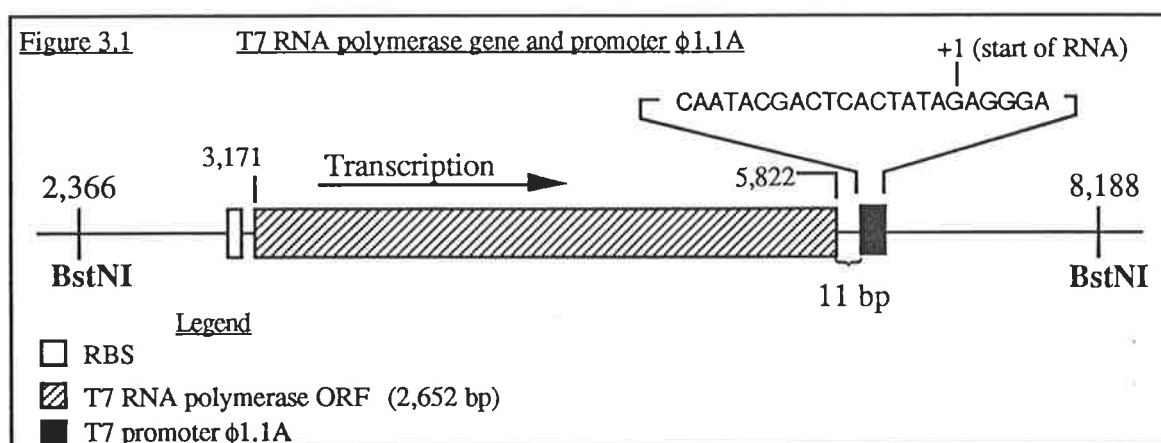
**ISOLATING AND CLONING THE PHAGE T7 RNA POLYMERASE GENE  
AND EXPRESSION OF RECOMBINANT T7 RNA POLYMERASE IN *E.COLI***

### 3.1 INTRODUCTION

Historically, the way to obtain T7 RNA polymerase was to superinfect a large number of *E.coli* with T7 phage, concentrate the cells by centrifugation and then purify the enzyme (Chamberlin et al. 1970). A significant problem with this approach is that the expression of T7 RNA polymerase in T7 phage-infected *E.coli* occurs for approximately ten minutes. The T7 RNA polymerase from T7 phage-infected *E.coli* represents only a small proportion of the total cellular protein from infected cells (Studier 1972).

The large scale purification of the enzyme is greatly facilitated by first cloning the T7 RNA polymerase gene into *E.coli* and then being able to induce expression of the gene for several hours before attempting to harvest the bacteria.

The coliphage T7 genome is 39,936 bp and linear (Dunn and Studier 1983). The T7 RNA polymerase gene coordinates are 3,171 bp to 5,822 bp relative to the start of the template strand of the genome. Situated twelve base pairs downstream of the RNA polymerase gene, starting at position 5,834, is a promoter ( $\phi$ 1.1A) to which the T7 RNA polymerase will bind (see figure 3.1). This promoter element belongs to the consensus 'family' of 17 promoters on the T7 genome that recognise T7 RNA polymerase (Dunn and Studier 1983). These promoters have a conserved 23 bp sequence which extends from -17 to +6 relative to the start site of the mRNA. The DNA sequence of  $\phi$ 1.1A is not recognised as a promoter element by *E.coli* RNA



polymerase, however it has high specificity for T7 RNA polymerase.

The phage T7 promoters have been exploited as a research tool by being cloned next to multiple cloning sites in DNA cloning vectors. These transcription vectors, such as pSP72

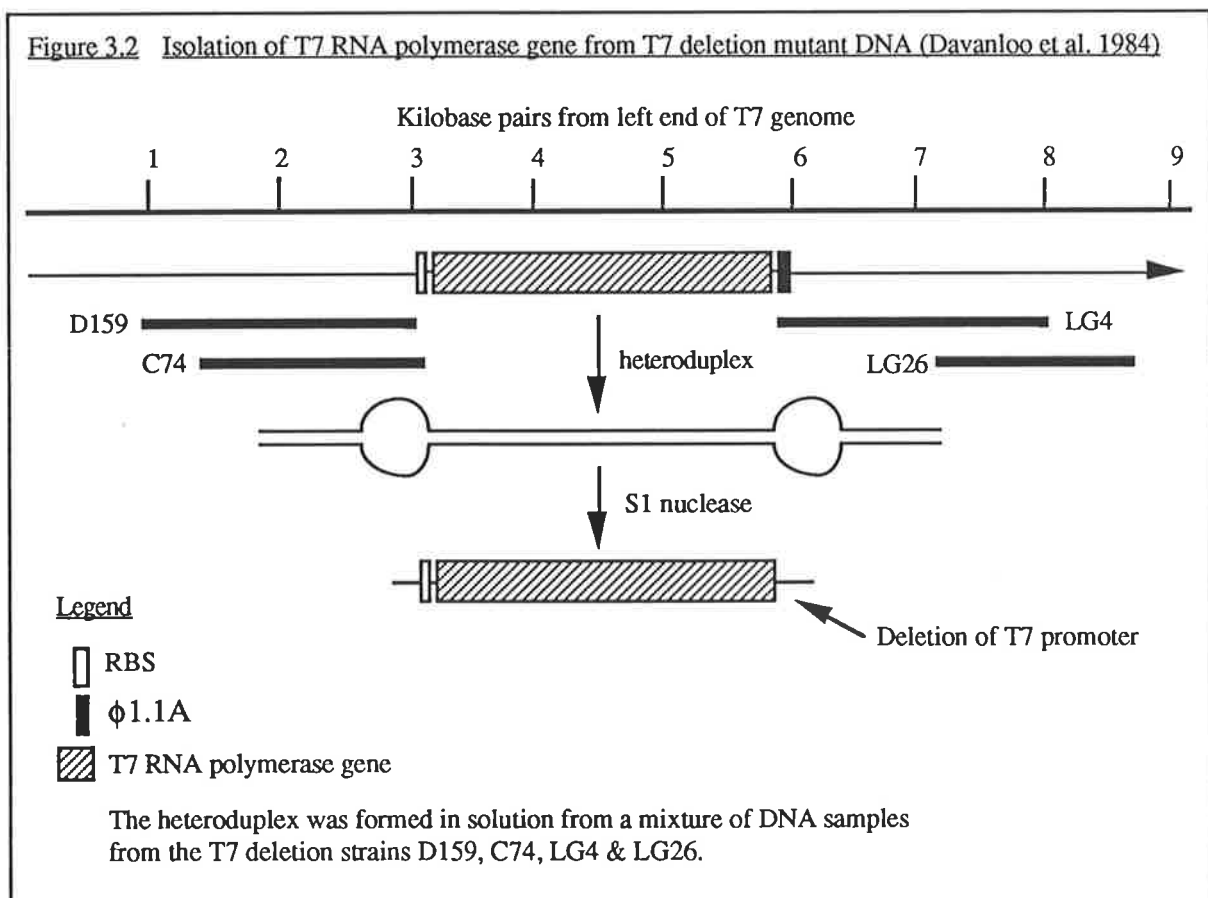
(Krieg and Melton 1987), enable RNA to be synthesized from a foreign DNA fragment within the multiple cloning site *in vitro*, in the presence of T7 RNA polymerase. The specificity of the T7 promoter prevents *de novo* expression, by *E.coli* RNA polymerase, of the foreign DNA fragment in the host bacterium .

Cloning a DNA fragment that encodes both the RNA polymerase gene and promoter  $\phi 1.1A$  into *E.coli* (via an expression vector) can be to be lethal to the host bacterium (Davanloo et al. 1984). Expression vectors give stringent, but possibly not complete control over expression of the inserted foreign gene (Hasan and Szybalski 1987). Thus, a low level expression of the T7 RNA polymerase gene, if cloned into *E.coli* with promoter  $\phi 1.1A$ , could result in enough enzyme to initiate transcription from promoter  $\phi 1.1A$ . It may transcribe the entire recombinant vector and in doing so produce more RNA polymerase which would augment continuation of the process. An alternate possibility is that this recombinant vector is lethal by virtue of depleting the ribonucleotide pool during the unregulated synthesis of short transcripts close to promoter  $\phi 1.1A$ .

Hence, the high specificity of T7 RNA polymerase for the T7 promoter ensures that neither the T7 RNA polymerase gene or promoter  $\phi 1.1A$  alone are lethal to *E.coli*. However, when both the T7 RNA polymerase gene and promoter  $\phi 1.1A$  are cloned in to *E.coli* in tandem, it appears that they can act operationally as an RNA transcription engine and kill the host. Nevertheless, with adequate control of gene expression, the combination of the T7 RNA polymerase gene and a T7 promoter *in trans* can be exploited as a tool for expressing another foreign gene which is to be controlled by the T7 promoter (Tabor and Richardson 1985, Plotch et al. 1989, Guan et al. 1990). In addition, the cloning of this gene for the commercial scale production of T7 RNA polymerase is facilitated by isolating of a DNA fragment bearing the T7 RNA polymerase gene, but not bearing promoter  $\phi 1.1A$ .

Before a functional T7 RNA polymerase gene was first cloned by Davanloo et al. (1984), the entire nucleotide sequence of the T7 genome was already known (Dunn and Studier 1983), as was a corrected sequence for the T7 RNA polymerase gene (Moffatt et al. 1984). Computer analysis of the DNA sequence had indicated that no restriction sites existed that could separate the gene from promoter  $\phi 1.1A$  without also digesting the gene. They were able

to isolate the gene from a heteroduplex prepared between two double deletion strains of T7 phage. Essentially, their heteroduplex molecule included two non-complimentary regions of DNA which formed S1 nuclease-sensitive single strand loops, with the RNA polymerase gene lying between the loops. Purification of the nuclease-insensitive double strand DNAs by agarose gel electrophoresis enabled the isolation of a DNA fragment bearing the RNA polymerase gene but lacking promoter  $\phi 1.1A$  (see figure 3.2). The T7 RNA polymerase purified from this cloned gene is biochemically identical to the same enzyme purified from T7 phage-infected *E.coli*.



### 3.2 ISOLATING AND CLONING THE T7 RNA POLYMERASE GENE

#### 3.2.1 INTRODUCTION

A simple approach to removing promoter  $\phi 1.1A$  away from the gene is by isolating a restriction fragment that has the promoter near one terminus and then subjecting the fragment to limited digestion with an exonuclease, such as Bal 31 nuclease. This will generate a population of

DNA fragments varying in length and may include fragments that encode the RNA polymerase gene, but not promoter  $\phi 1.1A$ . In practice, this approach has at least one problem. The deletion of terminal nucleotides from a DNA fragment by an exonuclease rarely involves the simultaneous binding of the enzyme to both termini. The asymmetry of digestion may complicate the sizing of exonuclease-digested DNA by agarose gel electrophoresis. It cannot be assumed that a band of DNA seen on an agarose gel will include the DNA fragments that correspond to identically sized DNA fragments predicted from an equation relating fragment-length to digestion time. The time function is valid for describing the population behaviour of the DNA fragments that are being digested, but is not appropriate for describing the outcome for a single fragment of DNA. This is significant because it is single fragments of DNA which will be cloned into the bacteria. The greater the variation in exonuclease digestion products, the greater will be the number of transformants to be screened for a particular DNA clone. Consequently it is expedient to somehow arrange for the transformants on an agar plate to be more likely to contain a desired DNA fragment, rather than some other DNA fragment.

The use of exonuclease deletion is thus tempered by the stochastic nature of the technique. Success tends to depend on luck rather than being the most likely result of the experimental strategy.

The inherent randomness of exonuclease digestion can be exploited to advantage when searching for a DNA fragment that is not fully characterised. For example, even in the absence of a detailed genetic map of the phage T7 genome it would be possible to isolate a functional T7 RNA polymerase gene, lacking promoter  $\phi 1.1A$ . This could be achieved by digesting a variety of restriction endonuclease-treated T7 genomic DNA samples with Bal 31 nuclease over a time course, isolating appropriately sized DNA fragments from an agarose gel and cloning the fragments into *E.coli*. If one were prepared to assay the clones for T7 RNA polymerase activity there exists a finite chance of finding the gene in a 'clonable' form. When a functional assay for a gene-product exists, this approach is feasible, but tends to be very unpredictable. Most importantly, success using this approach could be extremely hard to reproduce.

### **3.2.2 Experimental strategy for isolating the T7 RNA polymerase gene**

The development of phage M13 site-directed oligonucleotide mutagenesis (Zoller and Smith 1982, 1983) offers a simple solution to the problem of generating a specific DNA fragment using exonuclease digestion. As site-directed mutagenesis allows cloned DNA fragments to be modified, the asymmetric nature of Bal 31 nuclease can be exploited to generate, from a larger precursor, a DNA fragment bearing the T7 RNA polymerase gene with a deletion of part of the open reading frame and/or a deletion of promoter  $\phi 1.1A$ . Nuclease Bal 31 could generate a range of variously sized DNA fragments which would be cloned into *E.coli*. Colonies of the bacteria that grow on selective media are likely to contain the cloning vector but are unlikely to contain both a functional RNA polymerase gene and promoter  $\phi 1.1A$ .

Sizing the DNA insert would allow DNA fragments that may encode slightly truncated RNA polymerase genes to be identified. This would be determined precisely by dideoxy sequencing the termini of the inserts. Site-directed oligonucleotide mutagenesis would then be carried out to reconstruct the gene. The optimal Bal 31 nuclease digestion product would be a DNA fragment encoding a full-length RNA polymerase gene but lacking promoter  $\phi 1.1A$ .

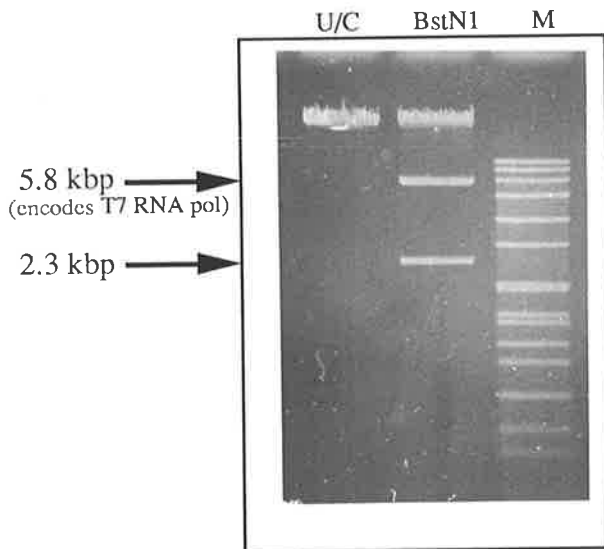
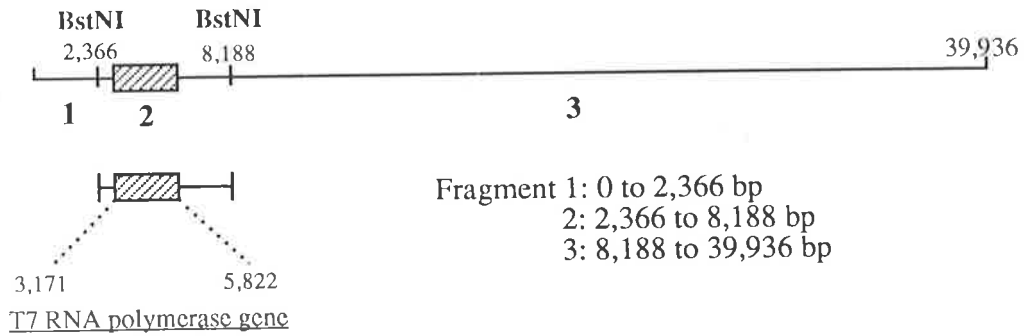
If the 3.6 kbp BstNI-DraI restriction fragment of T7 genomic DNA was digested with Bal 31 nuclease (figure 3.3), there could be several DNA species generated which may encode all or most of the T7 RNA polymerase gene in a form which is not lethal to *E.coli*. The minimum digestion of the 3.6 kbp BstNI-DraI T7 DNA fragment which would generate a DNA fragment that is not lethal when cloned into *E.coli* would involve a small deletion of either the 5' terminus of the gene or a deletion of promoter  $\phi 1.1A$ . Site-directed oligonucleotide mutagenesis could then be used to repair the deletion(s). In practice, this technique allows for deletions of 30-50 bp to be repaired using a synthetic oligonucleotide (Zoller and Smith 1983).

The digestion of T7 genomic DNA with Bal 31 nuclease to produce T7 RNA polymerase gene deletions, and the repair of the deletions by site-directed oligonucleotide mutagenesis, was considered a feasible approach to cloning and subsequently expressing the T7 RNA polymerase gene in *E.coli*. The results of this work are described in this chapter.



Figure 3.5

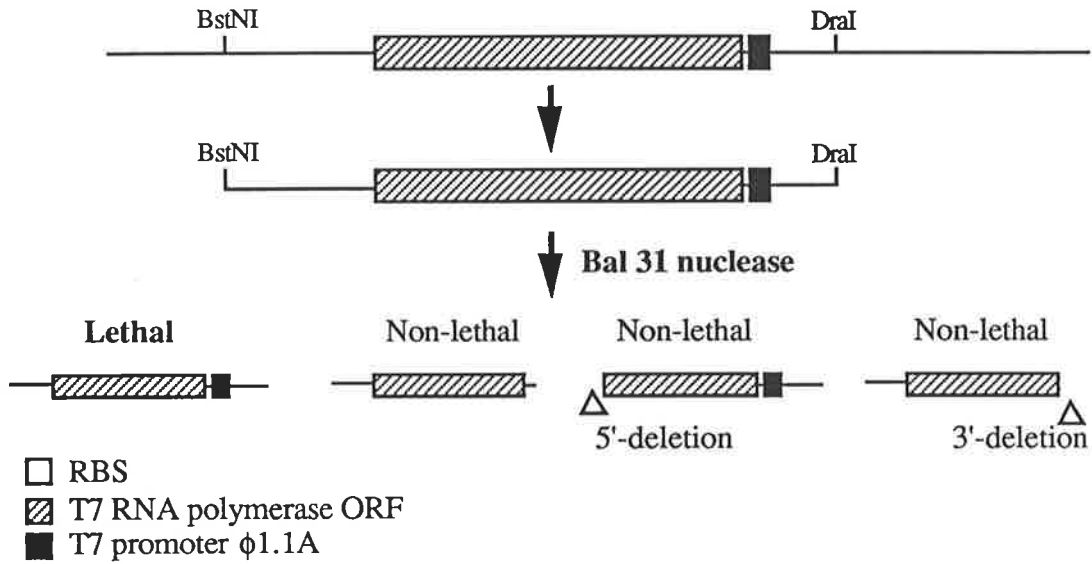
BstNI restriction map of T7 genome



U/C: uncut T7 DNA  
M: EcoR1-digested phage SPP-1 DNA

Restriction digest: 1 $\mu$ g T7 DNA digested with 10 units of BstN1 or 8 units of DraI in Low salt buffer, for 1 hour at 37°C.  
DNA electrophoresed through 1% agarose-TAE gel and stained in 10 $\mu$ g/ml ethidium bromide solution.

**Figure 3.3** Digestion of T7 genomic DNA to yield full or partial length RNA polymerase gene after Bal 31 nuclease digestion

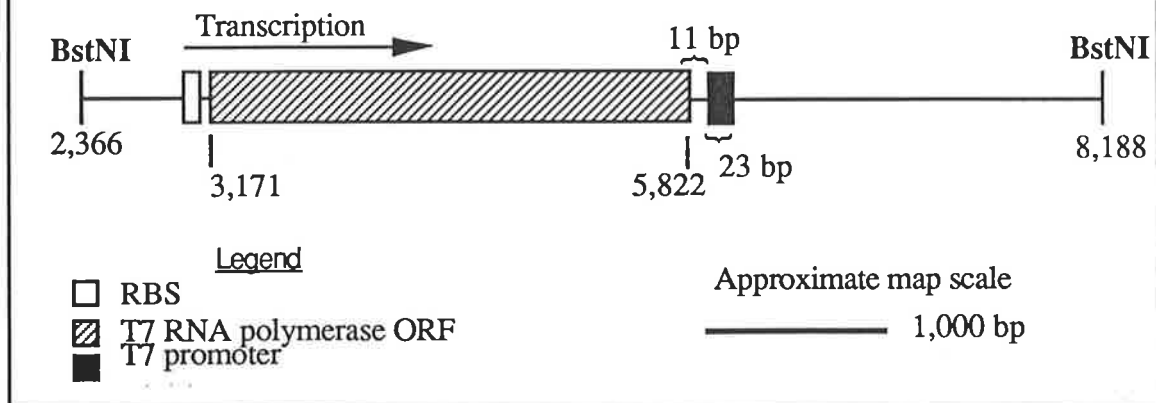


### 3.3 METHODS and RESULTS

#### 3.3.1 Preparation of starting material

The starting material for the following experiments was the 5,822 bp BstNI restriction fragment of T7 genomic DNA (see Fig 3.4). Digestion of T7 genomic DNA with BstNI yields

**Figure 3.4** Starting material for exp.3.3.2: 5,822 bp BstNI restriction fragment of T7 genomic DNA



three restriction fragments which are  $\approx 2.3$  kbp, 5.8 kbp and 31.7 kbp in length (see figure 3.5). The 5.8 kbp BstNI restriction fragment of T7 DNA encodes the T7 RNA polymerase

gene and promoter  $\phi$ 1.1A.

T7 DNA (100  $\mu$ g) was digested with 140 units of BstNI in low salt buffer in a reaction volume of 200  $\mu$ l for 2 hours at 37°C. The digested DNA was loaded onto a 20 x 20 cm 1% agarose-TAE submarine gel as two 170  $\mu$ l samples and electrophoresed for 1 hour at 350 mA. After staining with 10  $\mu$ g/ml ethidium bromide solution for 10 minutes, the two 5.8 kbp bands encoding the T7 RNA polymerase gene were removed, transferred to a dialysis bag and electroeluted until examination under UV light showed all dye had been eluted from the gel slices. The DNA was then phenol extracted and precipitated with ethanol. Following centrifugation, the DNA pellet was washed twice with 95 % ethanol, dried *in vacuo*, and dissolved in 170  $\mu$ l of TE buffer. To determine the DNA concentration, 10 $\mu$ l of the DNA solution and DNA markers (1 $\mu$ g of EcoRI-digested phage SPP-1 DNA) were electrophoresed through a 1% agarose-TAE submarine gel and stained with 10  $\mu$ g/ml ethidium bromide solution for 10 minutes. Comparison of the DNA fluorescence of the two DNA samples indicated that the 5.8 kbp BstNI T7 DNA fragment was at a concentration of approximately 100 ng/ $\mu$ l.

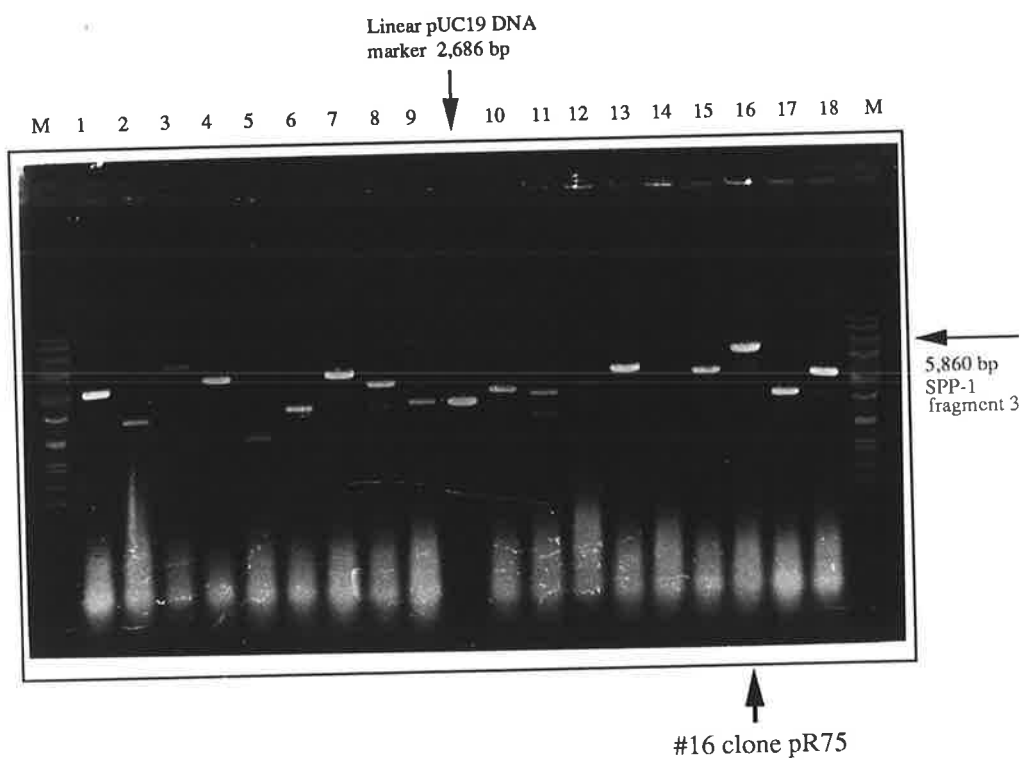
### 3.3.2 Bal 31 digestion to remove promoter $\phi$ 1.1A

The aim was to digest the 5.8 kbp BstNI T7 DNA restriction fragment with DraI, which cuts the DNA fragment 140 bp from promoter  $\phi$ 1.1A, followed by limited digestion with Bal 31nuclease (see figure 3.6).

Five micrograms of the 5.8 kbp BstNI fragment was dried *in vacuo*. The pellet was redissolved in 16  $\mu$ l of water and digested with 40 units of DraI, in low salt buffer, in a volume of 20  $\mu$ l, for 1 hour at 37°C. The reaction volume was then increased by adding 15  $\mu$ l of water and 4  $\mu$ l of 10 x Bal 31 buffer. The digestion was initiated on adding 1  $\mu$ l of 0.75 units/ $\mu$ l Bal 31 nuclease and terminated at 30,90,108, and 115 seconds by transferring 10  $\mu$ l of the reaction solution to an Eppendorf tube containing 10  $\mu$ l of 40 mM EGTA. Each sample of DNA was then purified using the Bio-101 GeneClean protocol and redissolved in 8  $\mu$ l of water. To this was added 1  $\mu$ l of 10 x End Fill buffer and 1  $\mu$ l of 1 unit/ $\mu$ l Klenow polymerase. The

Figure 3.7

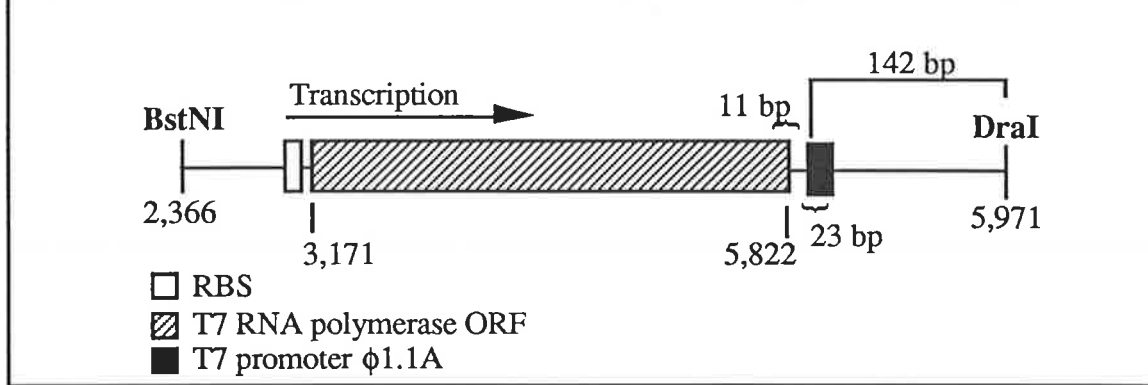
Candidate T7 RNA polymerase clones: Exp. 3.3.3



M: EcoR1-digested phage SPP-1 DNA

Plasmid DNA from the 24 clones described in Exp. 3.3.3 was isolated and 2 $\mu$ g of each plasmid was digested with 10 units of HindIII for 90 min at 37°C. The linear plasmids were electrophoresed through 1% agarose-TAE gels and stained with 10 $\mu$ g/ml ethidium bromide.

**Figure 3.6** 5.7 kbp BstNI-DraI T7 DNA fragment: starting material for exp. 3.3.2



solutions were incubated at 37°C for 15 minutes prior to adding 1 µl of 1mM dNTPs and incubating a further 30 minutes at 37°C. The blunt-end, Bal 31 nuclease digested DNA was purified on a 1% agarose-TAE submarine gel. The DNA was extracted from the agarose using the GeneClean protocol and redissolved in 20 µl of water.

### 3.3.3 Cloning the Bal 31 nuclease-digested DNA fragments

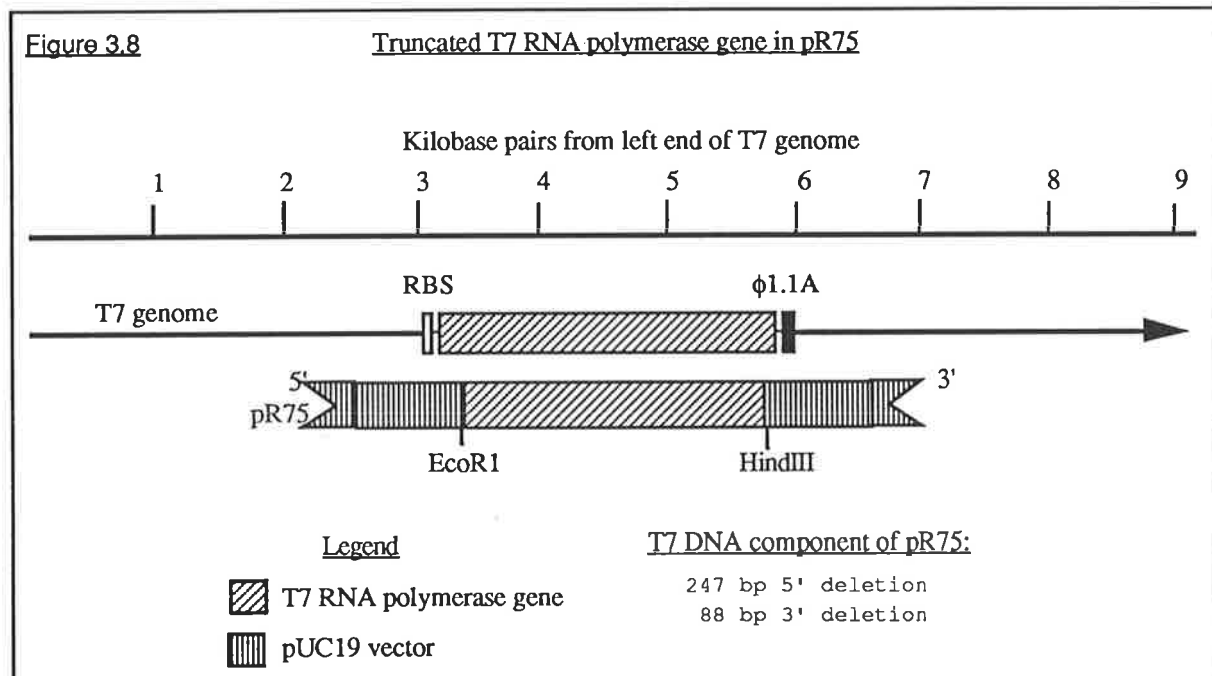
Each DNA sample (3 µl) was ligated with 20 ng of Sma 1-cut, dephosphorylated pUC19 DNA, in a 10 µl solution containing 50 mM Tris-Cl, pH 7.6, 5 mM MgCl<sub>2</sub>, 0.5 mM ATP, 0.5 mM DTT, 5 % PEG 8000 and 1 unit of T4 DNA ligase. The ligation reaction was done at 25°C overnight and then used to transform 200 µl of competent *E.coli* JM101. The transformed bacteria were grown overnight at 37°C on L-plates containing 75 µg/ml ampicillin. Colonies of bacteria were only found on the plates corresponding to the 30 second (6 colonies) and 90 second (18 colonies) Bal 31 nuclease digestions. Plasmids from the 24 bacterial colonies were isolated and digested with HindIII in order to estimate the size of the cloned T7 DNA fragments (see figure 3.7).

The best candidate plasmid for harbouring a full-length T7 RNA polymerase gene corresponded to the DNA from plasmid mini prep #16, designated pR75, as this plasmid has a linear size of about 5 kbp. As the pUC19 vector is 2,686 bp, the T7 DNA component of pR75 is about 2.3 kbp.

### 3.3.4 DNA sequencing the T7 DNA component of pR75

A sample of pR75 DNA (2 µg) was digested with EcoR1 and HindIII, which excised the ≈ 2.3

kbp T7 DNA component from the multiple cloning site polylinker of pR75. The 2.3 kbp T7 DNA was purified by agarose gel electrophoresis and then subcloned into EcoR1/HindIII-cut M13mp18 and M13mp19 to determine the terminal nucleotide sequences and hence the exact size of the cloned DNA fragment. Sequence analysis showed that the T7 DNA component of pR75 maps to T7 genome coordinates 3,417 to 5,735 bp, which is entirely within the T7 RNA polymerase open reading frame (see figure 3.8). This 2.3 kbp insert was used as a DNA probe for the next experiment.



### 3.3.5 Shotgun cloning of Bal 31 nuclease-digested T7 DNA fragments

The aim of experiment 3.3.2 was to digest the 5.8 kbp BstNI restriction fragment of T7 genomic DNA with Bal 31 nuclease and search for digested DNA fragments which can be cloned into *E.coli* and which are also large enough ( $\cong$  2.7 kbp) to encode the T7 RNA polymerase gene or this gene having a small ( $\cong$  30 bp) deletion.

In this modification of experiment 3.3.2 the aim was to:

1. Digest T7 genomic DNA using BstN1 and Dra1.
2. Subject the resultant DNA fragments to Bal 31 nuclease.
3. Separate the fragments by agarose gel electrophoresis.
4. Clone the appropriate size fragments into *E.coli*.

The transformed bacteria would then be probed for the RNA polymerase gene using the 2.3 kbp T7 DNA insert from pR75. This approach minimizes manipulation of the DNA prior to cloning.

The T7 DNA (10 µg) was digested with 20 units of BstN1 and 40 units of Dra1 in low salt buffer in a volume of 20 µl for 2 hours at 37°C. Following this, the reaction volume was increased with 15 µl of water, 4 µl of 10 x Bal 31 buffer, and 1 µl of 0.75 units/µl Bal 31 nuclease. The reaction proceeded for 90 seconds at 37°C after which 25 µl of 200 mM EGTA was added to stop the reaction. The DNA was purified using Bio-101 GeneClean: To the 65 µl of DNA solution was added 163 µl of GeneClean NaI solution, and the DNA/NaI solution split into 2 x 114 µl aliquots in Eppendorf tubes. Each tube then received 5 µl of the GeneClean glassmilk solution and the GeneClean protocol was followed, yielding 2 x 20 µl of aqueous DNA solution. The combined 40 µl solution was dried with ethanol *in vacuo* and the DNA pellet redissolved in 10 µl of water. Next the DNA was end-filled to ensure flush termini prior to the blunt ligation into the cloning vector: To the 10 µl DNA solution was added 5 µl of water, 2 µl of 10 x End Fill buffer and 2 µl of 1 unit/µl Klenow polymerase. The reaction proceeded for 15 minutes at 37°C followed by the addition of 1 µl of 12.5 mM dNTPs and a further 20 minutes at 37°C. The DNA fragments were then separated by electrophoresis through a 1% agarose-TAE submarine gel. After staining with 10 µg/ml ethidium bromide solution for 10 minutes, the gel was examined under UV light and a gel slice containing DNA fragments large enough to bear the full-length T7 RNA polymerase gene was excised. These DNA fragments correspond in size to DNA markers in the range 2,600 to 3,400 bp. DNA was purified from the agarose using the GeneClean protocol and was redissolved in 20 µl of water.

### **3.3.6 Cloning the 2,600 to 3,400 bp T7 DNA fragments into *E.coli***

The T7 DNA fragments (10 µl) were ligated to 100 ng of Sma1-cut, dephosphorylated pUC19 in a 25 µl solution containing 50 mM Tris-Cl, pH 7.6, 10 mM MgCl<sub>2</sub>, 10 mM DTT, 5 % PEG 8000 and 7 units of T4 DNA ligase. The reaction proceeded overnight at 25°C. The ligation mixture was used to transform 1 ml of competent *E.coli* MC1061. The transformed bacteria were grown overnight at 37°C on L-plates containing 75 µg/µl ampicillin, yielding about 600

colonies.

### **3.3.7 Identification of bacteria harbouring the T7 RNA polymerase gene**

This process involved fixing the bacteria to nitrocellulose and probing for the T7 RNA polymerase gene with a 2.3 kbp EcoR1-HindIII DNA insert from pR75. The 2.3 kbp DNA fragment encodes a partial length T7 RNA polymerase gene and was radiolabelled by nick translation.

#### **3.3.7.1 Preparation of probe for the T7 RNA polymerase gene**

Plasmid pR75 DNA was digested with EcoR1 and HindIII and the T7 insert (5 µg) was purified by agarose gel electrophoresis then radiolabelled by nick translation with 50 µCi of α-<sup>32</sup>P-dCTP (16 pmoles) with 80 % incorporation of the radiolabel, giving a specific activity of  $1.7 \times 10^7$  cpm/µg probe DNA. The nick translation reaction also included 50 mM Tris-Cl, pH 7.6, 10 mM MgSO<sub>4</sub>, 0.1 mM DTT, 25 mM each dATP, dGTP and dTTP, 1 µM dCTP and 2 µg of BSA in a volume of 30 µl. The reaction was initiated on adding 10 µl of a solution containing 10 units of *E.coli* DNA polymerase I and 16 units of DNase I, and proceeded for 90 minutes at 16°C. The probe was purified using Bio-101 GeneClean and redissolved in 500 µl of 10 mM Tris-Cl, pH 7.6, containing 0.1 mM EDTA and 5 mM β-mercaptoethanol.

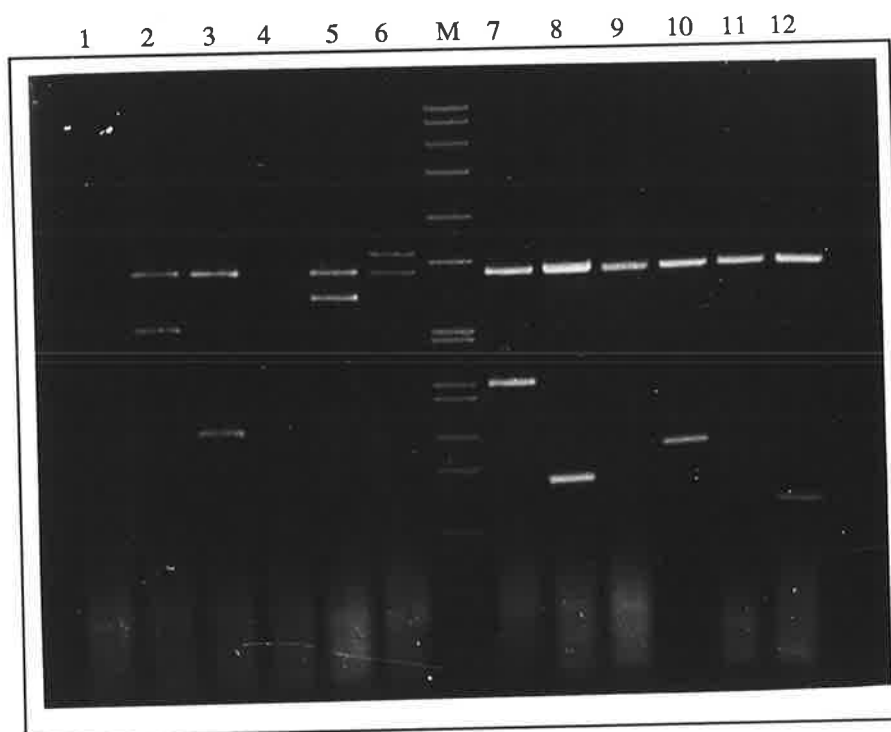
#### **3.3.7.2 Preparation of DNA-bound nitrocellulose filters and hybridization with the 2.3 kbp partial length T7 RNA polymerase gene probe**

Two hundred transformants were subcultured onto sterile nitrocellulose filters overlaid onto L-plates containing 75 µg/ml ampicillin and grown overnight at 37°C. The colonies of bacteria that grew on the nitrocellulose were treated by a modification of the technique described by Grunstein and Hogness (1975). The nitrocellulose filter was layered, colonies up, onto squares of Whatmann 3MM paper soaked in 0.5 M NaOH for 7 minutes. The filter was then similarly treated by washing twice with 1 M Tris-Cl, pH 7.6, for 2 minutes followed by a 4 minute



Figure 3.9

Candidate T7 RNA polymerase clones: Exp.3.3.6



↑  
#6 clone pR80

M: EcoR1-digested phage SPP-1 DNA

Plasmid DNA from the 12 clones generated in Exp. 3.3.6 was isolated, digested with 10 units each of EcoR1 and HindIII for 90 min at 37°C, and then electrophoresed through a 1% agarose-TAE gel. The DNA was stained with 10 µg/ml ethidium bromide.

wash with a solution containing 0.5 M Tris-Cl, pH 7.6 and 1.5 M NaCl. The filter was then dipped in chloroform and then into ethanol, dried at room temperature, and baked at 80°C for 2 hours *in vacuo*. The filter was then washed in 20 ml of hybridization solution (6 x SSC, 5 x Denhardtts, 0.1 % SDS, 0.2 mg/ml salmon sperm DNA) for 4 hours at 68°C. Following this, the radiolabelled probe was denatured by boiling for three minutes and added to the hybridization solution. Hybridization continued overnight at 68°C. The unbound probe was removed by washing the filter in high, then low salt conditions, at room temperature: First the filters were washed three times in a solution containing 2 x SSC and 0.1 % SDS. The first wash was for 10 minutes, the latter washes were 30 minutes. The final procedure involved washing twice in a solution containing 0.5 x SSC and 0.1 % SDS. The filters were then air dried and autoradiographed. Twelve clones were found which strongly bound the pR75-derived radiolabelled probe.

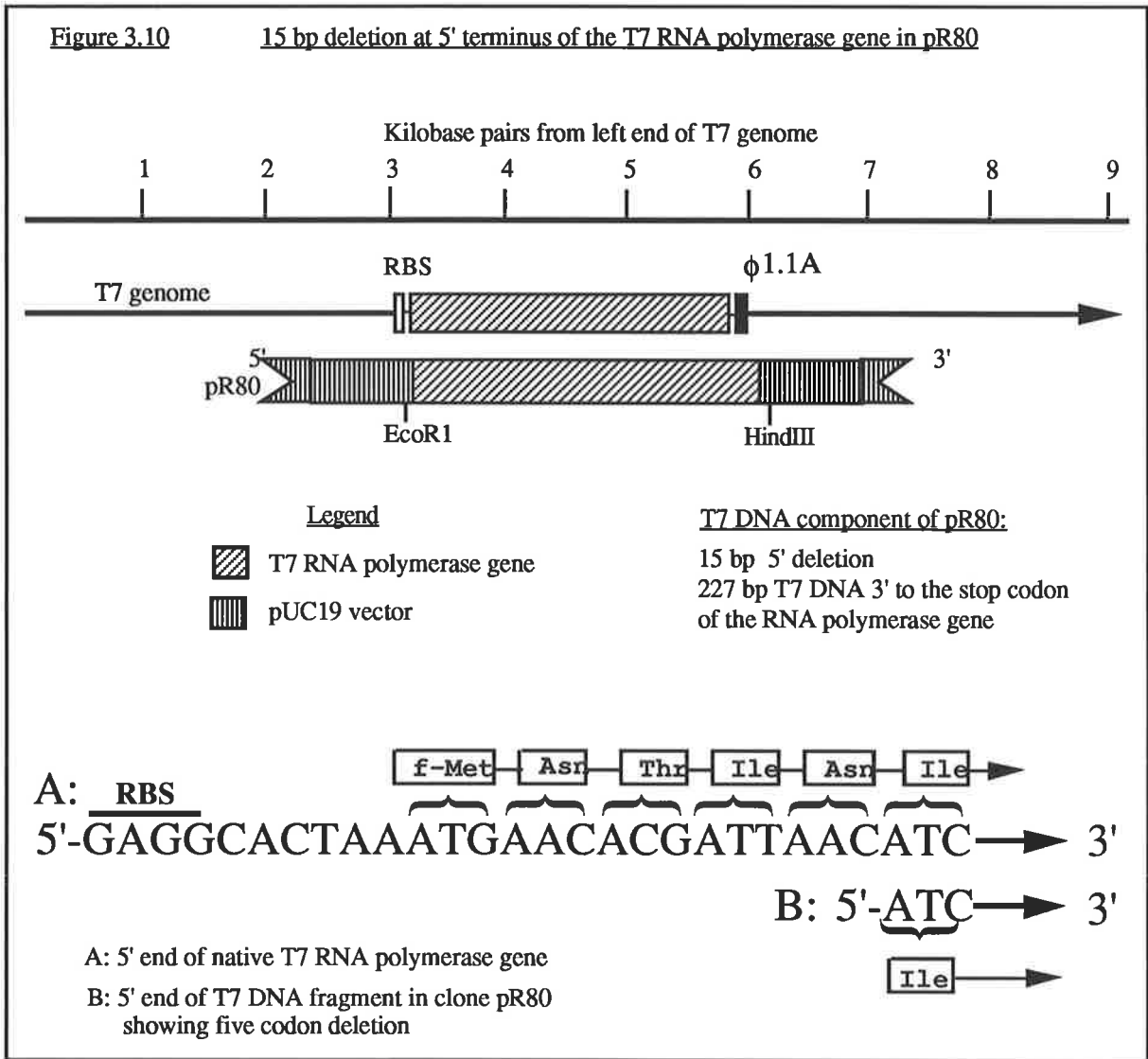
Plasmids from the twelve positive clones were isolated and then digested with EcoR1 and HindIII in order to determine the size of the T7 DNA insert. The DNA samples were electrophoresed in a 1% agarose-TAE submarine gel and stained in 10 µg/ml ethidium bromide for 10 minutes (see figure 3.9).

Clone #6 was the only transformant bearing a cloned T7 DNA fragment large enough to encode a full length T7 RNA polymerase gene. The recombinant plasmid from clone #6 was designated pR80. The size of the T7 DNA component of pR80 was then determined precisely by dideoxy sequencing the termini of the DNA.

### **3.3.8 DNA sequencing of the T7 DNA component of pR80**

Ten micrograms of pR80 was digested with EcoR1 and HindIII and the T7 DNA fragment subcloned into EcoR1/HindIII-cut M13mp18 and M13mp19. The M13 clones were called M13/6-18 and M13/6-19. Sequence analysis showed that the T7 DNA component of pR80 maps to T7 genome coordinates 3,186 to 6,049. This DNA fragment includes promoter  $\phi$ 1.1A but lacks the first 15 nucleotides of the open reading frame (see figure 3.10).

**Figure 3.10** 15 bp deletion at 5' terminus of the T7 RNA polymerase gene in pR80



As the T7 DNA component of pR80 is just 15 bp short of full length, it was possible to exploit oligonucleotide site-directed mutagenesis to engineer the T7 DNA into a full-length T7 RNA polymerase gene but lacking promoter  $\phi$ 1.1A.

### 3.3.9 Engineering the T7 DNA component of pR80

#### 3.3.9.1 Introduction

The template for the first round of mutagenesis was the single stranded DNA from M13/6-18. The repair of the partial length T7 RNA polymerase gene in pR80 to a full length gene was done in two stages. First was the removal of promoter  $\phi$ 1.1A, followed by insertion of the missing 5'-nucleotides. An overview of the reconstruction is shown diagrammatically in

**Figure 3.11** Engineering the T7 DNA component of pR80 into a full-length T7 RNA polymerase gene, with optimised RBS and lacking T7 promoter  $\phi$  1.1A

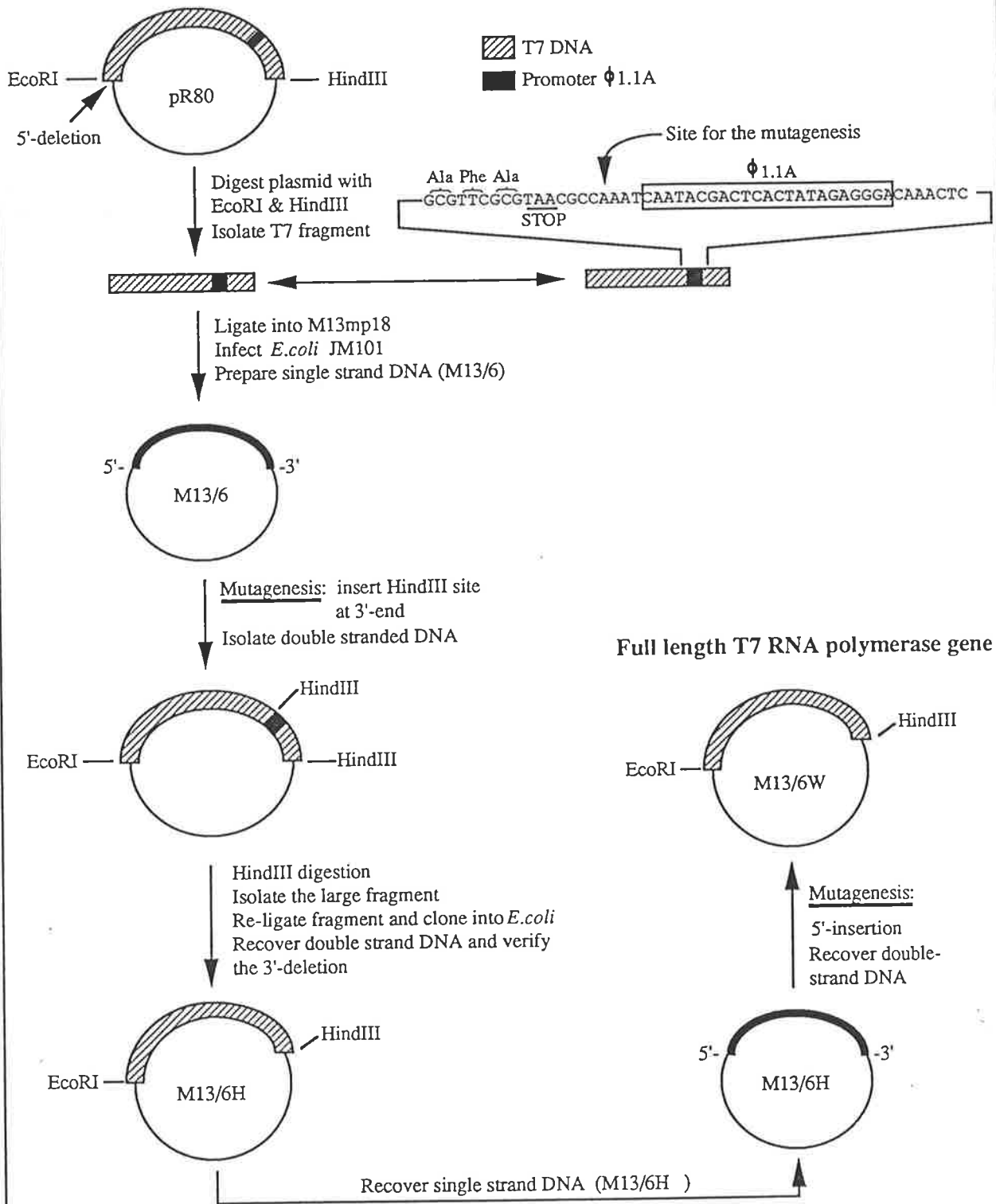


figure 3.11.

The approach taken to remove promoter  $\phi$ 1.1A away from the T7 RNA polymerase gene involved the site-directed mutagenesis of T7 DNA between the gene and promoter  $\phi$ 1.1A, from nucleotides 5,829 to 5,832. This mutagenesis transformed the DNA sequence, starting at nucleotide 5,827 (in bold), from 5'-AAATCA-3' to 5'-AAGCTT-3'; the latter DNA sequence is the recognition site for the endonuclease HindIII. Digestion of this construct with HindIII allowed removal of the promoter and maintained a HindIII site 3'- to the polymerase gene, maintaining portability of the polymerase gene as an EcoRI-HindIII cassette. This was followed by insertion of the missing 5'-nucleotides and also appropriate 5'-untranslated sequences by site-directed mutagenesis.

The expression vector into which the RNA polymerase gene was ligated is pPLC236, devised by Remaut et al. (1981) in which the  $P_L$  promoter from bacteriophage lambda is placed upstream of unique EcoRI and HindIII restriction sites. For a gene to be expressed in *E.coli* the coding sequence must be preceded by a consensus *E.coli* ribosome binding site (RBS), that is, a sequence which has some homology to the 3'-terminus of *E.coli* 16S ribosomal RNA as described by Shine and Dalgarno (1974). Vector pPLC236 does not supply initiation signals, so an RBS and 5'-ATG codon must be provided by the foreign DNA fragment.

### **Designing an optimised RBS for the T7 RNA polymerase gene**

There is considerable degeneracy in the RBS sequence recognised by *E.coli* ribosomes. Scherer et al.(1980) determined a consensus RBS sequence by comparing 68 RBS sequences from *E.coli* and found that the RBS "...shows a preference rather than an absolute requirement for a specific base in any given position". The relationship of the sequence described by Shine and Dalgarno (1974) and the consensus RBS (Scherer et al. 1980) is shown in figure 3.12.

**Figure 3.12**

**Comparative RBS sequences**

3'-terminus of *E.coli*  
16S rRNA

3'-**AUUCCUCCACUAG**••••

Shine-Delgarno

5'-UAAGGAGGUGAUC-3'

Consensus  
RBS

5'-UUAAAUUAAGGAGGUAUAUUAUG••••  
f-met

A review by Gren (1984) showing the DNA sequences upstream from 273 genes from *E.coli* and several bacteriophage highlights the variation of RBS sequences within a genome. For example, the RBS sequence 5'-GGAGG-3' is found in more than 10 % of the 273 sequences compiled by Gren. However this sequence is invariably just one of many RBS sequences found in a given genome.

Prior to cloning the engineered RNA polymerase gene a useful expression vector (pEV601) became available (Brumby 1986). This vector, derived from pKT52 (Talmadge and Gilbert 1980) has a synthetic RBS and spacer having the sequence 5'-GGAGGTAACATATG-3'. The ATG is the initiation codon. The sequence 5'-CATATG-3' is the recognition site for the endonuclease NdeI. Having the sequence 5'-GGAGGTAACAT-3' as the synthetic RBS-spacer in the engineered RNA polymerase gene allows the gene to be cloned into pEV601 via the NdeI-HindIII restriction sites and into pPLC236 via the EcoRI-HindIII restriction sites (see figure 3.13).

**Figure 3.13**

**Cloning vectors for the reconstructed T7 RNA polymerase gene**

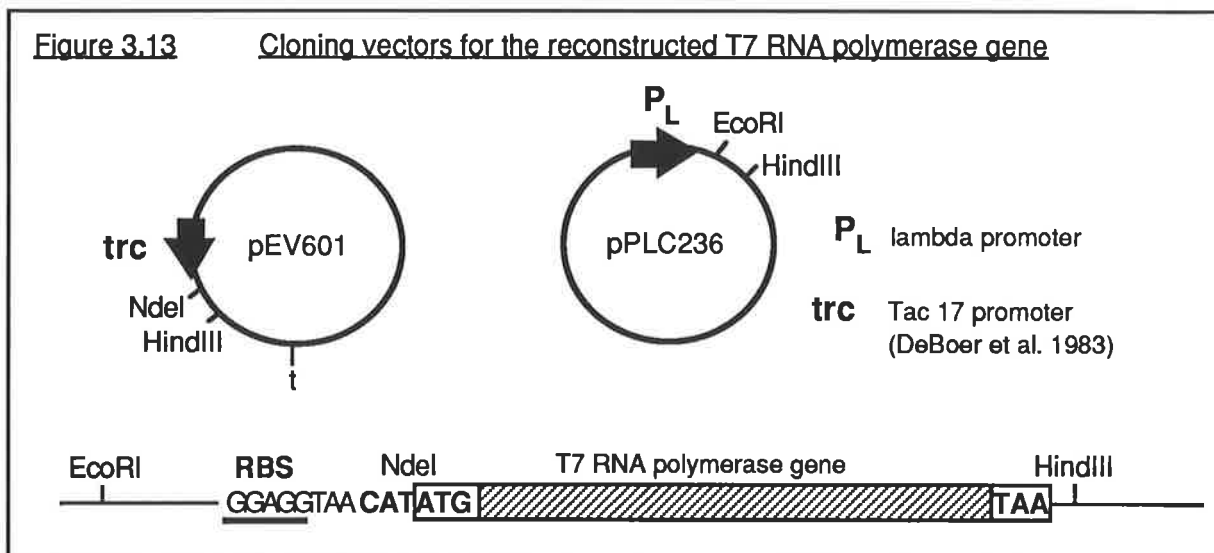


Table 3.1 RBS and spacer sequences ahead of the initiation codons from T7 phage

T7 gene	Sequence 5' to initiation codon	(from Dunn and Studier 1983)
1	<b>GAGGCACTAAATG</b>	.....RNA polymerase
2	<b>GAGGTCAATGACTATG</b>	.....Inactivates host RNA polymerase
3	<b>AGGAGGCGACATG</b>	.....Endonuclease
4	<b>AGGAGGGAATTGCATG</b>	.....Primase
5	<b>AGGAGAAATCAATATG</b>	.....DNA polymerase
6	<b>AGGATGTGTCTAATG</b>	..... Exonuclease
7	<b>GGAGGTAAGAAGTGATG</b>	.....Host range
8	<b>GGAGACACATTTAATG</b>	.....Head-tail protein
9	<b>TAAGGAGACAATAATAATG</b>	...Head assembly protein
10	<b>AAGGAGATATACATATG</b>	.....Head protein
11	<b>AGGAGGAACTATATG</b>	.....Tail protein
12	<b>AGGAGGCTCTAATG</b>	.....Tail protein
13	<b>GGAGGGTTTTCTTATG</b>	.....Internal virion protein
14	<b>AGGAGGATAACCATATG</b>	.....Internal virion protein
15	<b>GGAGGTAATGAGCTATG</b>	.....Internal virion protein
16	<b>AGGAGGCCCTAAATG</b>	.....Internal virion protein
17	<b>TAAGGAGGTCAAATG</b>	.....Internal virion protein
18	<b>AGGAGTATATGTATG</b>	
19	<b>AGGAGGCAATG</b>	

**TAAGGAGGTGATC** (Shine-dalgarno 1974)

14 **AGGAGGATAACCATATG**  
**GGAGGTAACATATG** (RBS and spacer in the  
reconstructed T7 RNA polymerase)

RBS sequences ( Shine-Dalgarno 1974) and the initiation sequences are shown in bold.





at a concentration of 111 ng/μl. The oligonucleotide (100 ng) was then 5'-phosphorylated by incubation in a 10 μl solution containing 50 mM Tris-Cl, pH 7.6, 10 mM MgCl<sub>2</sub>, 1mM ATP, 1 mM DTT and 3 units of T4 polynucleotide kinase. The reaction proceeded for 1 hour at 37°C followed by 5 minutes at 70°C to inactivate the kinase. The volume was increased to 25 μl with water, bringing the 5'-phosphorylated 30-mer oligonucleotide concentration to 4 ng/μl.

#### **3.3.9.2.1 Mutagenesis Reaction**

The mutagenesis reaction involved annealing the SR-463 oligonucleotide to single strand M13/6 DNA followed by primer extension and transformation of *E.coli*. The 13 μl annealing reaction contained 4 ng of phosphorylated SR-463, 4 ng of universal sequencing primer, 1 μl of M13/6 DNA (from a 25 μl DNA prep), 5 μl of 500 mM Tris-Cl, pH 7.6/100 mM MgCl<sub>2</sub> and 5 μl of 200 mM NaCl. The solution was heated to 70°C, then 30°C for five minutes. The volume was adjusted to 50 μl with the addition of 5 μl of 10 mM ATP, 5 μl of 10 mM DTT, 5 μl of 0.5 mM dNTPs, 1 μl of 2 units/μl T4 DNA ligase, 2 μl of 3 units/μl Klenow polymerase and 19 μl of water. The reaction proceeded for 4 hours at 25°C. 0.5 μl of this solution was then used to transform 200 μl of competent *E.coli* JM101. The transformed bacteria were then mixed with 3 ml YT-sloppy agar, poured onto minimal A plates and incubated overnight at 37°C.

#### **3.3.9.3 Detection of mutant M13/6 DNA**

The resultant phage were lifted onto a sterile nitrocellulose filter and baked for 2 hours at 80°C *in vacuo*. The filter was then incubated for 2 hours at 42°C in a hybridization solution containing 90 mM Tris-Cl, pH 7.6, 900 mM NaCl, 6 mM EDTA, 5 x Denhardts, 100 μg/ml salmon sperm DNA and 0.5% NP40 detergent.

The probe for mutant detection was prepared by incubating 111 ng of the 30-mer oligonucleotide primer (SR-463) with 50 μCi (10 pmoles) of γ-<sup>32</sup>P-ATP, 50 mM Tris-Cl, pH 7.6, 10 mM MgCl<sub>2</sub>, 1 mM DTT and 3.5 units polynucleotide kinase in a volume of 10 μl. The reaction proceeded for 1 hour at 37°C, then 5 minutes at 70°C.

The radiolabelled 30-mer probe was added to the hybridization solution and incubated

Figure 3.15 Insertion of HindIII restriction site between the partial length T7 RNA polymerase gene and T7 promoter  $\phi$ 1.1A

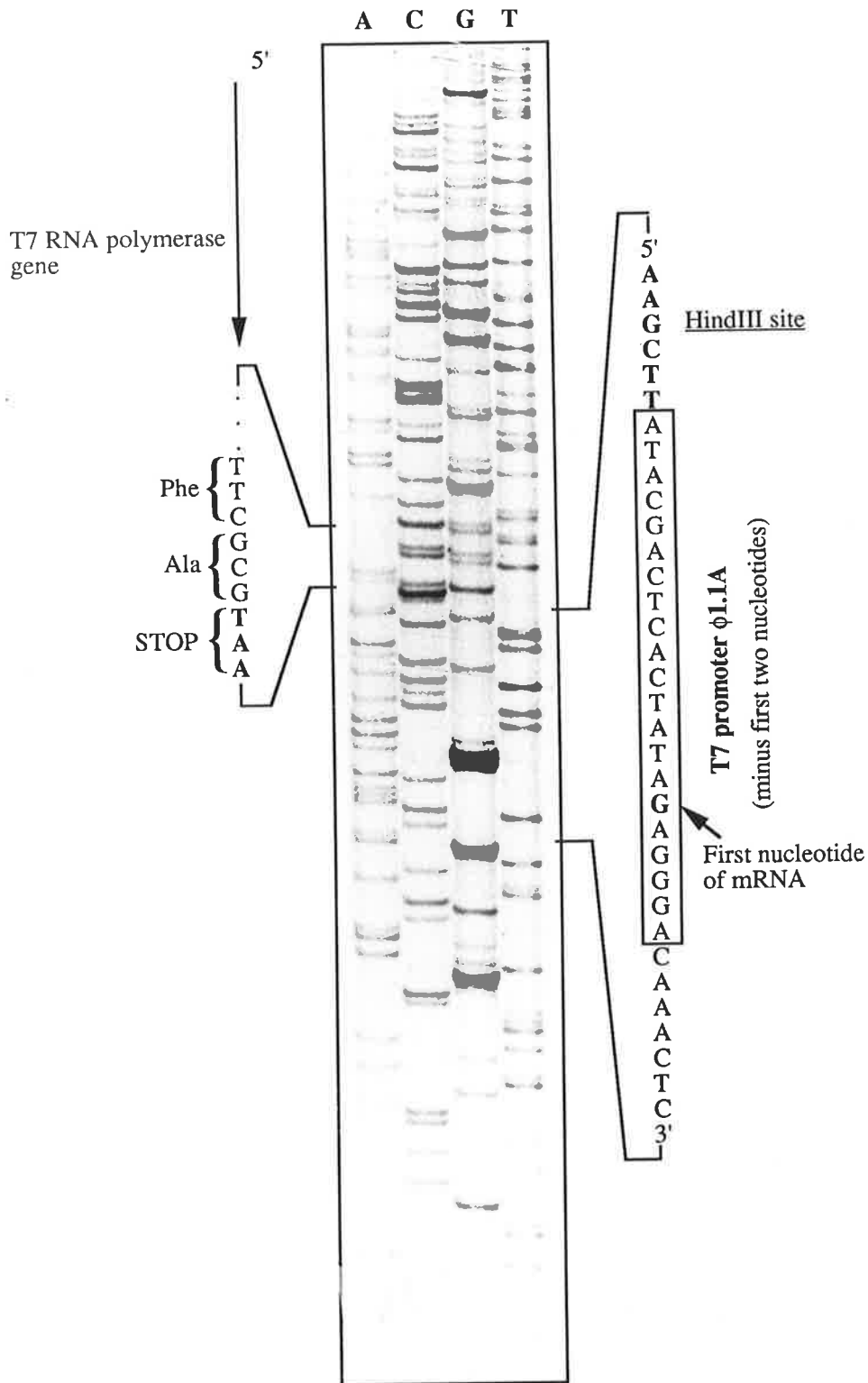
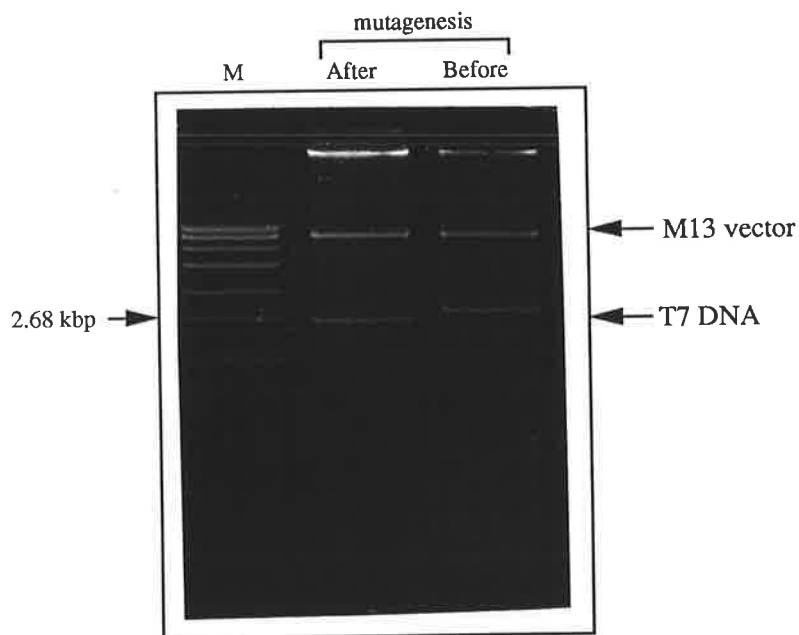
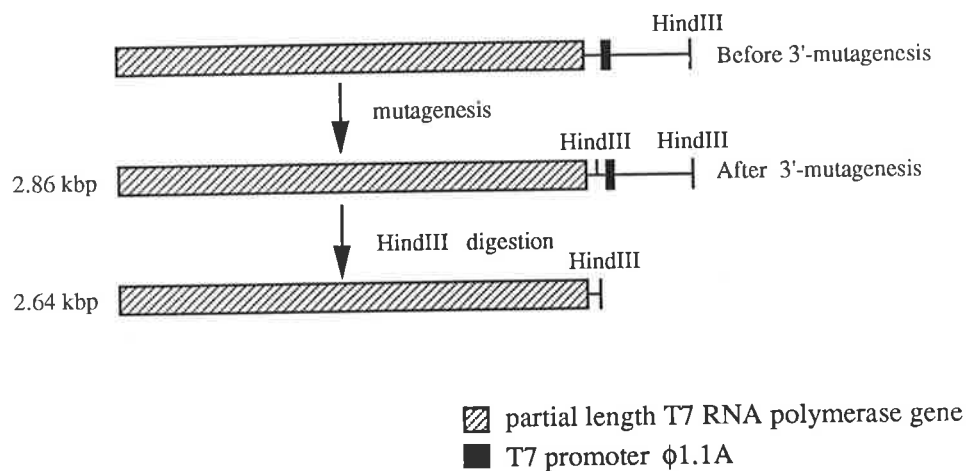


Figure 3.16 Removal of T7 promoter  $\phi$ 1.1A from the engineered, partial length T7 RNA polymerase gene by digestion with HindIII



M: EcoR1-digested phage SPP-1 DNA

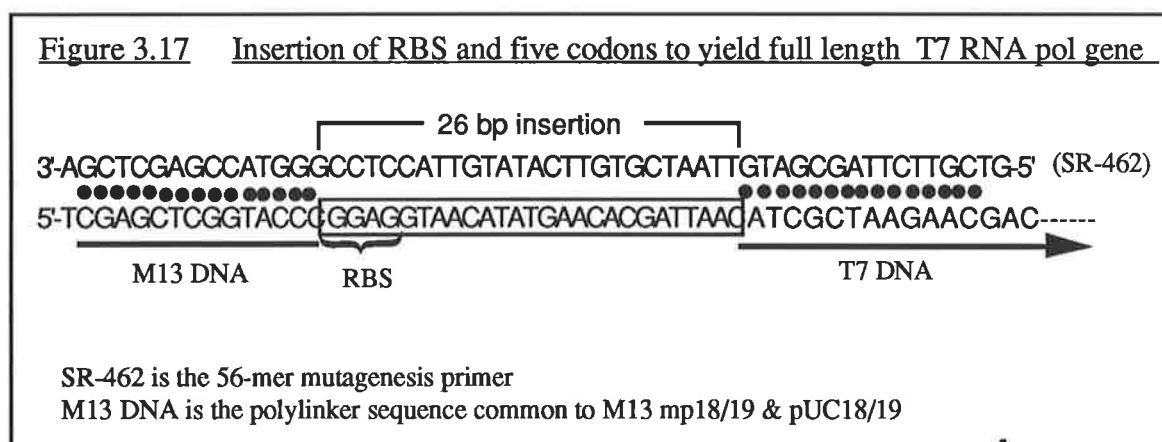
1  $\mu$ g of the RF M13 DNA from phage bearing mutated or non-mutated T7 DNA was digested with 14 units of EcoR1 and 5 units of HindIII for 90 min at 37<sup>o</sup> C. The DNA was then electrophoresed through a 1% agarose-TAE gel and stained in 10  $\mu$ g/ml ethidium bromide.

overnight at 42°C. The filter was then washed twice at room temperature in 6 x SSC prior to 2 x 30 minute washes at 69°C in TMAC solution (3 M tetramethyl ammonium chloride, 50 mM Tris-Cl, pH 8.0, 2 mM EDTA and 0.1 % SDS). The dissociation temperature for a 30-mer oligonucleotide, in the presence of TMAC, is about 74°C (Wood et al.1985). The filter was autoradiographed overnight at -80°C using a tungsten screen.

Mutant and wild type phage were identified and DNA from 12 mutant phage was sequenced to verify the mutation, as shown by figure 3.15. Double stranded DNA was subsequently isolated from a mutant phage, digested with EcoR1 and HindIII, and the partial length T7 RNA polymerase gene was subcloned into EcoR1/HindIII-cut M13mp18. This construct was designated M13/6H (see figure 3.16).

### 3.3.9.4 Repair of the 5'-deletion in M13/6H

Restoration of this construct to a full length open reading frame required insertion of five codons. Also needed was a ribosome binding site and Nde1 restriction site to exploit two highly efficient expression vectors as described (see figure 3.17).



Single stranded DNA of M13/6H was isolated and subjected to the the 5'-insertion mutagenesis under essentially the same conditions as previously described for the HindIII insertion mutagenesis. The 56-mer mutagenesis primer was designated SR-462 and has the sequence (5' → 3' reading left to right):

dGTCGTTCTTAGCGATCTTAATCGTGTTCATATGTTACCTCCGGGTACCGAGCTCGA

The dissociation temperature for a 56-mer oligonucleotide in the presence of TMAC is 85°C

Figure 3.18

Insertion of RBS, 6 nt spacer and five codons to restore the T7 DNA in M13/6H to a full-length T7 RNA polymerase gene

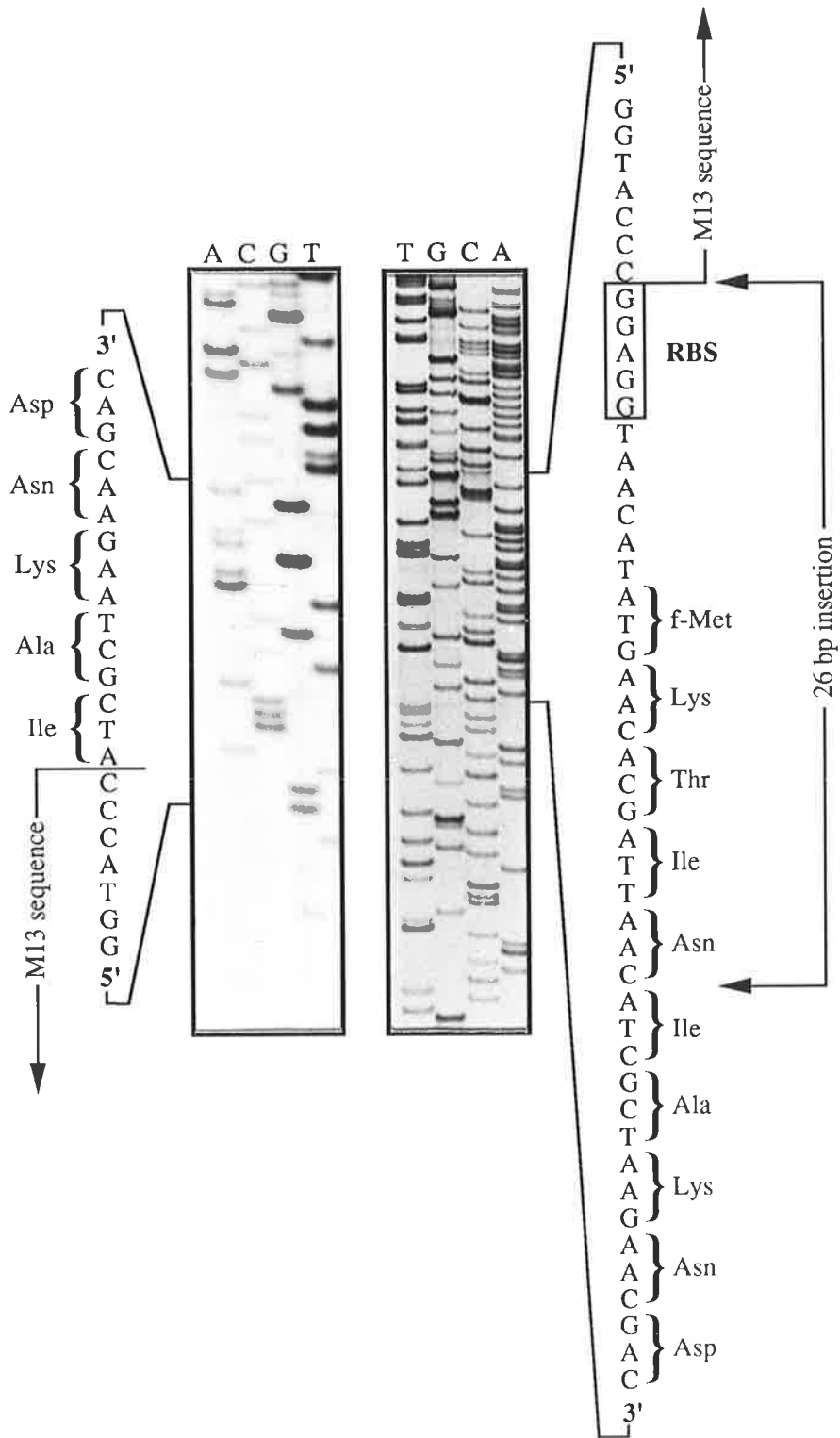
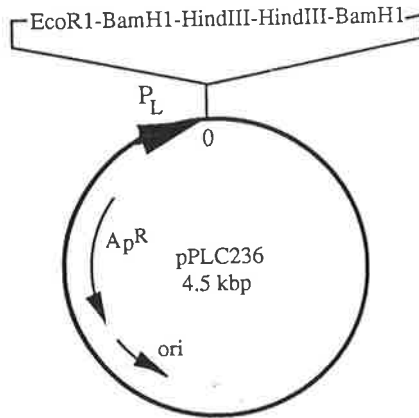
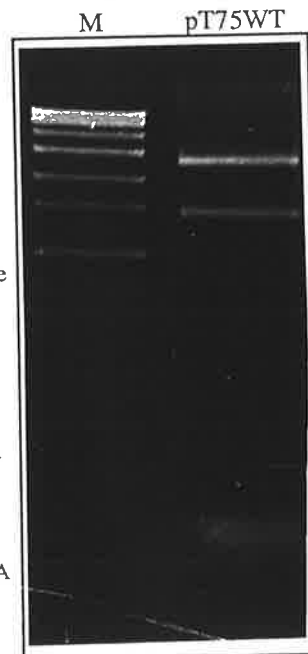


Figure 3.19

Expression vector pPLC236 and verification of transformation of *E. coli* E832 with pPLC236 harboring the reconstructed T7 RNA polymerase gene.



Ap<sup>R</sup>: β-lactamase gene  
ori: Origin of replication



← vector pPLC236 (4.5 kbp)

← T7 RNA polymerase gene (2.7 kbp)

pT75WT is pPLC236 harboring the reconstructed T7 RNA polymerase gene

5 μg of pT75WT was digested with 14 units each of EcoR1 and HindIII for 60 minutes at 37°C and then electrophoresed through a 1% agarose-TAE gel and stained with 10 μg/ml ethidium bromide.

M: EcoR1-digested phage SPP-1 DNA

(Wood et al. 1985). The probed filter was washed in TMAC solution at 80°C. The mutation was verified by dideoxy DNA sequencing (see figure 3.18). The M13mp18 clone harbouring the reconstructed full length T7 RNA polymerase gene was designated M13/6W.

### **3.3.10 Expression of T7 RNA polymerase in *E.coli* E832**

Double stranded M13/6W DNA was isolated, digested with EcoR1 and HindIII and the T7 DNA ligated into the expression vector pPLC236 (see figure 3.19). This construct, designated pT75WT, was used to transform *E.coli* E832, which expresses cI857<sub>ts</sub> protein. The transformed, competent *E.coli* E832 were grown for 20 hours on L-plates containing 75 µg/ml ampicillin. The incubation temperature was 30°C to avoid release of cI857 control over promoter P<sub>L</sub>; however no transformants were found, probably because of poor transformation efficiency. The pT75WT DNA was subsequently used to transform *E.coli* JM101. After identification of clones bearing pT75WT, the supercoiled pT75WT DNA was isolated from the *E.coli* JM101 and successfully used to transform *E.coli* E832.

The recombinant *E.coli* E832 bearing pT75WT (designated RS101) was grown at 30°C in L-broth containing 100 µg/ml ampicillin until the A<sub>600</sub> reached 1.0. The 5 ml culture was then transferred to a 42°C water bath for another 2 hour incubation to induce expression of the T7 RNA polymerase.

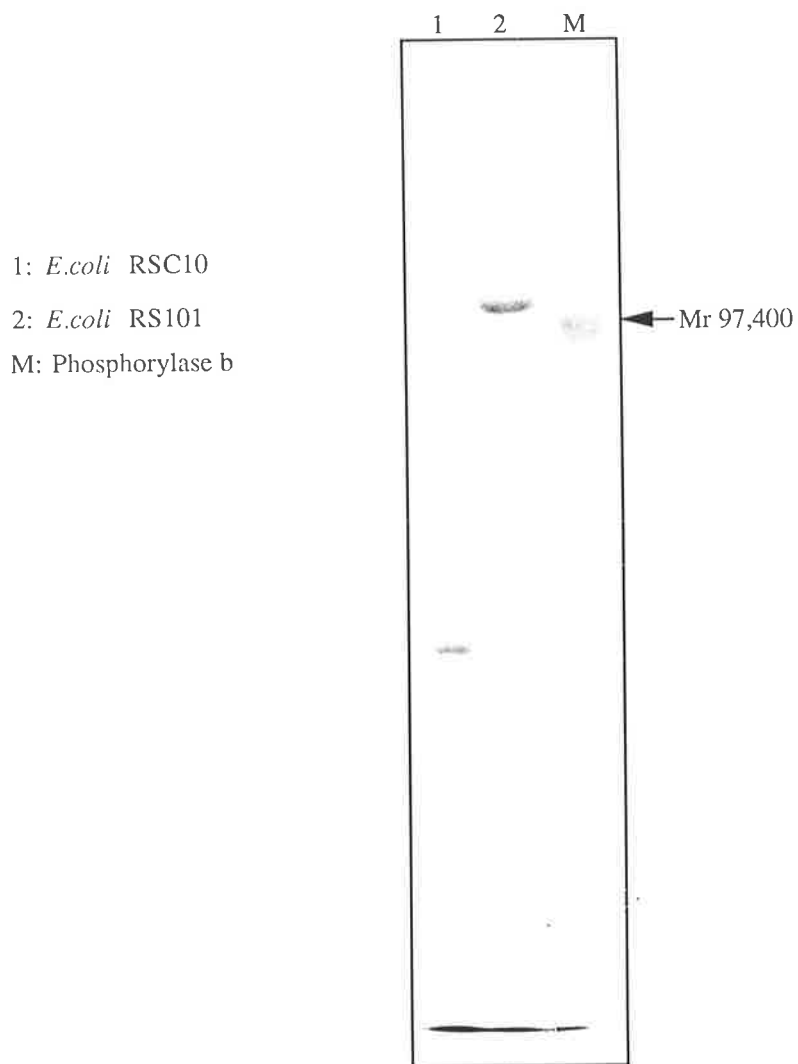
The control bacterium in these experiments was *E.coli* E832 that had been transformed with an incomplete T7 RNA polymerase gene, prepared by the ligation of the EcoR1/HindIII T7 fragment of M13/6H into EcoR1/HindIII-digested pPLC236 and used to transform *E.coli* E832. This clone, designated RSC10 is identical to RS101 except for its T7 DNA-bearing plasmid which lacks the 5'-region of the T7 RNA polymerase gene.

### **3.3.11 Assays of T7 RNA polymerase expression**

#### **3.3.11.1 Detection of protein**

Protein from the heat-induced T7 RNA polymerase gene was examined by SDS-PAGE (Laemmli 1970) and by enzyme assay. A 20 µl sample of induced RS101 was spun for 5 minutes in an Eppendorf centrifuge at room temperature. The pelleted bacteria were

Figure 3.20 Expression of reconstructed T7 RNA polymerase gene in *E. coli* RS101



Note: *E. coli* RSC10 is *E. coli* strain E832 which has been transformed with vector pPLC236 that harbors the T7 RNA polymerase gene with the 5'-deletion of the RBS and the first five codons of the gene.

*E. coli* RS101 is *E. coli* strain E832 which has been transformed with vector pPLC236 that harbors the completely reconstructed T7 RNA polymerase gene.



resuspended in 20  $\mu$ l of gel loading buffer, heated for 5 minutes at 95°C and electrophoresed through an SDS-polyacrylamide gel.

Unlike the control clone RSC10, the protein pattern from RS101 includes a significant band which has an apparent Mr of 98,000. This is the correct size for the RNA polymerase and represents the most abundant protein expressed by the bacteria (see figure 3.20).

### 3.3.11.2 Assay of enzyme activity

A 1 ml sample of induced RS101 was digested with 1 mg of lysozyme for 5 minutes at 25°C. The sample was sonicated for 3 x 10 seconds and spun in an Eppendorf centrifuge for 15 minutes at 5°C. A 1  $\mu$ l aliquot of supernatant was assayed in a 50  $\mu$ l solution containing 40 mM Tris-Cl, pH 8.0, 8 mM MgCl<sub>2</sub>, 5 mM  $\beta$ -mercaptoethanol, 1 mM KH<sub>2</sub>PO<sub>4</sub>, 0.4 mM of each NTP, 5  $\mu$ Ci (1.7 pmoles)  $\alpha$ -<sup>32</sup>P-UTP and 1  $\mu$ g T7 DNA. The reaction proceeded for 10 minutes at 37°C and was stopped by spotting 40  $\mu$ l onto a Whatman GF/A filter and washing in 3 x 100 ml of 1 M HCl /1% Na<sub>4</sub>P<sub>2</sub>O<sub>7</sub> /1% KH<sub>2</sub>PO<sub>4</sub> at 4°C. The filters were washed briefly in ethanol, dried and measured by liquid scintillation counting. Clone RS101 had no detectable T7 RNA polymerase activity.

Several factors may contribute to the protein being enzymically inactive:

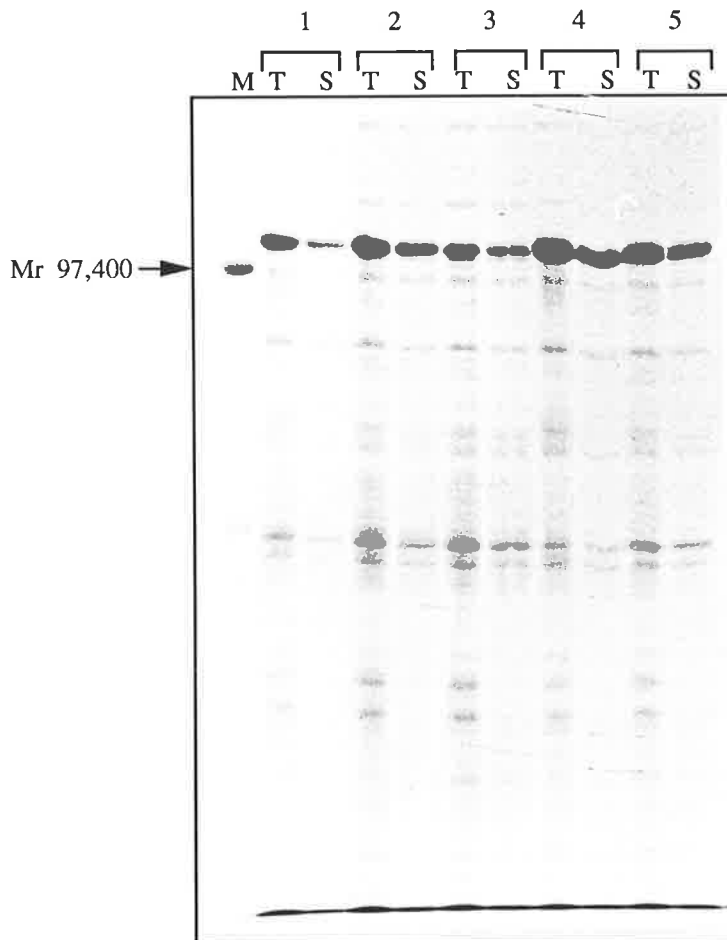
1. The protein may be unable to fold correctly when expressed at very high levels.
2. The gene includes an in-frame mutation(s).
3. The Mr 98,000 protein band is either not T7 RNA polymerase or initiation of translation is within the gene, giving a slightly truncated protein.

The latter two arguments were resolved by DNA sequencing the gene and by direct protein sequencing the amino-terminus of the Mr 98,000 protein.

It has been demonstrated that overexpressing some recombinant enzymes results in protein which lacks measurable enzyme activity, often because the proteins are packaged into inclusion bodies, which can sometimes be seen as lumps in the host bacteria, and which may be isolated to reconstitute the protein to its native conformation (Ho et al. 1982, George et al. 1985, Schoner et al. 1985).

Observation of induced *E.coli* RS101 under phase contrast light microscopy was

Figure 3.21 Expression of reconstructed T7 RNA polymerase gene in *E.coli* RS101



Legend

M: phosphorylase b ( Mr 97,400 )

1-5 are samples of IPTG-induced *E.coli* RS101

Clone RS101 was grown to an  $A_{600}$  of 1.0 and then induced by the addition of IPTG to 100 mM for 3 hours at 37°C.

T: Total. 30  $\mu$ l of the induced culture was spun at 3,000 rpm in an Eppendorf centrifuge for 5 minutes. The supernatant was discarded and the pellet dissolved in 30  $\mu$ l of loading buffer. The pellet thus represents the total protein from induced RS101. The sample was heated to 90°C for 5 minutes and 20  $\mu$ l loaded onto the polyacrylamide gel.

S: Supernatant. 1 ml of the induced culture was spun at 3,000 rpm in an Eppendorf centrifuge for 5 minutes. The supernatant was discarded and the pellet resuspended in 500  $\mu$ l of 40 mM Tris-Cl, 10 mM MgCl<sub>2</sub>, 2 mM  $\beta$ -mercaptoethanol. The culture was then sonicated for 15 seconds and then spun for 20 minutes in an Eppendorf centrifuge. To 15  $\mu$ l of the supernatant was added 15  $\mu$ l of loading buffer. The sample was heated to 90°C for 5 minutes and 20  $\mu$ l loaded onto the gel.

therefore done to see if inclusion bodies, and thus insoluble T7 RNA polymerase, could be found, but the bacteria had a normal phenotype.

### 3.3.11.3 Is the Mr 98,000 protein from *E.coli* RS101 soluble or insoluble?

*E.coli* RS101 was grown in 100 ml of L-broth containing 100 µg/ml ampicillin, at 37°C until the A<sub>600</sub> reached 1.0, and then induced by the addition of IPTG to 100 mM and incubated a further 3 hours at 37°C.

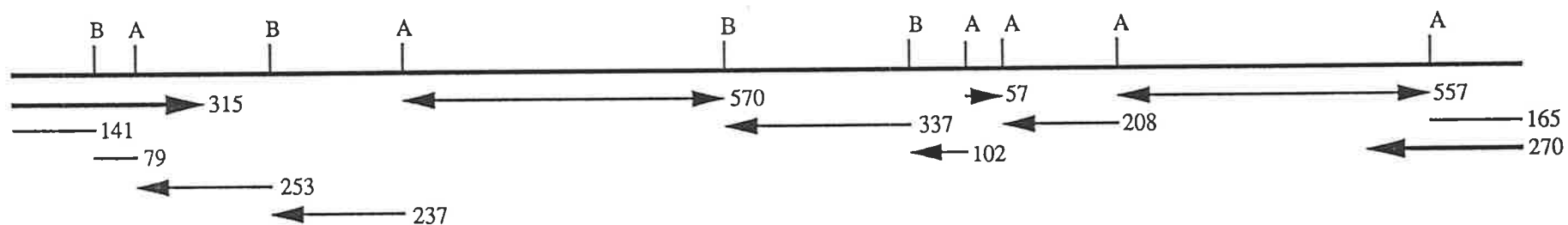
The total cellular protein was determined by centrifuging a 30 µl sample of the induced culture, resuspending the pelleted cells in 30 µl of loading buffer, heating at 90°C for 5 minutes and then electrophoresing through a 10% polyacrylamide gel (with a 4% stacking gel).

Determination of whether the Mr 98,000 protein was soluble involved centrifuging lysed cells prior to loading onto the polyacrylamide gel. Five 1 ml samples of the induced culture were first centrifuged and the pellets resuspended in 500 µl of 40 mM Tris-Cl, pH 8, 10 mM MgCl<sub>2</sub> and 2 mM β-mercaptoethanol. The cells were then sonicated for 15 seconds and then centrifuged for 20 minutes in an Eppendorf centrifuge. To 15 µl of each supernatant was added 15 µl of 2 x loading buffer. The samples were heated to 90°C for 5 minutes prior to loading onto the polyacrylamide gel.

Figure 3.21 compares total cellular protein (T) from lysed cells to the protein remaining in the supernatant (S) after centrifugation of the lysed cells. Although the five samples that were loaded onto the gel are identical aliquots from a single flask of bacterial culture, figure 3.21 shows an apparent increase in solubility from sample 1 to sample 5. The only difference between these samples is the way they were handled; that is, sample 1 was manipulated first, sample 2 second etc. Therefore, the variation in solubility from sample 1 to sample 5 may indicate that protein is leaching from the pellets into the supernatant.

In a variation of this experiment, the pelleted cells were resuspended in loading buffer and then sonicated and centrifuged. The solubilisation of the proteins by the SDS in the load buffer resulted in the total protein and soluble protein tracks being identical (data not shown). This indicated that the apparent reduction in soluble proteins shown by figure 3.21 is due to the

Figure 3.22 Cloning strategy for sequencing the 2.7 kbp reconstructed T7 RNA polymerase gene from pT75WT



legend

A: HincII

B: HaeIII

Numbers indicate the fragment size in base pairs (except for sequence data from the termini of the 2.7 kbp fragment. Terminal sequence data avoided having to clone the 141, 79 and 165 bp fragments.)

Arrows indicate the direction of sequencing

protein being in an insoluble form and not an artefact of the extra centrifugation step in preparing the 'soluble protein' sample.

#### **3.3.11.4 DNA sequencing strategy**

Plasmid T75WT (10 µg) was digested with EcoR1 and HindIII and the 2.7 kbp T7 DNA fragment was purified by agarose gel electrophoresis and then digested with HincII and HaeIII. The digested DNA was then purified on a 2 M urea , 6% polyacrylamide-TBE gel. The DNA was stained with 0.5% toluidine blue and eight bands were isolated and the DNA eluted from the gel slices and then subcloned into Sma1-digested M13mp18 and M13mp19 (see figure 3.22). The DNA sequence of the gene matched the sequence described by Moffatt et al.(1984).

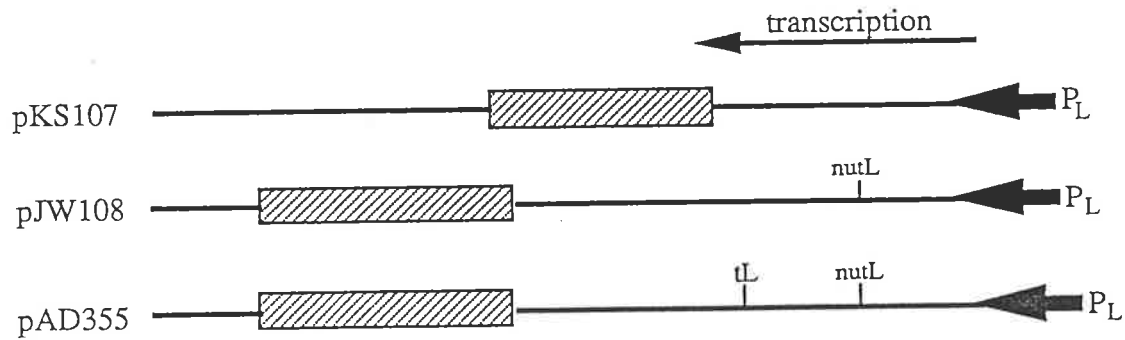
#### **3.3.11.5 Protein sequencing**

Induced *E.coli* RS101 was electrophoresed on an SDS-polyacrylamide gel as described, then electroblotted onto a sheet of polyvinylidene difluoride (Millipore immobilon) essentially as described by Matsudaira (1988).

The polyacrylamide gel was allowed 24 hours to polymerize prior to use as this has been found to improve the likelihood of gaining protein sequence data from proteins which have been isolated from a polyacrylamide gel (B. Forbes per. comm.).The immobilon was first wetted in methanol for 2 minutes, then soaked in transfer buffer (10 mM 3-[cyclohexylamino]-1-propane sulfonic acid, pH 11.0, 10 % (v/v) methanol). Prior to the transfer, the gel was washed in water to reduce its glycine content. The transfer utilized a Hoefer model TE22 transfer apparatus and the transfer current was maintained at 400 mA for 30 minutes.

The membrane was then washed in water for 5 minutes before staining in 50% methanol/5% acetic acid/0.5% Coomassie blue R. After 10 minutes staining, the membrane was destained in 50% methanol/ 10% acetic acid for 10 minutes. The Mr 98,000 band was excised and given to the Adelaide University Centre for Gene Technology Protein Chemistry Unit, where the protein was loaded onto an Applied Biosystem 470A protein sequencer that was linked to a Brownlee Labs HPLC and Waters 441 UV detector. Analysis of HPLC profiles from the protein sequencer unambiguously revealed the first seventeen amino acids of

Figure 3.23 Attenuation of gene expression from the lambda  $P_L$  promoter



Adapted from Das and Wolska (1984)

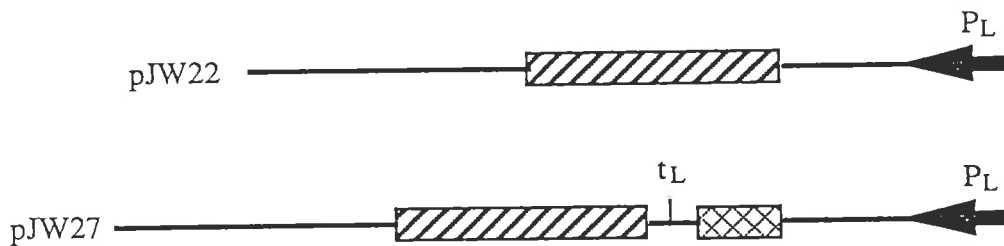
Legend

pKS107 Reference plasmid.  
 pJW108 Similar to pKS107 but with a longer promoter-gene distance.  
 pAD355 Similar to pJW108 but with a transcription terminator between promoter and the gene.

t<sub>L</sub> leftward transcription terminator from phage lambda

nutL binding site for N gene product


 gal K



Adapted from White and Richardson (1988)

Legend

 T7 gene 19

 phage lambda N gene

t<sub>L</sub> leftward transcription terminator from phage lambda

the Mr 98,000 protein, which matched the first seventeen amino acids of T7 RNA polymerase.

Given that pT75WT appears to faithfully encode the T7 RNA polymerase gene and since *E.coli* RS101 expresses protein that has a normal amino-terminal sequence for T7 RNA polymerase, the T7 RNA polymerase is clearly translated from the first AUG codon 3' to the ribosome binding site and is the product of a gene lacking mutations. The lack of demonstrable enzyme activity therefore appears to be due to some factor other than the primary structure of the protein. / /

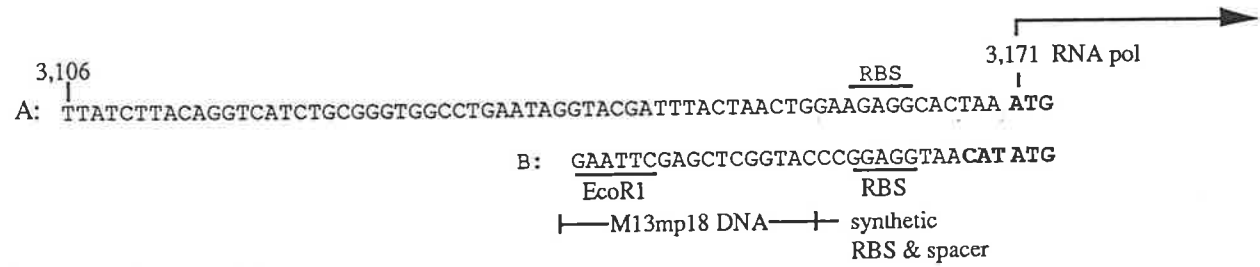
### 3.4 DISCUSSION

The work describe in this chapter involved the isolation of the T7 RNA polymerase gene from the T7 genome and cloning of the gene and expression of the T7 RNA polymerase protein in *E.coli*. That the protein did not show functional enzyme activity was not anticipated. It is possible that there are many reasons why the recombinant T7 RNA polymerase was enzymically inactive.

The simplest explanation is that a mutation in the T7 RNA polymerase gene, possibly gained during *in vitro* mutagenesis (Exp. 3.3.9.2.1 and 3.3.9.4) resulted in an amino acid(s) substitution which altered the structure and biological activity of the protein. The T7 RNA polymerase gene from pT75WT was sequenced (Exp. 3.3.11.4), but only 42% of the gene was sequenced from both strands of the DNA. Although the complete nucleotide sequence matched the known sequence, the possibility of an error in the DNA sequencing cannot be excluded. Complete sequencing of the DNA of both strands of the gene from pT75WT would have increased the confidence in recognising the recombinant T7 RNA polymerase gene as having a wild-type DNA sequence.

The use of the lambda P<sub>L</sub> promoter to express the recombinant T7 RNA polymerase gene may have contributed to the recombinant protein being expressed to a high level but not being enzymically active. It has been reported that the expression of proteins of similar size to T7 RNA polymerase, under P<sub>L</sub> control from expression vectors in *E.coli*, can be expressed in both soluble and insoluble forms (White and Richardson 1988). For example, it was reported by White and Richardson (1988) that phage T7 gene 19, which is involved in phage capsid maturation and packaging, when cloned downstream of the lambda P<sub>L</sub> promoter

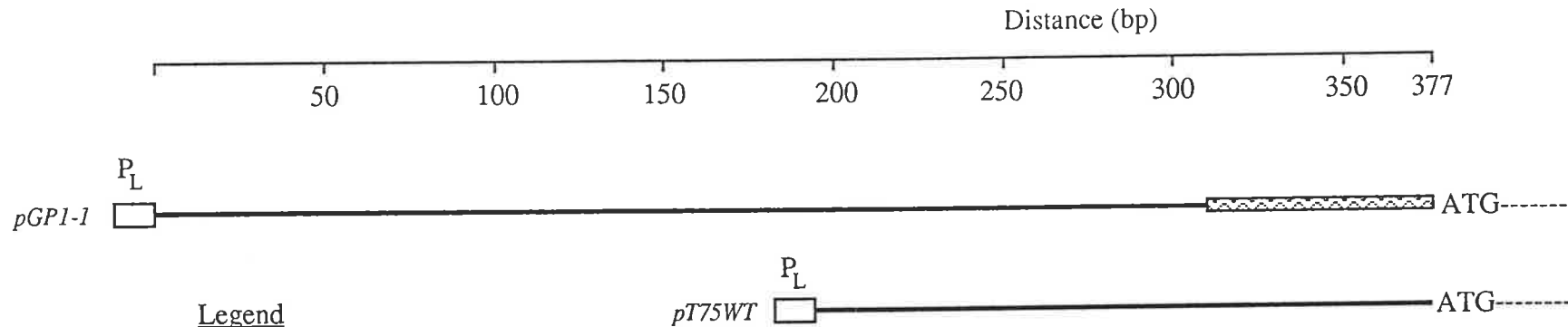
Figure 3.24 (a) The T7 DNA sequence upstream of the T7 RNA polymerase gene in pGP1-1 (Tabor and Richardson 1985)



Legend

- A: 5' T7 DNA sequence from pGP1-1 (Tabor and Richardson 1985)  
 B: 5' DNA sequence of the cloned fragment harboring the T7 RNA polymerase gene from pT75WT (Exp 3.3.10)

Figure 3.24 (b) Scale map indicating distance from Promoter to T7 RNA polymerase gene initiation (ATG) codon : Comparison of pGP1-1 and pT75WT.



Legend

- Vector DNA
- T7 DNA 5' to T7 RNA polymerase gene initiation codon
- $P_L$  Leftward promoter from phage  $\lambda$



and expressed in *E.coli* at 42°C, yielded protein which was insoluble. However, attenuating the expression level of T7 gene 19 from the P<sub>L</sub> promoter resulted in the recovery of T7 gene 19 protein which was soluble. This attenuation was achieved by inserting the phage lambda leftward transcription terminator ( $\lambda t_{L1}$ ) and the N-protein gene from phage lambda between the P<sub>L</sub> promoter and T7 gene 19 (see figure 3.23).

It had previously been demonstrated that the level of gene expression from P<sub>L</sub> *in vitro* can be reduced by the combination of  $\lambda t_{L1}$  and the lambda N gene when inserted between P<sub>L</sub> and a reporter gene (Das and Wolska 1984: see figure 3.23). The antitermination effect of the N gene product overcomes the cessation of transcription that would otherwise occur if  $\lambda t_{L1}$  were placed between P<sub>L</sub> and a gene. The report by Das and Wolska (1984) describes the quantitative estimate of galactokinase activity from the gal K cistron in several plasmids which differ in the presence or absence of  $\lambda t_{L1}$  between the gal K cistron and P<sub>L</sub>. The plasmid pAD355 (figure 3.23) resulted in about 55% of the galactokinase activity as that directed by plasmid pKS107.

White and Richardson (1988) argue that expression of T7 gene 19 from P<sub>L</sub> at 42°C is attenuated because the plasmid pJW27 encodes  $\lambda t_{L1}$  and the lambda N gene between P<sub>L</sub> and gene 19. Whether the combination of  $\lambda t_{L1}$  and the lambda N gene would actually reduce gene expression to the same degree *in vivo* is unknown. A comparative study by Brunner and Bujard (1987) of nine promoters, including P<sub>L</sub>, demonstrated the poor correlation between promoter strength *in vitro* and promoter strength *in vivo*. Although the effects of promoter strength and the effects of N-protein on gene expression are different phenomena, Brunner and Bujard's observations indicate the caution required when extrapolating data from studies of recombinant genetic constructs *in vitro*, as an explanation for observations of similar genetic constructs operating *in vivo*.

An alternate explanation of the success of pJW27 over pJW22 (which directs the synthesis of inactive gene 19 product) could possibly be related to another observation of Das and Wolska (1984); that pJW108 directs galactokinase synthesis to 78% of the level produced from pKS107 (see figure 3.23). Plasmid pJW108 differs from pKS107 by having about 2 kbp of lambda DNA and pBR322 DNA inserted between P<sub>L</sub> and the gal K cistron. That is, the

Figure 3.25

The open reading frames upstream of the T7 RNA polymerase gene in pGP1-1 (Tabor and Richardson 1985)

Translation →

1. Leu Ser Tyr Arg Ser Ser Ala Gly Gly Leu Asn Arg Tyr Asp Leu Leu Thr Gly Arg Gly Thr Lys S  
TTA TCT TAC AGG TCA TCT GCG GGT GGC CTG AAT AGG TAC GAT TTA CTA ACT GGA AGA GGC ACT AAA TGA AC

2. Tyr Leu Thr Gly His Leu Arg Val Ala S RBS f-Met Asn  
T TAT CTT ACA GGT CAT CTG CGG GTG GCC TGA ATAGGTACGATTTACTAACTGGAAGAGGCACTAA **ATG** AAC

3. Ile Leu Gln Val Ile Cys Gly Trp Pro Glu S Val Arg Phe Thr Asn Trp Lys Arg His S f-Met Asn  
TT ATC TTA CAG GTC ATC TGC GGG TGG CCT GAA TAG GTA CGA TTT ACT AAC TGG AAG AGG CAC TAA **ATG** AAC

S = stop codon

The initiation codon for the T7 RNA polymerase gene is shown in bold

attenuation of gal K expression from pJW108 appears to be due to the distance between  $P_L$  and the gal K cistron. Plasmid pJW27 (encoding T7 gene 19) has about 650 bp more DNA between  $P_L$  and T7 gene 19 than does pJW22. White and Richardson (1988) reported the expression of T7 gene 19 protein from pJW27 to be about 30% of the expression level observed with pJW22, but the minimum attenuation required in T7 gene 19 expression to yield soluble, biologically active protein is unknown. Hence the relative contributions of placing  $\lambda_{tL1}$  and the lambda N gene between  $P_L$  and gene 19, and the effect of increasing the distance between  $P_L$  and T7 gene 19 is difficult to quantify.

The role of promoter-gene spacing may explain the report by Tabor and Richardson (1985) of enzymically active T7 RNA polymerase expression in *E.coli*, from the lambda  $P_L$  promoter, after 42°C heat induction. Their clone harbours the construct pGP1-1 which encodes a T7 DNA fragment bearing the RNA polymerase gene and 65 bp of DNA 5' to the initiation codon (see figure 3.24 a). The precise difference between the expression in *E.coli* of pGP1-1 and pT75WT (Exp. 3.3.10) is unknown, partly because of the different host genotypes and also because the distances vary between promoter and RBS (see figure 3.24 b). If the magnitude of gene expression has a critical value, beyond which the tertiary structure of large proteins is compromised, it is possible that the difference in promoter-gene spacing between pGP1-1 and pT75WT could account for the observable difference in gene expression between the two clones.

Another difference between the expression, in *E.coli*, of pGP1-1 and pT75WT (Exp. 3.3.10) is that from pGP1-1 "Transcription from the  $P_L$  promoter also results in the synthesis of a Mr 6,000 protein containing the first 33 amino acids of the N protein of  $\lambda$  fused to 22 amino acids encoded by the region before gene 1, and terminating 2 bp before the start codon for gene 1." (Tabor and Richardson 1985). However, this is unconvincing given that the T7 DNA fragment in pGP1-1, when read in three reading frames cannot code for protein terminating 2 bp short of the initiation codon for the T7 RNA polymerase gene (see figure 3.25). The longest open reading frame upstream of the T7 RNA polymerase gene ends with a UGA stop codon, which starts within the AUG initiation codon. If *E.coli* RNA polymerase read through the UAG stop codon in frame 3, this could allow the next in-frame stop codon

(UAA) to prevent translation into the T7 RNA polymerase open reading frame.

In pT75WT there is partial sequence for the N protein gene between the P<sub>L</sub> promoter and the T7 RNA polymerase gene. However, translation of N gene protein does not occur because the initiation signal for the gene is missing.

Hence, another possible explanation for the observed difference between the expression, in *E.coli*, of pGP1-1 and pT75WT (Exp. 3.3.10) could be a limited ribosome access to the mRNA, due to translation through the RBS ahead of the T7 RNA polymerase gene.

Whether attenuating the expression level of the T7 RNA polymerase gene would also facilitate the recovery of enzymically active T7 RNA polymerase was then determined. This work is described in the following chapter.

## **CHAPTER 4**

### **BACTERIOPHAGE T7 RNA POLYMERASE**

**-ISOLATING THE PHAGE T7 RNA POLYMERASE GENE BY THE  
POLYMERASE CHAIN REACTION**

**-EXPRESSION OF THE PCR-DERIVED T7 RNA POLYMERASE GENE IN  
*E.COLI***

## 4.1 INTRODUCTION

### Attenuated expression of T7 RNA polymerase in *E.coli*

The strength of *E.coli* RNA polymerase promoters is a function of the nucleotide sequence at the -35 and -10 regions and also on the number of nucleotides (spacer length) between the -35 and -10 regions (Auble et al. 1986, Brosius et al. 1985, Deuschle et al. 1986). One means of altering the expression level from a given promoter is to reconstruct the promoter. This is feasible with comparative data of promoter sequence variants that have been correlated to promoter function.

A well characterised example of this are the variants of the *tac* promoter. The promoter  $P_{tac}$  (DeBoer et al. 1983) is a hybrid promoter which has the -35 region of the *E.coli* tryptophan operon promoter and the -10 region of the lac UV5 promoter, with a 16 bp spacer. These promoters, called  $P_{tac}$  or TAC16 and two variants, TAC17 or  $P_{trc}$  and TAC18 or  $P_{tic}$  are identical except for the spacer length of 16, 17 or 18 bp. The TAC17 promoter has the greatest promoter strength *in vitro* (Mulligan et al. 1985) and the TAC16 promoter has the greatest promoter strength *in vivo* (Brosius et al. 1985). The attenuation of TAC16 promoter strength *in vivo* was demonstrated by fusing the three TAC promoters to the *E.coli* 4.5S RNA gene and quantifying the amount of RNA transcript by a densitometer scan of the RNA after electrophoresis in a polyacrylamide gel. The TAC17 promoter produced 90% as much 4.5S RNA as the TAC16 promoter and TAC18 produced only 65% as much 4.5 S RNA as the TAC16 promoter. The experiment was not run in duplicate, hence the actual quantitative difference between TAC16 and TAC17 directed gene expression is questionable. Similar results are reported for a mutant lac promoter in which the mutant promoter having a 20 bp spacer shows 25% of the expression observed with the wild type promoter having an 18 bp spacer (Mandecki and Reznikoff 1982).

Another, more conservative approach to attenuating gene expression is to take advantage of a promoter which is known to have weaker promoter strength *in vivo*. A comparison of 14 promoters (Deuschle et al. 1986) demonstrated that several controllable promoters have strengths much less than the lambda  $P_L$  promoter. The  $P_L$  promoter was used in the expression of the reconstructed T7 RNA polymerase gene described in the previous

Table 4.1 Gene expression in *E.coli* from phage lambda and *E.coli* promoters

Vector	Prm	Ti	Active	Mr	Gene	Reference
pTP2	P <sub>L</sub>	42°C	YES	17	T7 gene 3	Pham & Coleman (1985)
pYS6	P <sub>L</sub>	42°C	YES	34	T4 gene 32	Shamoo et al.(1986)
pJW22	P <sub>L</sub>	42°C	NA	66	T7 gene 19	White & Richardson (1988)
pJW27	P <sub>L</sub>	42°C	NA	66	T7 gene 19	White & Richardson (1988)
pGP-1	P <sub>L</sub>	42°C	YES	98	T7 RNA pol	Tabor & Richardson (1985)
pTL43W	P <sub>L</sub>	42°C	NO	94	T4 DNA pol	Lin et al.(1987)
pTL43W	P <sub>L</sub>	40°C	YES	94	T4 DNA pol	Lin et al.(1987)
pRC23	P <sub>L</sub>	42°C	NO	55	Poliovirus RNA polymerase	Plotch et al.(1989)
pT7-POL	T7	37°C	YES	55	Poliovirus RNA polymerase	Plotch et al.(1989)
pTL43Q	P <sub>tac</sub>	42°C	NO	94	T4 DNA pol	Lin et al.(1987)
pTL43Q	P <sub>tac</sub>	40°C	YES	94	T4 DNA pol	Lin et al.(1987)
1CM21	P <sub>R'</sub>	37°C	YES	33	T4 PNK	Midgley & Murray (1985)
pCM56	P <sub>lacUV5</sub>	37°C	YES	100	T3 RNA pol	Morris et al.(1986)
pUC12N	P <sub>lac</sub>	37°C	YES	66	HIV RT	Hizi et al.(1988)

Legend

- Prm promoter  
Ti temperature of gene induction  
Mr apparent molecular weight in kD  
NA not applicable

chapter. The experimental strategy employed by Deuschle et al. (1986) involved substituting various promoters ahead of a reporter gene, inducing gene expression and then probing bacterial cell lysates with a single stranded M13 probe which bound to RNA transcripts from the reporter gene. Unfortunately the promoter sequences were inserted into the expression plasmid as restriction fragments. Hence the promoter-reporter gene distance, transcript length ahead of the reporter gene and the restriction fragment size vary in each construct. These factors highlight the difficulty of experimentally relating the function of a promoter element to its nucleotide sequence, let alone the secondary and tertiary structures that may be involved in promoter function.

Two other factors which correlate to the expression of soluble, active protein when expressed *in vivo* are the temperature of induction and the size of the protein (see table 4.1). The expression of cloned phage T4 DNA polymerase (Mr 94,000) from the lambda P<sub>L</sub> promoter in an expression vector (pTL43W) was found to be insoluble and inactive when gene expression was induced at 42°C, but was found to be soluble and active when the induction temperature of the bacteria was 40°C (Lin et al. 1987). This is in contrast to the expression of cloned T7 gene 3 (Mr 17,000) from the P<sub>L</sub> promoter in another expression vector (pTP2) which was soluble and active after induction at 42°C (Pham and Coleman 1985).

If the expression of T7 RNA polymerase in a soluble and active form depends on a lesser rate of gene expression than its expression in *E.coli* RS101, it may be optimal to express the T7 RNA polymerase gene under the control of a promoter that has a weaker promoter strength *in vivo* and to keep the gene induction temperature below 40°C.

The approach taken to reduce the expression level of T7 RNA polymerase in *E.coli* was to re-clone the gene into another expression vector (pEDV-1) downstream of a controllable promoter (P<sub>lac</sub>), which exhibits weaker promoter strength *in vivo* than the P<sub>L</sub> promoter (Deuschle et al. 1986). Expression of the T7 RNA polymerase gene in *E.coli* was then induced at 37°C by adding IPTG to the bacterial culture.

Prior to cloning the T7 RNA polymerase gene into pEDV-1, the T7 RNA polymerase gene was isolated from T7 genomic DNA by exploiting the polymerase chain reaction (PCR). This technique potentially allows for DNA fragments as large as the T7 RNA polymerase gene





to be isolated from genomic DNA, with precisely defined termini (Saiki et al. 1986, Scharf et al. 1986). This allowed a direct comparison to the previous approach of repairing Bal 31 nuclease-digested DNA with site-directed mutagenesis, as a means of isolating a specific DNA fragment from genomic DNA, independently of restriction sites.

## 4.2 METHODS and RESULTS

### 4.2.1 Construction of expression vector pEDV-1

The cloning/expression vector described in this work was derived from pUC18. The  $P_{lac}$  promoter in pUC18 lies upstream from the coding region of a  $\beta$ -galactosidase gene, within which is the EcoR1  $\rightarrow$  HindIII multiple cloning site polylinker. pUC18 was designed to allow a short peptide from the  $\beta$ -galactosidase gene to be fused to the translation product of a gene cloned into the multiple cloning sites, and can be useful if the foreign gene lacks prokaryote initiation signals.

The T7 RNA polymerase gene which is to be extracted from T7 genomic DNA by PCR (described below) includes 5'-untranslated sequence, including a RBS; hence, pUC18 was modified by engineering a translation stop codon at the start of the multiple cloning site polylinker. The stop codon allows initiation of translation to occur from the first AUG codon of the T7 RNA polymerase gene mRNA transcript rather than having the enzyme fused to a  $\beta$ -galactosidase peptide.

Digestion of pUC18 with EcoR1, followed by end filling and religation of the blunt ends results in the generation of the translation stop codon, in frame with the  $\beta$ -galactosidase gene (see figure 4.1).

5  $\mu$ g of pUC18 DNA was first digested with 10 units of EcoR1 in medium salt buffer for 90 minutes at 37°C, in a reaction volume of 17  $\mu$ l. The DNA was then end-filled by adding 1  $\mu$ l of 10 x end-fill buffer, 1  $\mu$ l of 1 mM dNTPs and 1  $\mu$ l (1 unit) of Klenow polymerase and incubated for 30 minutes at 37°C. The DNA was then purified on a 1% agarose-TAE gel, isolated from the agarose with Bio-101 GeneClean and dissolved in water to a concentration of 150 ng/ $\mu$ l. The blunt ends of the DNA were then religated together by incubating 1  $\mu$ l of the DNA in a 20  $\mu$ l solution containing 50 mM Tris-Cl, pH 7.6, 5 mM MgCl<sub>2</sub>, 0.5 mM ATP, 0.5



mM DTT and 2 units of T4 DNA ligase overnight at 25°C. The entire ligation mixture was used to transform 200 µl of competent *E.coli* JM101. The transformed bacteria were grown overnight at 37°C on L-plates containing 100 µg/ml ampicillin. Plasmids from six bacterial colonies were isolated and digested with EcoR1 or HindIII, in medium salt buffer. All plasmids failed to be digested by EcoR1, indicating the absence of the EcoR1 restriction site. The multiple cloning site polylinker in pEDV-1 still maintained the unique Hind III site downstream of the Sma I site. The HindIII digestion of pEDV-1 was to verify that the plasmids could be cut (see figure 4.1).

#### **4.2.2 Isolation of the T7 RNA polymerase gene from phage T7 DNA by PCR**

##### **4.2.2.1 Design of the PCR oligonucleotide primers**

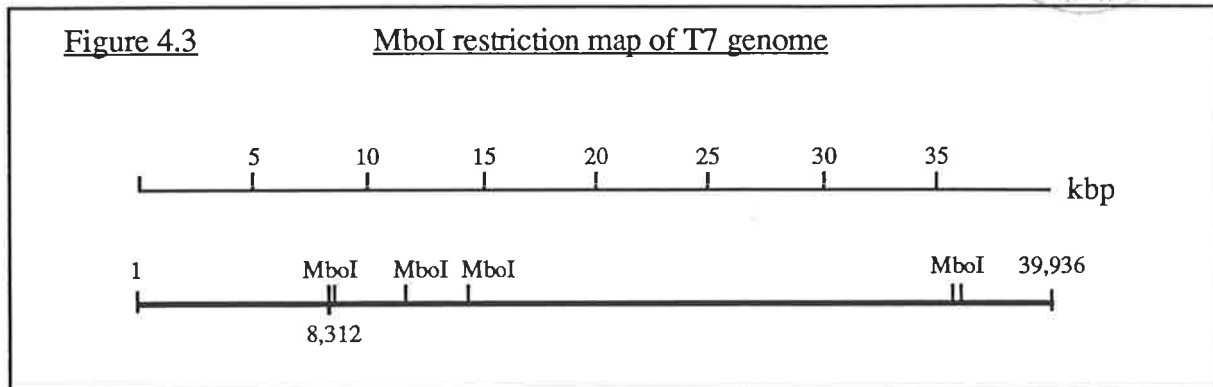
Given that the 5'-terminal nucleotide of the PCR oligonucleotide primers determines the limits of the amplified DNA sequence, it is possible to extract a target sequence with absolute precision. The *E.coli* clone described by Davanloo et al.(1984) which produces enzymically active T7 RNA polymerase, exploits a similar promoter (lac UV5) to the lac promoter in pEDV-1, and as the DNA fragment which harbours the T7 RNA polymerase gene does not appear to affect the expression of the T7 RNA polymerase detrimentally, it was decided to extract, from T7 phage genomic DNA, a DNA fragment identical to that cloned by Davanloo et al.(1984).

The T7 RNA polymerase clone described by Davanloo et al.(1984) is *E.coli* strain HMS174 harbouring the recombinant vector pAR1219, within which is the T7 RNA polymerase gene, encoded in a T7-derived DNA fragment that extends from 3,145 to 5,841 (T7 coordinates). Two 24-mer oligonucleotide primers, obtained from Bresatec, were designed for use in the PCR to amplify the region of T7 genomic DNA from 3,145 bp to 5,841 bp inclusive (see figure 4.2).

##### **4.2.2.2 PCR amplification of phage T7 DNA**

The PCR protocol was determined empirically and was based on theoretical considerations

described by Saiki (1989).



Ten micrograms of T7 genomic DNA was digested in medium salt buffer with 14 units of MboI, in a 20  $\mu$ l reaction volume for 90 minutes at 37°C. The DNA was electrophoresed in a 1% agarose-TAE gel and stained with 10  $\mu$ g/ml ethidium bromide. The first Mbo I restriction site from the left end on the T7 genome is at position 8,312. This 5'-terminal 8,312 bp MboI restriction fragment (see figure 4.3) was isolated from the agarose using Bio-101 GeneClean and the DNA dissolved in 20  $\mu$ l of water.

The 50  $\mu$ l reaction mixture contained about 150 ng of the MboI-digested T7 DNA, 10 mM Tris-Cl, pH 8.4, 50 mM KCl, 1.5 mM MgCl<sub>2</sub>, 2  $\mu$ g of BSA, 200  $\mu$ M of each dNTP, 1  $\mu$ g of each 5'-phosphorylated 24-mer oligonucleotide primer and 3 units of Cetus Amplitaq polymerase. The reaction mixture was incubated in a Cetus Thermal Cycler with denaturation for 30 seconds at 94°C, annealing for 1 minute at 55°C and extension at 72°C for 5 minutes, through 30 cycles.

#### 4.2.2.3 Cloning the PCR-derived DNA into *E.coli*

The PCR mixture from Exp. 4.2.2.3 was electrophoresed through a 1% agarose-TAE gel and stained with 10  $\mu$ g/ml ethidium bromide. Visualisation of the gel under long wavelength UV light showed a single band which co-migrated with a 2.7 kbp DNA marker (data not shown). The DNA was purified from the agarose with Bio-101 GeneClean and dissolved in 20  $\mu$ l of water. A 10  $\mu$ l sample of the DNA was ligated to 100 ng of SmaI-cut, dephosphorylated

pEDV-1 in a 20  $\mu$ l solution containing 50 mM Tris-Cl, pH 7.6, 10 mM MgCl<sub>2</sub>, 10 mM DTT, 5% PEG 8000 and 7 units of T4 DNA ligase. The reaction proceeded overnight at 25°C. The ligation mixture was used to transform 200  $\mu$ l of competent *E.coli* JM101. The transformed bacteria were grown overnight at 37°C on L-plates containing 100  $\mu$ g/ $\mu$ l ampicillin, yielding 88 colonies. There were no background colonies .

#### **4.2.3 Assay for enzymically active T7 RNA polymerase *E.coli* clones**

The bacterial colonies were toothpicked into 3 ml of L-broth containing 100  $\mu$ g/ml ampicillin and incubated at 37°C until the A<sub>600</sub> reached 1.0, and were then induced with IPTG to a final concentration of 1 mM. A 1 ml sample of each induced culture was sonicated for 3 x 10 seconds and spun in an Eppendorf centrifuge for 15 minutes at 5°C. A 1  $\mu$ l aliquot of supernatant was assayed in a 50  $\mu$ l solution containing 40 mM Tris-Cl, pH 8.0, 8 mM MgCl<sub>2</sub>, 5 mM  $\beta$ -mercaptoethanol, 1 mM KH<sub>2</sub>PO<sub>4</sub>, 0.4 mM of each NTP, 10  $\mu$ Ci (3.3 pmoles)  $\alpha$ -<sup>32</sup>P-UTP and 1  $\mu$ g T7 DNA. The reaction proceeded for 10 minutes at 37°C and was stopped by spotting 40  $\mu$ l onto a Whatman GF/A filter and washing in 3 x 100 ml of 1 M HCl /1% Na<sub>4</sub>P<sub>2</sub>O<sub>7</sub> /1% KH<sub>2</sub>PO<sub>4</sub> at 4°C. The filters were washed briefly in ethanol, dried and measured by liquid scintillation counting.

Only one of the bacterial cultures, designated RS51, expressed active T7 RNA polymerase and was used in the following experiments.

#### **4.2.4 Validation of expression vector pEDV-1**

Verification that the T7 RNA polymerase expressed from *E.coli* RS51 is not a fusion protein was achieved by electrophoresing IPTG-induced bacteria on a polyacrylamide gel, transferring the proteins electrophoretically to a polyvinylidene difluoride membrane, staining with Coomassie blue R and then loading the Mr 98,000 protein onto an Applied Biosystems protein sequencer. The method is described in detail in Exp. 3.3.11.5.

Analysis of the HPLC profiles from the Protein Sequencer revealed the amino terminal sequence.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
<i>Met</i>	<i>Asn</i>	<i>Thr</i>	<i>Ile</i>	<i>Asn</i>	<i>Ile</i>	<i>Ala</i>	<i>Lys</i>	<i>Asn</i>	<i>Asp</i>	<i>Phe</i>	?	<i>Asp</i>	<i>Ile</i>	?	<i>Leu</i>	<i>Ala</i>
<i>Met</i>	<i>Asn</i>	<i>Thr</i>	<i>Ile</i>	<i>Asn</i>	<i>Ile</i>	<i>Ala</i>	<i>Lys</i>	<i>Asn</i>	<i>Asp</i>	<i>Phe</i>	<i>Ser</i>	<i>Asp</i>	<i>Ile</i>	<i>Glu</i>	<i>Leu</i>	<i>Ala</i>

(Note: the protein sequence data is shown above the known T7 RNA polymerase amino-terminal amino acid sequence. '?' indicates unreadable data).

The first eleven amino terminus amino acids of the sequenced Mr 98,000 protein unambiguously matched the corresponding amino acids in T7 RNA polymerase, indicating the presence of a functional stop codon 5' to the T7 RNA polymerase initiation codon.

#### 4.2.5 T7 RNA polymerase expression in *E.coli* RS51

The commercial utility of *E.coli* RS51 was determined by inducing a liquid culture with IPTG and sampling the induced culture at hourly intervals. The samples were then assayed for T7 RNA polymerase activity and were also examined by SDS-PAGE.

A sample of *E.coli* RS51 was incubated at 37°C in 50 ml of L-broth containing 100 µg/ml ampicillin until the A<sub>600</sub> reached 1.0. Two 1 ml aliquots of the culture were spun in an Eppendorf centrifuge for 5 minutes and the pellets stored at -20 °C. These samples were used as controls for determining the level of T7 RNA polymerase expression prior to induction by IPTG. The remaining culture was induced by adding IPTG to a final concentration of 0.5 mM, with continued incubation at 37°C. At intervals of 2.25, 3, 4, 5, 7.6, 9 and 23 hours, 1 ml aliquots of culture were were spun in an Eppendorf centrifuge for 5 minutes and the pellets stored at -20°C.

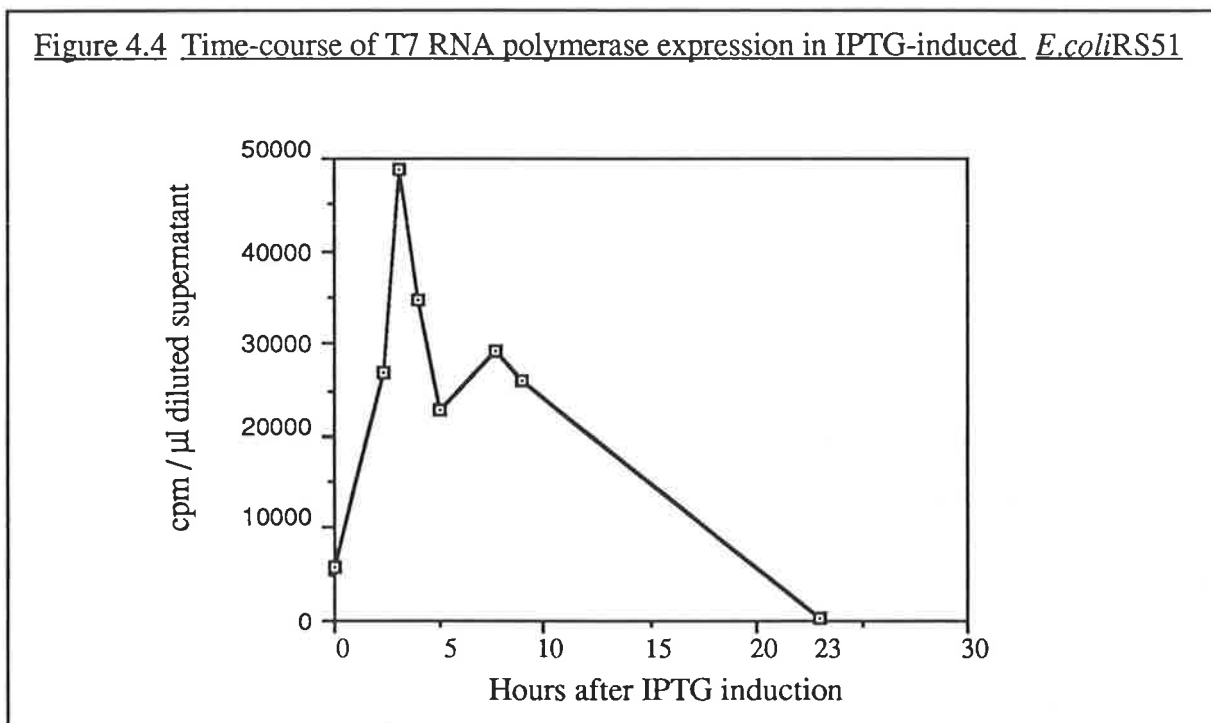
##### 4.2.5.1 Assay of enzyme activity

Shortly before assaying for T7 RNA polymerase activity, the pelleted cells were resuspended in 200 µl of 40 mM Tris-Cl, pH 8, 8 mM MgCl<sub>2</sub> and 1mM β-mercaptoethanol, sonicated for 10 seconds and then spun in an Eppendorf centrifuge for 20 minutes. The supernatant was diluted 1:10 into a solution containing 40 mM Tris-Cl, pH 8, 8 mM MgCl<sub>2</sub>, 1mM β-mercaptoethanol, 100 µg/ml BSA and 5% glycerol. The dilution was found to be

necessary to keep the amount of  $^{32}\text{P}$ -labelled RNA to a level that would give an accurate assay from the scintillation counter.

A 1  $\mu\text{l}$  aliquot of diluted supernatant was assayed in a 50  $\mu\text{l}$  solution containing 40 mM Tris-Cl, pH 8.0, 8 mM  $\text{MgCl}_2$ , 5 mM  $\beta$ -mercaptoethanol, 1 mM  $\text{KH}_2\text{PO}_4$ , 0.4 mM of each NTP, 2.5  $\mu\text{Ci}$  (0.8 pmoles)  $\alpha$ - $^{32}\text{P}$ -UTP and 1  $\mu\text{g}$  T7 DNA. The reaction proceeded for 10 minutes at  $37^\circ\text{C}$  and was stopped by spotting 10  $\mu\text{l}$  onto a Whatman GF/A filter and washing in 3 x 100 ml of 1 M HCl /1%  $\text{Na}_4\text{P}_2\text{O}_7$  /1%  $\text{KH}_2\text{PO}_4$  at  $4^\circ\text{C}$ . The filters were washed briefly in ethanol, dried and measured by liquid scintillation counting.

The time-course demonstrated that optimal recovery of T7 RNA polymerase from *E.coli* RS51 is 3 hours after IPTG induction (see figure 4.4).



#### 4.2.5.2 Examination of *E.coli* RS51 by SDS-PAGE.

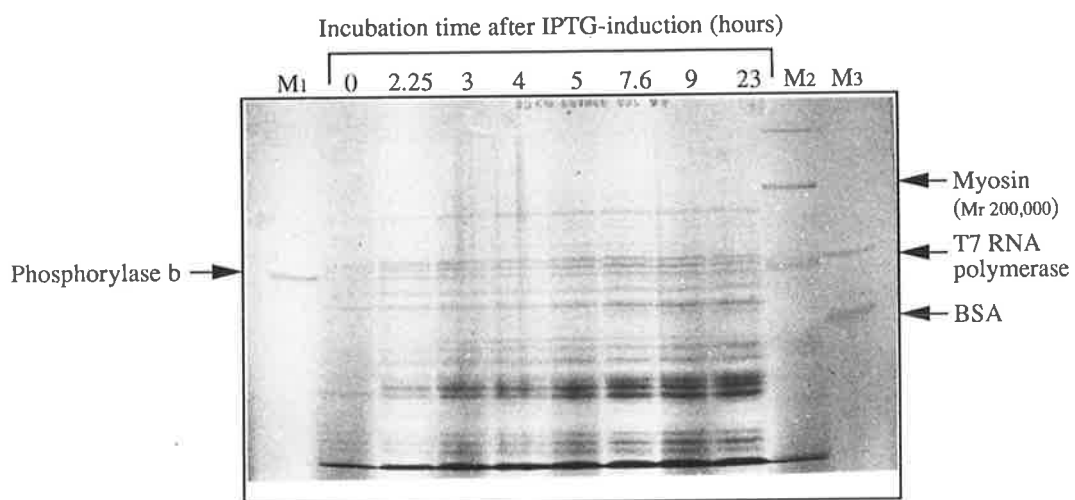
Pelleted bacteria from Exp.4.2.5 were resuspended in 100  $\mu\text{l}$  of loading buffer, heated at  $90^\circ\text{C}$  for five minutes and 20  $\mu\text{l}$  samples were then loaded onto an SDS-polyacrylamide gel as described in the Methods chapter (2.1.8) except that a 4% stacking gel was used.

Each 20  $\mu\text{l}$  sample of *E.coli* RS51 that was loaded onto the SDS-polyacrylamide gel



Figure 4.5

Expression of T7 RNA polymerase from *E.coli* RS51 (Exp. 4.2.5)



Legend

M1: Phosphorylase b Mr 97,400

M2: Amersham Rainbow protein markers

M3: T7 RNA polymerase + BSA (Bresatec)

contained the equivalent of 200  $\mu$ l of bacterial culture. Since each track of the gel received the same volume of bacteria, the amount of protein in each track was primarily a function of cell number. Consequently, the difference in protein expression between the pre-induced and IPTG-induced samples was difficult to quantify. Nevertheless, the data demonstrates that the IPTG-induced samples expressed a protein which co-migrated with, and had the same intensity as a sample of T7 RNA polymerase (from Bresatec Ltd.) that was used as a marker (see figure 4.5).

The T7 RNA polymerase marker was approximately 200 units of enzyme, consequently the culture of induced *E.coli* RS51 is expressing at least 1 million units of T7 RNA polymerase per litre of culture.

### 4.3 DISCUSSION

Clearly, the PCR protocol is appropriate for extracting a 2.7 kbp DNA fragment from genomic DNA. The highly specific assay for T7 RNA polymerase enabled the rapid identification of positive clones from among the transformants. The nucleotide sequence of the PCR-derived T7 DNA fragment in *E.coli* RS51 has yet to be determined, so it remains possible that a mutation(s) could have been introduced by the PCR. A report by Tindall and Kunkel (1988) indicated a Taq polymerase error rate of  $10^{-4}$  in the amplification of the  $\beta$ -galactosidase gene, therefore on that error rate, the 2.7 kbp T7 RNA polymerase gene could be expected to average about 7 nucleotide substitutions. Since only one of the 88 transformants from Exp.4.2.2.3 was expressing T7 RNA polymerase, it is possible that most of the other transformants were full-length T7 RNA polymerase clones but with loss of function mutations. The 87 transformants not expressing T7 RNA polymerase were not studied further.

The induction of T7 RNA polymerase expression in *E.coli* RS51 by IPTG was followed 3 hours later by a peak level of T7 RNA polymerase activity. The measured level of T7 RNA polymerase activity decreased after 3 hours, dropping to half-maximal level by 10 hours after induction. The expression of T7 RNA polymerase reported by Davanloo et al. (1984) and Tabor and Richardson (1985) also exhibited a peak of activity 3 hours after induction, followed by a rapid decrease in activity. The kinetics of T7 RNA polymerase

expression and degradation in *E.coli* RS51 are unknown. Hence, the decreasing level of enzyme activity after 3 hours could reflect a number of factors, such as a decreasing rate of enzyme synthesis with a constant rate of degradation or a constant rate of synthesis with an increasing rate of degradation.

The proteolytic cleavage of T7 RNA polymerase in *E.coli* is known to occur between amino acid 172 (lysine) and amino acid 173 (arginine), giving rise to Mr 78,000 and Mr 20,000 fragments (Tabor and Richardson 1985). The protease mostly responsible for this is believed to be an outer membrane endoprotease, OmpT (Grodberg and Dunn 1988). These authors isolated *ompT* deletion strains by examining the ability of various *E.coli* strains for their ability to cleave T7 RNA polymerase. One of the *ompT* deletion strains, an *E.coli* B strain (BL21), was already being used for the expression of T7 RNA polymerase from the cloned gene (Studier and Moffatt 1986). The advantage of expressing T7 RNA polymerase from *ompT* deletion strains is the ability to greatly increase the time of induction without losing the protein through proteolysis.

Apart from the need to test the value of PCR in isolating a large DNA fragment without mutations, the work described in this chapter aimed to determine the value of attenuating the level of T7 RNA polymerase expression.

The level of expression, from a lac promoter, of T7 RNA polymerase in *E.coli* RS51 was very low in comparison to the level of expression observed in *E.coli* RS101 (Exp.3.3.10) which was controlled by the lambda P<sub>L</sub> promoter. A quantitative estimate of the amount of T7 RNA polymerase protein produced by IPTG-induced *E.coli* RS51 has yet to be determined, nonetheless the qualitative data provided by the protein gel (Exp.4.2.5.2) indicated that the T7 RNA polymerase accounts for only a few percent of the total cellular protein.

The use of an IPTG-inducible lac promoter for expression of T7 RNA polymerase in *E.coli* RS51 certainly contributed to the yield of enzymically active protein, but the magnitude of its contribution is unknown. The differences between *E.coli* strains RS101 and RS51 also included different temperatures of gene induction and variation in the RBS and spacer sequences, either of which can effect the expression of protein in an active, soluble form (Stanssens et al. 1985, White and Richardson 1988).

## **CHAPTER 5**

### **BACTERIOPHAGE SP6 RNA POLYMERASE**

**-ISOLATING THE PHAGE SP6 RNA POLYMERASE GENE BY THE  
POLYMERASE CHAIN REACTION**

**-EXPRESSION OF THE PCR-DERIVED SP6 RNA POLYMERASE GENE IN  
*E.COLI***

## 5.1 INTRODUCTION

Bacteriophage SP6 has a double stranded DNA genome of about 43,500 bp that encodes an RNA polymerase which is highly specific for SP6 promoters (Chamberlin et al. 1970). The SP6 promoters are not recognised by other RNA polymerases. The protein has an Mr of 96,000 and is a single polypeptide (Butler and Chamberlin 1982).

SP6 RNA polymerase can be obtained from SP6-infected *S.typhimurium*. However, unlike the large scale purification of T7 RNA polymerase from T7-infected *E.coli*, the yields of SP6 RNA polymerase are reasonable, and the enzyme is relatively stable. Nonetheless, cloning the SP6 RNA polymerase gene allows for a significantly greater yield of SP6 RNA polymerase/gram of bacteria and also avoids having to prepare large SP6 phage stocks.

The SP6 promoter was used to construct the first in a series of highly efficient transcription vectors, pSP64/5 (Melton et al. 1984, Krieg and Melton 1984). The commercial potential of SP6 RNA polymerase, and the need for large amounts of the enzyme for RNA transcriptions, provided the impetus for cloning the SP6 RNA polymerase gene into *E.coli*.

Prior to the cloning of the SP6 RNA polymerase gene by Kotani et al.(1987) the locus for the gene was unknown, though a restriction map of the SP6 genome was available (Kassavetis et al.1982). However considerable evidence indicated that SP6 genomic organization was similar to that of phage T7. Transcription of SP6 DNA with SP6 RNA polymerase was found to be from one strand of the genome and resulted in discreet transcripts, indicating multiple SP6 promoters (Kassavetis et al.1982).

Given that both SP6 and T7 phage are similar in having phage-encoded, phage promoter-specific RNA polymerases, our initial attempt at locating the SP6 RNA polymerase gene assumed the possibility that the gene was located near an SP6 promoter, which could complicate cloning the gene, but which could be a useful marker.

Essentially, the strategy involved probing SP6 DNA restriction fragments with a 5'-<sup>32</sup>P-labelled 25-mer oligonucleotide which matches the DNA sequence of the SP6 promoter in the transcription vector pSP64. After hybridizing the oligonucleotide probe to the SP6 DNA restriction fragments, the probe was washed from the filter in the presence of tetramethyl ammonium chloride (Wood et al. 1985). This negated the dissimilar binding kinetics between

A-T and G-C pairs, allowing for sensitive mapping of homologous sequences. The probe identified a 1.94 kbp region of the SP6 genome (39.45 kbp to 41.39 kbp) which bound the probe, indicating that if the SP6 promoters have a conserved sequence, they are not identical. The existence of multiple conserved-sequence SP6 promoters was later demonstrated by Brown et al.(1986).

The report by Kotani et al.(1987) of a cloned SP6 RNA polymerase gene ended our need to find the gene locus. The development of the polymerase chain reaction made the isolation, cloning and expression of SP6 RNA polymerase in *E.coli* a relatively trivial procedure.

## **5.2 METHODS and RESULTS**

### **5.2.1 Isolation of the SP6 RNA polymerase gene by PCR**

#### **5.2.1.1 Design of the PCR oligonucleotide primers**

The PCR oligonucleotide primers were designed from the DNA sequence data of Kotani et al. (1987). The PCR primers were designed to amplify a region of DNA that includes the RNA polymerase gene, its RBS and some DNA 3' to the stop codon (see figure 5.1).

#### **5.2.1.2 PCR amplification of phage SP6 DNA**

Five micrograms of SP6 genomic DNA was digested in medium salt buffer with 8 units of Bgl II, in a 20 µl reaction volume for 90 minutes at 37°C. Following this, 60 µl of water was added to the DNA solution, bringing the DNA concentration to 60 ng/µl.

The 50 µl PCR mixture contained 60 ng of the Bgl II-digested SP6 DNA, 10 mM Tris-Cl, pH 8.4, 50 mM KCl, 1.5 mM MgCl<sub>2</sub>, 2 µg of BSA, 200 µM of each dNTP, 1 µg of each 5'-phosphorylated 24-mer oligonucleotide primer and 4 units of Cetus Amplitaq polymerase. The reaction mixture was incubated in a Cetus Thermal Cycler with denaturation for 30 seconds at 94°C, annealing for 1 minute at 55°C and extension at 72°C for 5 minutes, through 30 cycles.

**Figure 5.** Design of PCR primers for extracting the SP6 RNA polymerase gene from SP6 genomic DNA

A: 5' - GCGCTCAATTAAGTTTTCTAGTACCGCAT RBS GAGGATACAAGATGCAA' ·····  
f-Met Gln

B: 5' - GTACCGCATGAGGATACAAGATGC - 3'

A: 5'-terminus of SP6 RNA polymerase gene (Kotani et al. 1987)

B: RRS-782 24-mer oligonucleotide PCR primer

Asp Ser Glu Tyr Val Phe Ala STOP

C: ······ GGATTCTGAATACGTATTTGCCTAATAGAACAA' ······ 3'

|||||

D: 3' - AAGACTTATGCATAAACGGATTAT -5'

C: 3'-terminus of SP6 RNA polymerase gene (Kotani et al. 1987)

D: RRS-783 24-mer oligonucleotide PCR primer

### 5.2.2 Cloning the PCR-derived DNA into *E.coli*

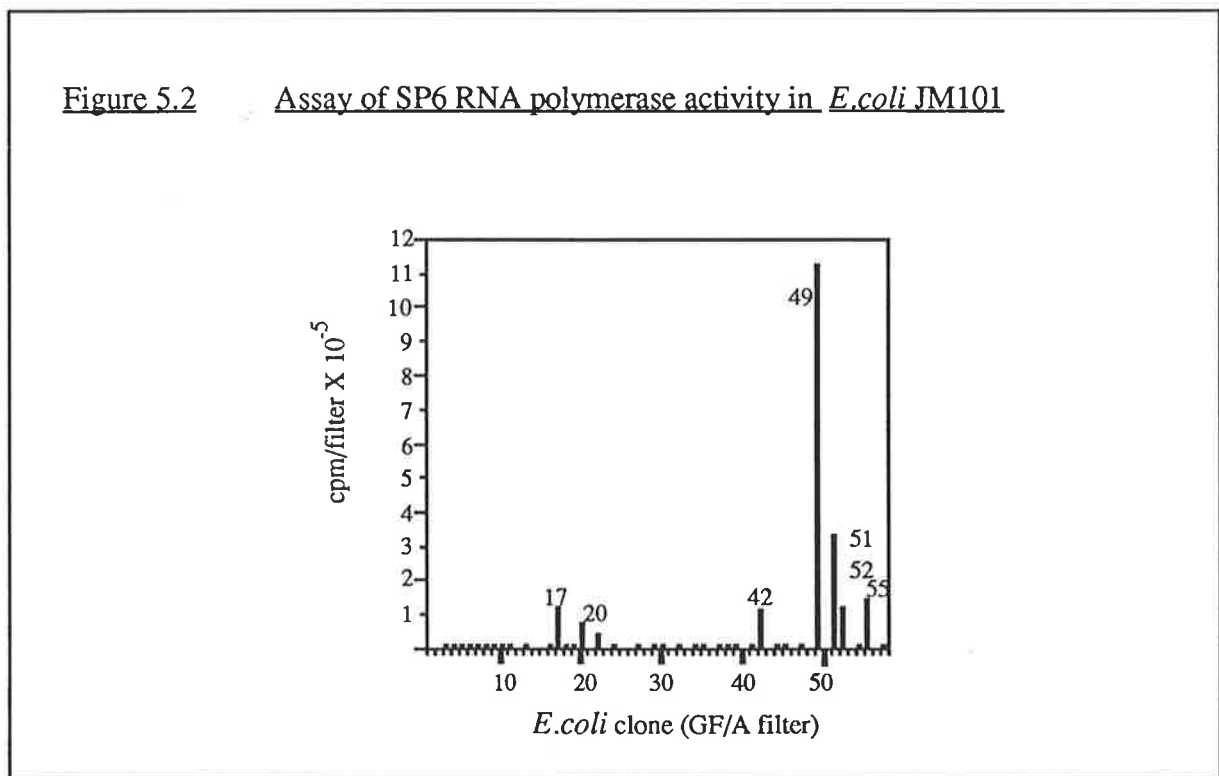
The PCR mixture from Exp. 5.2.1.2 was electrophoresed through a 1% agarose-TAE gel and stained with 10 µg/ml ethidium bromide. Visualisation of the gel under long wavelength UV light showed a single band which co-migrated with a 2.7 kbp DNA marker (data not shown). The DNA was purified from the agarose with Bio-101 GeneClean and dissolved in 20 µl of water. A 10 µl sample of the DNA was ligated to 100 ng of Sma1-cut, dephosphorylated pEDV-1 in a 20 µl solution containing 50 mM Tris-Cl, pH 7.6, 10 mM MgCl<sub>2</sub>, 10 mM DTT, 5% PEG 8000 and 7 units of T4 DNA ligase. The reaction proceeded overnight at 25°C. The ligation mixture was used to transform 200 µl of competent *E.coli* JM101. The transformed bacteria were grown overnight at 37°C on L-plates containing 100 µg/µl ampicillin, yielding 104 colonies. There were no background colonies.

### 5.2.3 Assay for enzymically active SP6 RNA polymerase *E.coli* clones

Sixty of the bacterial colonies were toothpicked into 3 ml of L-broth containing 100 µg/ml ampicillin and incubated at 37°C until the A<sub>600</sub> reached 1.0, and were then induced with IPTG

to a final concentration of 1 mM. A 1 ml sample of each induced culture was sonicated for 3 x 10 seconds and spun in an Eppendorf centrifuge for 15 minutes at 5°C. A 1 µl aliquot of supernatant was assayed in a 50 µl solution containing 40 mM Tris-Cl, pH 8.0, 8 mM MgCl<sub>2</sub>, 5 mM β-mercaptoethanol, 1 mM KH<sub>2</sub>PO<sub>4</sub>, 0.4 mM of each NTP, 10 µCi (3.3 pmoles) α-<sup>32</sup>P-UTP and 1 µg SP6 DNA. The reaction proceeded for 10 minutes at 37°C and was stopped by spotting 40 µl onto a Whatman GF/A filter and washing in 3 x 100 ml of 1 M HCl /1% Na<sub>4</sub>P<sub>2</sub>O<sub>7</sub> /1% KH<sub>2</sub>PO<sub>4</sub> at 4°C. The filters were washed briefly in ethanol, dried and measured by liquid scintillation counting.

Eight of the bacterial cultures expressed active SP6 RNA polymerase and the most active of these, designated *E.coli* RS49, was used in the following experiments (see figure 5.2).



#### 5.2.4 Protein sequencing the SP6 RNA polymerase from *E.coli* RS49

*E.coli* RS49 was used by Bresatec for the commercial preparation of SP6 RNA polymerase, essentially as described by Butler and Chamberlin (1981) except the SP6 RNA polymerase



gene was induced by adding IPTG to the culture, to a final concentration of 0.4 mM, after the culture had reached an  $A_{600}$  of 0.8. The bacteria were then pelleted by centrifugation and ruptured by two passes through a French pressure cell. The *E. coli* RS49-derived SP6 RNA polymerase used in this and the following experiment was donated by Bresatec.

Approximately 20  $\mu$ g of *E. coli* RS49-derived SP6 RNA polymerase was electrophoresed on a polyacrylamide gel, and then transferred electrophoretically to a polyvinylidene difluoride membrane, stained with Coomassie blue R and then loaded onto an Applied Biosystems protein sequencer. The method is described in detail in Exp. 3.3.11.5. The purification step was necessary to remove the BSA component in the enzyme storage buffer.

Protein sequencing unambiguously revealed the first nineteen amino acids of the purified SP6 RNA polymerase, which exactly matched the sequence predicted by Kotani et al. (1987) from their DNA sequence data.

#### **5.2.5 Qualitative SP6 RNA polymerase transcription assay**

*E. coli* RS49-derived SP6 RNA polymerase was compared to SP6 RNA polymerase identically purified from SP6-infected *S. typhimurium*, to verify that RNA transcripts from the cloned enzyme showed no aberrations.

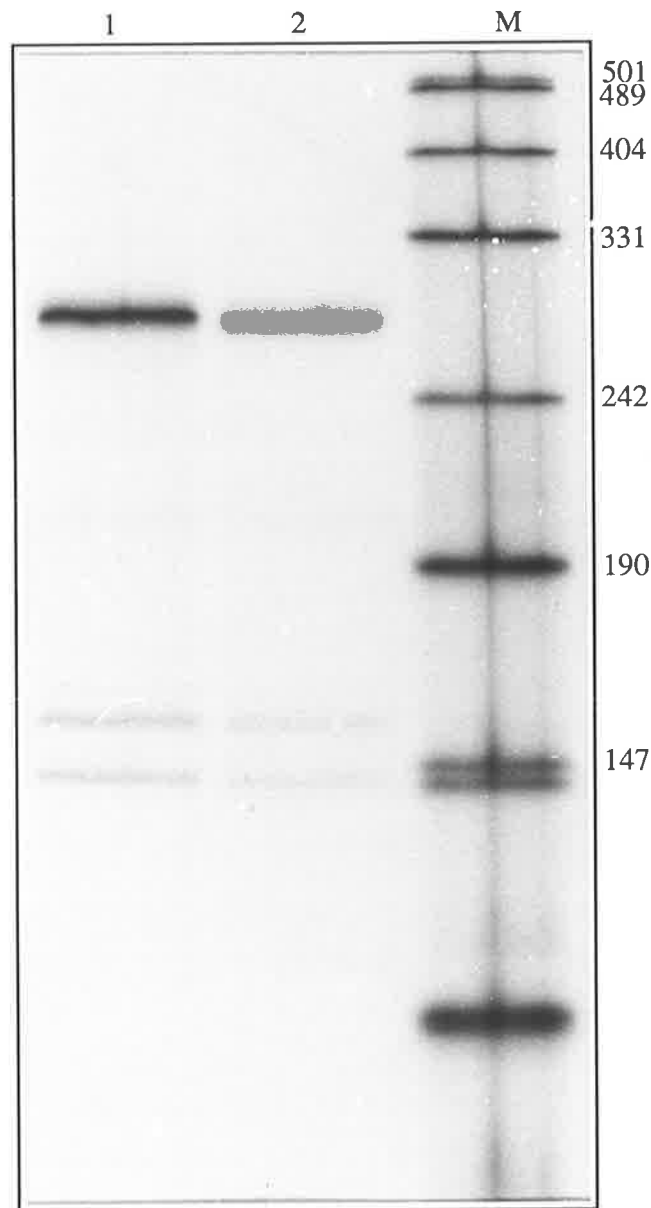
The 20  $\mu$ l transcription reactions contained 40 mM Tris-Cl, pH 7.6, 6 mM  $MgCl_2$ , 0.5 mM ATP, 0.5 mM GTP, 0.5 mM CTP, 150  $\mu$ M UTP, 17 pmoles  $\alpha$ - $^{32}P$ -UTP (3,000 Ci/mM), 2  $\mu$ g BSA, 10 mM DTT, 1  $\mu$ g transcription template (see below) and 4 units SP6 RNA polymerase.

One unit of enzyme is defined as the amount required to catalyse the incorporation of UTP into an acid-insoluble form at the rate of 1 nmole/hour in a 10 minute reaction at 37°C (Butler and Chamberlin 1982). The assay template, pSH-1, is an SP6/T7 transcription vector (pSP72) that harbours a 406 bp DNA fragment. Digestion of pSH-1 with *Sac*I, followed by transcription with SP6 RNA polymerase results in the generation of a 262 nucleotide major transcript.

The transcription products were analysed by electrophoresis through a 15%

Figure 5.3

Comparison of SP6 RNA polymerase RNA transcripts:  
Cloned enzyme vs SP6-infected *S.typhimurium*-derived enzyme



The assay template, pSH-1, contains a 406 bp insert next to an SP6 promoter. The vector was linearized with *Sac*I. Transcription with SP6 RNA polymerase generates a 262 nt major transcript.

1: SP6-infected *S.typhimurium*-derived enzyme  
2: Enzyme from *E.coli* RS49

M: End-filled *Hpa*II-digested pUC18 markers

polyacrylamide, 8 M urea gel, followed by autoradiography (see figure 5.3). The *E.coli* RS49-derived SP6 RNA polymerase produced identical transcripts to those directed from pSH-1 by the SP6 RNA polymerase purified from SP6-infected *S.typhimurium*.

### 5.3 DISCUSSION

Isolation of the SP6 RNA polymerase gene from genomic DNA by PCR was successful and posed few technical difficulties. The selection of an *E.coli* clone harbouring the SP6 RNA polymerase gene was made on the basis of a functional assay. Consequently, it was possible to screen the bacterial transformants directly and quickly, so in the event of a very high mutation rate, it would still be possible to identify a positive clone from a large background of undesirable bacterial colonies.

An interesting result from Exp.5.2.3 is that clone number 51 showed 30% of the SP6 RNA polymerase activity as clone number 49 (*E.coli* RS49). The SP6 RNA polymerase gene in the latter clone has not been DNA sequenced, but the enzyme isolated from the *E.coli* JM101 bearing this PCR-derived gene was purified to homogeneity, and was found to direct the synthesis of identical RNA transcripts from an SP6 promoter as did the SP6 RNA polymerase purified from phage SP6-infected *Salmonella typhimurium*. Whether the SP6 RNA polymerase expressed by clone number 51 is produced at a lower level than in *E.coli* RS49 or is partially inactive has yet to be determined.

## GENERAL DISCUSSION

This work describes the gene cloning and expression, in *E.coli*, of the RNA polymerases from bacteriophage T7 and SP6.

The major problem complicating the cloning of the full-length RNA polymerase genes from both T7 and SP6 was known to be the deleterious effect to *E.coli* of being transformed with plasmids encoding the RNA polymerase gene and also the phage-specific RNA polymerase promoter (Davanloo et al.1984, Kotani et al.1987). Expression of the T7 and SP6 RNA polymerases from the cloned genes firstly involved the isolation of DNA fragments from which the phage-specific promoters had been removed. The approaches used by Davanloo et al.(1984) and Kotani et al.(1987) toward that end were very different.

Davanloo et al.(1984) took advantage of the complete genomic nucleotide sequence and genetic mapping data of phage T7 in producing a heteroduplex DNA species from T7 double-deletion strains. Their heteroduplex DNA encoded the T7 RNA polymerase gene and its RBS but lacked the T7 promoter  $\phi$ 1.1A.

By contrast, the SP6 RNA polymerase gene locus was unknown. Butler and Chamberlin (1982) had previously determined the apparent molecular weight of SP6 RNA polymerase and in doing so predicted the approximate size of the SP6 RNA polymerase gene. Furthermore, Kassavetis et al.(1982) had demonstrated various similarities of SP6 RNA polymerase to the RNA polymerase from phage T7. In particular, Kassavetis et al.(1982) demonstrated that transcription of SP6 DNA with SP6 RNA polymerase resulted in discreet transcripts toward the right terminus of the genome, suggesting the presence of multiple SP6 promoters. SP6 RNA polymerase was also found to be highly specific for the SP6 promoter. Together, these findings indicated that both T7 RNA polymerase and SP6 RNA polymerase were specific for their respective promoters and transcribed their genomes from promoters which are spaced to control the expression of late function genes.

Kotani et al.(1987) DNA-sequenced the SP6 genome until they identified an open reading frame long enough to encode the SP6 RNA polymerase protein that was also near the left terminus of the genome. Having determined the SP6 RNA polymerase gene locus, Kotani

et al.(1987) cloned an SP6 DNA fragment which encoded the RNA polymerase gene into *E.coli*. They were unable to identify bacterial colonies harbouring the SP6 RNA polymerase gene. By reasoning that an SP6 promoter was located on the DNA fragment they had attempted to clone into the bacteria, they digested the DNA fragment with Bal 31 exonuclease to produce a range of DNA fragments. Some of these were successfully cloned into *E.coli*. Reducing the size of the fragment which was lethal to *E.coli* potentially enabled the isolation of a full-length SP6 RNA polymerase gene without the SP6 promoter. Transformation of *E.coli* with the Bal 31 nuclease-digested DNA was expected to produce viable clones harbouring either a partial-length SP6 RNA polymerase gene with or without the SP6 promoter, or a full-length SP6 RNA polymerase gene without the SP6 promoter. One of their *E.coli* clones expressed enzymically active SP6 RNA polymerase.

The difficulties for other laboratories in cloning the T7 and SP6 RNA polymerase genes or any gene first cloned by the methods used by Davanloo et al.(1984) and Kotani et al.(1987) are not completely overcome simply because the gene loci and DNA sequence is known.

Our initial approach to cloning the T7 RNA polymerase gene used combination of imprecise and highly specific techniques. A T7 DNA restriction fragment encoding the RNA polymerase gene and the T7 promoter was digested with Bal 31 nuclease and a partial-length T7 RNA polymerase gene was isolated and cloned into *E.coli*. Reconstruction of the gene to a full-length T7 RNA polymerase gene was accomplished by exploiting the site-directed mutagenesis technique developed by Zoller and Smith (1983). Our second approach to cloning the T7 RNA polymerase gene used the PCR to extract the gene from T7 DNA. Both approaches to cloning the T7 RNA polymerase gene were successful.

The PCR approach involved only one step to isolate the T7 RNA polymerase gene away from the genome, whereas the first approach required considerable manipulation of T7 DNA before a full-length T7 RNA polymerase gene was generated. However the first approach, in using site-directed mutagenesis to repair a partial-length gene, used *E.coli* to repair the deletion. The main limitation of our first approach was determined by the size of the DNA fragment that could be cloned into the M13 RF DNA. This problem could have been overcome by initially cloning the partial-length T7 RNA polymerase gene into a plasmid that

also contained an f1 origin of replication. This would have enabled the isolation of single-stranded DNA for the mutagenesis reaction from the double-stranded plasmid by superinfection of the bacterial host with helper phage. Any limitations to the use of single-stranded DNA for site-directed mutagenesis could also be overcome to some extent by exploiting restriction site-independent techniques for site-directed mutagenesis of double-stranded plasmid DNA.

We also used the PCR to isolate the SP6 RNA polymerase gene, which was subsequently cloned into *E.coli* and found to express enzymically active SP6 RNA polymerase. A significant limitation of the PCR is the error rate of misincorporated nucleotides caused by the use of Taq DNA polymerase *in vitro* (Tindall and Kunkell 1988). However if a functional assay for a protein is available, the PCR-induced mutations may not be a fundamental problem as non-mutated proteins may still be identified from a background of PCR mutants.

## REFERENCES

- Auble, D.T., Allen, T.L. and deHaseth, P.L.: Promoter recognition by *Escherichia coli* RNA polymerase. *J. Bio. Chem.* **261** (1986) 11202-11206.
- Bernard, H.-U., Remaut, E., Hershfield, M.V., Das, H.K., Helinski, D.R., Yanofsky, C. and Franklin, N.: Construction of plasmid cloning vehicles that promote gene expression from the bacteriophage lambda P<sub>L</sub> promoter. *Gene* **5** (1979) 59-76.
- Birnboim, H.C. and Doly, J.: A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nuc. Acids Res.* **7** (1979) 1513-1523.
- Brosius, J., Erfle, M. and Storella, J.: Spacing of the -10 and -35 regions in the *tac* promoter. *J. Biol. Chem.* **260** (1985) 3539-3541.
- Brown, J.E., Klement, J.F. and McAllister, W.T.: Sequences of three promoters for the bacteriophage SP6 RNA polymerase. *Nuc. Acids Res.* **14** (1986) 3521-3526.
- Brumby, A.: A vector for high expression of native proteins. (1986) Honours Thesis, University of Adelaide.
- Brunner, M. and Bujard, H.: Promoter recognition and promoter strength in the *Escherichia coli* system. *EMBO Journal* **6** (1987) 3139-3144.
- Butler, E.T. and Chamberlin, M.J.: Bacteriophage SP6-specific RNA polymerase. *J. Biol. Chem.* **257** (1982) 5772-5778.
- Casadaban, M.J. and Cohen, S.N.: Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. *J. Mol. Biol.* **138** (1980) 179-207.
- Chamberlin, M.J., McGrath, J. and Waskell, L.: New RNA polymerase from *Escherichia coli* infected with bacteriophage T7. *Nature* **228** (1970) 227-231.
- Das, A. and Wolska, K.: Transcription antitermination *in vitro* by lambda N gene product: requirement for a phage nut site and the products of host nusA, nusB, and nusE genes. *Cell* **38** (1984) 165-173.
- Davanloo, P., Rosenberg, A.H., Dunn, J.J. and Studier, F.W.: Cloning and expression of the gene for bacteriophage T7 RNA polymerase. *Proc. Natl. Acad. Sci. USA* **81** (1984) 2035-2039.
- DeBoer, H.A., Hui, A., Comstock, L.J., Wong, E. and Vasser, M.: Portable Shine-Dalgarno regions: a system for a systematic study of defined alterations of nucleotide sequences within *E. coli* ribosome binding sites. *DNA* **2** (1983) 231-235.

- Deuschle, U., Kammerer, W., Gentz, R. and Bujard, H.: Promoters of *Escherichia coli*: a hierarchy of *in vivo* strength indicates alternate structures. *EMBO Journal* **5** (1986) 2987-2994.
- Dunn, J.J. and Studier, F.W.: Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *J. Mol. Biol.* **166** (1983) 477-535.
- George, H.J., L'Italien, J.J., Pilacinski, W.P., Glassman, D.L. and Krzyzek, R.A.: High-level expression in *Escherichia coli* of biologically active bovine growth hormone. *DNA* **4** (1985) 273-281.
- Gonzalez, N., Wiggs, J. and Chamberlin, M.J.: A simple procedure for resolution of *Escherichia coli* RNA polymerase holoenzyme from core polymerase. *Arch. Biochem. Biophys.* **182** (1977) 404-408.
- Gren, E.J.: Recognition of messenger RNA during translational initiation in *Escherichia coli*. *Biochemie* **66** (1984) 1-29.
- Grodberg, J. and Dunn, J.J.: *ompT* encodes the *Escherichia coli* outer membrane protease that cleaves T7 RNA polymerase during purification. *J. Bacteriol.* **170** (1988) 1245-1253.
- Grunstein, M. and Hogness, D.S.: Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. *Proc. Natl. Acad. Sci. USA* **72** (1975) 3961-3965.
- Guan, K., Haun, R.S., Watson, S.J., Geahlen, R.L. and Dixon, J.E.: Cloning and expression of a protein-tyrosine-phosphatase. *Proc. Natl. Acad. Sci. USA* **87** (1990) 1501-1505.
- Hanahan, D.: Techniques for transformation of *E. coli*. In Glover, D.M. (ed.) *DNA Cloning*, Vol. 1. IRL Press, Oxford, U.K., 1985, pp. 109-135.
- Hasan, N. and Szybalski, W.: Control of cloned gene expression by promoter inversion *in vivo*: construction of improved vectors with a multiple cloning site and the  $p_{tac}$  promoter. *Gene* **56** (1987) 145-151.
- Hizi, A., McGill, C. and Hughes, S.H.: Expression of soluble, enzymatically active, human immunodeficiency virus reverse transcriptase in *Escherichia coli* and analysis of mutants. *Proc. Natl. Acad. Sci. USA* **85** (1988) 1218-1222.
- Ho, Y., Lewis, M. and Rosenberg, M.: Purification and properties of a transcriptional activator. *J. Biol. Chem.* **257** (1982) 9128-9134.
- Holland, L.J. and Wamg, L.J.: Efficient recovery of functionally intact mRNA from agarose gels via transfer to an ion-exchange membrane. *Nuc. Acids Res.* **11** (1983) 3283-3300.
- Kassevetis, G.A., Butler, E.T., Roulland, D. and Chamberlin, M.J.: Bacteriophage SP6-specific RNA polymerase. *J. Biol. Chem.* **257** (1982) 5779-5788.



- Kotani, H., Ishizaki, Y., Hiraoka, N. and Obayashi, A.: Nucleotide sequence and expression of the cloned gene of bacteriophage SP6 RNA polymerase. *Nuc. Acids Res.* **15** (1987) 2653-2664.
- Krieg, P.A. and Melton, D.A.: Functional messenger RNAs are produced by SP6 *in vitro* transcription of cloned cDNAs. *Nuc. Acids Res.* **12** (1984) 7057-7070.
- Krieg, P.A. and Melton, D.A.: *In vitro* RNA synthesis with SP6 RNA polymerase. *Methods in Enzymology* **155** (1987) 397-415.
- Laemmli, U.K.: Cleavage of structural proteins during the assembly of the head of phage T4. *Nature* **227** (1970) 680-685.
- Lin, T., Rush, J., Spicer, E.K. and Konisberg, W.H.: Cloning and expression of T4 DNA polymerase. *Proc. Natl. Acad. Sci. USA* **84** (1987) 7000-7004.
- Mandecki, W. and Reznikoff, W.S.: A lac promoter with a changed distance between -10 and -35 regions. *Nuc. Acids Res.* **10** (1982) 903-912.
- Maniatis, T., Fritsch, E.F. and Sambrook, J. *Molecular Cloning: A Laboratory Manual.* (1982) Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Matsudaira, P.: Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. *J. Biol. Chem.* **262** (1988) 10035-10038.
- Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. and Green, M.R.: Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nuc. Acids Res.* **12** (1984) 7035-7056.
- Messing, J.: A multipurpose cloning system based on single-stranded DNA bacteriophage M13. *Recombinant DNA Technical Bulletin* **2** (1979) 43-48.
- Midgley, C.A. and Murray, N.E.: T4 polynucleotide kinase; cloning of the gene (*pseT*) and amplification of its product. *EMBO Journal* **4** (1985) 2695-2703.
- Moffatt, B., Dunn, J.J. and Studier, F.W.: Nucleotide sequence of the gene for bacteriophage T7 RNA polymerase. *J. Mol. Biol.* **173** (1984) 265-269.
- Morris, C.E., Klement, J.F. and McAllister, W.T.: Cloning and expression of the bacteriophage T3 RNA polymerase gene. *Gene* **41** (1986) 193-200.
- Mulligan, M.E., Brosius, J. and McClure, W.R.: Characterization *in vitro* of the effect of spacer length on the activity of *Escherichia coli* RNA polymerase at the TAC promoter. *J. Biol. Chem.* **260** (1985) 3529-3538.

- Ner, S.S., Goodin, D.B., Pielak, G.J. and Smith, M.: A rapid droplet method for Sanger dideoxy sequencing. *Biotechniques* **6** (1988) 408-412.
- Pham, T.T. and Coleman, J.E.: Cloning, expression and purification of gene 3 endonuclease from bacteriophage T7. *Biochemistry* **24** (1985) 5672-5677.
- Plotch, S.J., Palant, O. and Gluzman, Y.: Purification and properties of poliovirus RNA polymerase expressed in *Escherichia coli*. *J. Virology* **63** (1989) 216-225.
- Remaut, E., Stanssens, P. and Fiers, W.: Plasmid vectors for high-efficiency expression controlled by the  $p_L$  promoter of coliphage lambda. *Gene* **15** (1981) 81-93.
- Saiki, R.K.: The design and optimization of the PCR. in Erlich, H.A. (ed.) *PCR Technology: Principles and Applications for DNA Amplification*. Stockton Press, N.Y., 1989, 7-16.
- Saiki, R.K., Bugawan, T.L., Horn, G.T., Mullis, K.B. and Erlich, H.A.: Analysis of enzymatically amplified  $\beta$ -globin and HLA-DQ $\alpha$  DNA with allele-specific oligonucleotide probes. *Nature* **324** (1986) 163-166.
- Sanger, F. and Coulson, A.R.: The use of thin acrylamide gels for DNA sequencing. *FEBS Lett.* **87** (1978) 107-110.
- Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H. and Roe, B.A.: Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. *J. Mol. Biol.* **143** (1980) 161-178.
- Sanger, F., Nicklen, S. and Coulsen, A.R.: DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74** (1977) 5463-5467.
- Scharf, S.J., Horn, G.T. and Erlich, H.A.: Direct cloning and sequence analysis of enzymatically amplified genomic sequences. *Science* **233** (1986) 1076-1078.
- Scherer, G.F.E., Walkinshaw, M.D., Arnott, S. and More, D.J.: The ribosome binding sites recognized by *E. coli* ribosomes have regions with signal character in both the leader and protein coding segments. *Nuc. Acids Res.* **8** (1980) 3895-3907.
- Schoner, R.G., Ellis, L.F. and Schoner, B.E.: Isolation and purification of protein granules from *Escherichia coli* overproducing bovine growth hormone. *Biotechnology* **3** (1985) 151-154.
- Serghini, M.A., Ritzenthaler, C. and Pinck, L.: A rapid and efficient 'mini-prep' for isolation of plasmid DNA. *Nuc. Acids Res.* **17** (1989) 3604.
- Shamoo, Y., Adari, H., Konigsberg, W.H., Williams, K.R. and Chase, J.W.: Cloning of T4 gene 32 and expression of the wild-type protein under  $\lambda$  promoter  $P_L$  regulation in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **83** (1986) 8844-8848.

- Shine, J. and Dalgarno, L.: The 3'-terminal sequence of *E. coli* 16S RNA: complementarity to nonsense triplets and ribosome binding sites. *Proc. Natl. Acad. Sci. USA* **71** (1974) 1342-1346.
- Stanssens, P., Remaut, E. and Fiers, W.: Alterations upstream from the Shine-Dalgarno region and their effect on bacterial gene expression. *Gene* **36** (1985) 211-223.
- Studier, F.W. *Science* **176** Bacteriophage T7. (1972) 367-376.
- Studier, F.W. and Moffatt, B.A.: Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189** (1986) 113-130.
- Tabor, S. and Richardson, C.C.: A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. USA* **82** (1985) 1074-1078.
- Talmadge, K. and Gilbert, W.: Construction of plasmid vectors with unique PstI cloning sites in a signal sequence coding region. *Gene* **12** (1980) 235-241.
- Tindall, K.R. and Kunkel, T.A.: Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Biochemistry* **27** (1988) 6008-6013.
- Vogelstein, B. and Gillespie, D.: Preparative and analytical purification of DNA from agarose. *Proc. Natl. Acad. Sci. USA* **76** (1979) 615-619.
- Wallace, R.B., Schold, M., Johnson, M.J., Dembek, P. and Itakura, K.: Oligonucleotide directed mutagenesis of the human  $\beta$ -globin gene : A general method for producing specific point mutations in cloned DNA. *Nuc. Acids Res.* **9** (1981) 3647-3656.
- White, J.H. and Richardson, C.C.: Gene 19 of Bacteriophage T7. *J. Biol. Chem.* **263** (1988) 2469-2476.
- Wood, W.I., Gitschier, J., Lasky, L.A. and Lawn, R.M.: Base composition-independent hybridization in tetramethylammonium chloride: a method for oligonucleotide screening of highly complex gene libraries. *Proc. Natl. Acad. Sci. USA* **82** (1985) 1585-1588.
- Yanisch-Perron, C., Vieira, J. and Messing, J.: Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33** (1985) 103-109.
- Zoller, M.J. and Smith, M.: Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any fragment of DNA. *Nuc. Acids Res.* **20** (1982) 6487-6500.
- Zoller, M.J. and Smith, M.: Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors. *Methods in Enzymology* **100** (1983) 468-500.