A MECHANISM FOR TRANSCRIPTIONAL INTERFERENCE BETWEEN CONVERGENT PROMOTERS IN THE DEVELOPMENTAL SWITCH OF BACTERIOPHAGE 186

This thesis is submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in the School of Molecular and Biomedical Sciences (Biochemistry), University of Adelaide

Benjamin Peter Callen, B. Sc. (Hons)

March 2003

TABLE OF CONTENTS

THESIS SUMMARY

Convergent transcription defines a situation where two promoters are arranged face-to-face, leading to a partial 5' overlap between their transcripts. Multiple examples of convergent promoters have been reported in a wide range of organisms (including bacteriophage, *E. coli*, plants, yeast and humans). In both eukaryotes and prokaryotes the simultaneous *in vivo* activity of such promoters generally leads to repression of transcription from the opposing promoter, a phenomenon termed transcriptional interference. Interference is often asymmetric, with a strong (the aggressive) promoter reducing the expression of a weak (the sensitive) promoter. This interaction has been used in various ways to effect regulation of gene expression.

A number of studies have investigated this phenomenon in both *E. coli* and eukaryotic systems however the exact mechanism(s) remain speculative. Theoretically, a wide range of molecular mechanisms of interference are possible. This study uses largely *in vivo* methods to eliminate some of these possibilities and sets out to further understand the mechanism operating for one set of converging promoters.

The convergent promoters used in this study are those of the lysis-lysogeny switch from bacteriophage 186. The bacteriophage 186 lytic promoter, *pR*, is ten times stronger than the lysogenic promoter, *pL*, found 62 bp downstream in a convergent orientation. In chapter 2 a single copy, promoter, reporter system was established to measure convergent promoter activity. It was shown that the stronger promoter reduced the transcriptional activity of the weaker promoter 5.6 fold, and that this interference is not reciprocal. No variation in interference was seen when different host strains were used.

In Chapter 3, this promoter system was used to determine what action of *pR* causes interference of *pL*. A minor role for *pR* bound RNAP during interference was demonstrated by the following observations: (i) Increasing the spacing between the promoters an extra 100 bp brought no loss in interference but rather a slight increase. This is inconsistent with a steric hindrance model of interference where RNAP bound at *pR* inhibits RNAP binding at *pL*. (ii) Placing the promoters in a close divergent orientation completely abolished interference, which is contrary to a competitive

inhibition model of interference. And (iii) The placement of a unidirectional intrinsic terminator between *pR* and *pL* significantly reduced interference from 9.2 fold to 3.3 fold. This is contrary to a mechanism of interference where RNAP bound at *pR* acts as a roadblock.

A major role of elongation from *pR* over *pL* during interference was shown by the observations that (i) Divergent promoters that also actively transcribe the 62 nt of 5' antisense transcript did not restore interference, contrary to a model of interference that involves antisense transcription. And (ii) that terminating elongation from *pR* prior to transcription over *pL* dramatically reduced interference.

The nature of interference caused by elongation over *pL* was then investigated in Chapter 4. Based on the *in vitro* observation that open complexes formed at *pL* were slow to escape and clear the promoter and that the activity of these complexes was reduced by collisions with converging elongating polymerase from *pR*. A 'sitting duck' mechanism of interference is proposed in which *pR* convergent transcription over *pL* negatively affects promoter initiation intermediates that form at *pL*.

In Chapter 5 the validity of this mechanism was supported by an investigation of transcriptional interference between another example of convergent promoters found in the developmental switch of the related bacteriophage, P2. The strong lytic promoter was shown to reduce the activity of the weak lysogenic promoter only 2.2 fold (contrary to the 30 fold interference found in the literature). The difference in interference with that of 186 promoters was shown to be partly due to the differences in sensitivity of the two lysogenic promoters. The *in vitro* rate of escape of open complexes formed at the P2 lysogenic promoter was shown to be much faster than that of 186 *pL*. The lack of interference for P2 promoters could therefore be explained by a reduced potential of its lysogenic promoter to form 'sitting ducks'. Thus, interference in the 186 system occurs because RNAP complexes at *pL* that are waiting to clear are sensitive to passing elongating polymerase from *pR*. This 'sitting duck' mechanism is likely to be important for promoters that are close together. Mechanisms involving collisions between elongating polymerases and promoter occlusion and a general mechanism for transcriptional interference by convergent promoters discussed in Chapter 6.

DECLARATION

This work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge, contains no material previously published or written by another person, except where due reference is made in the text.

I give consent to this copy of my thesis, when deposited in the University Library, being available for loan and photocopying.

Benjamin Peter Callen
March 2003

ACKNOWLEDGEMENTS

Sincere thanks to:

The discipline of Biochemistry it's academics and very important other staff, although it spent much of its time during the course of this thesis 'morphing', it thankfully maintained the freedom to allow quality research and teaching of fundamental principles of biology without the need of a fashionable, immediate economic outcome.

The grand-founding father of 186 research, Prof. J. Barry Egan. His supervision was always achieved with a greatly appreciated gentle undercurrent of learning, not just scientifically but fundamentally. He instilled an approach to life that questions whether you're creating tracks or simply running races, and maintained constant vigilance towards preventing studies of how the phases of the moon effect the molecular nature of a bacteriophage. Thankyou for all you've given, may you forever find yourself atop a genuine smelly tree at the end of the year!

To the Egan Lab: Keith, Ian, Rachel and Ali. To the untrained they may seem quiet and uninterested- HOWEVER there's not a better group of infinitely patient, 'eager' to help, enthusiastic for all the right reasons, gentle, enjoyable to provoke people. They are more friends and companions in phagy adventures than simply exceptionally helpful 'other people' in the lab.

To my many good friends especially fellow pre-doctors and old friends Duncan, Robyn and Tom. I'm constantly thankful to have such unique and essential company. Here's to long lunches and consoling conversation now and forever. Special mention must go out to Robyn for saving my sanity during the final moments.

To my extra family at Unley and Malvern Uniting Church, for its many activities and prayers that gave much needed peace, patience and strength. Especially my good mate Johnno for looking after me and being the type of person who always brings a resonating smile.

Finally, to my family both newly acquired (Ruth, Andrew, Mel and Jared) and since conception (Mum, Dad, Adam and Em), you can all now stop worrying, and breathe an emphatic breath of fresh air. Thank you for all sharing the weight with me in many different ways. I am very fond of you all, I hope I can return equally the love you've given me.

Most importantly to my very loving wife, Nicole Ruth you have been the most enjoyable portion of this thesis experience and I thank God greatly for gluing us together.

"From the moment I picked your book up until the moment I laid it down, I was convulsed with laughter. Someday I intend reading it." Graucho Marx

"I do not approve of anything that tampers with natural ignorance. Ignorance os like a delicate, exotic fruit, touch it and it is gone. The whole theory of modern education is radically unsound; fortunately in England at any rate, education produces no effect whatsoever. If it did it would have a serious effect on the upper classes and probably lead to violence in Grosvenor Square."
The Importance of Being Earnest, by Oscar Wilde.

So here's to one step towards the loss of ignorance...

CHAPTER ONE

INTRODUCTION

Chapter 1

Introduction

1.1 Convergent transcription and transcriptional interference.

The term 'convergent transcription' defines a situation where two promoters are arranged face-to-face, leading to a partial 5' overlap between their transcripts. The first description of convergent transcription was of an artificial example where the *his* and *rough b* operons of *Salmonella typhimurium* of opposing orientation were fused together by deletion (Levinthal and Nikaido, 1969). A symmetrical *in vivo* inhibition of expression from each operon was observed. The first natural example described was in the early control region of bacteriophage λ where two promoters P_R and P_{RE} were found to be positioned either side of the *cro* gene, directing convergent transcription (see fig 1.1) (Spiegelman *et al.*, 1972). Since then a large number of other examples both, natural and artificial, have been reported and studied. Several of these are described further in this chapter.

In both eukaryotes and prokaryotes the simultaneous *in vivo* activity of convergent promoters generally leads to repression of transcription from the opposing promoter (Elledge and Davis, 1989; Eszterhas *et al.*, 2002), a phenomenon termed transcriptional interference. Transcriptional interference is broadly defined as the perturbation of one transcription unit by another, which may include promoter interactions resulting from not only convergent transcription but also from divergent transcription or transcription of promoters arranged in tandem (Eszterhas *et al.*, 2002). In this study it specifically refers to repression at a transcriptional level caused by convergent transcription. This unusual method of repression is simple in design, requiring only a duplication and rearrangement of the same genetic component (a promoter). Interference is often asymmetric, with a strong (the aggressive) promoter reducing the expression of a weak (the sensitive) promoter. This is an effective mechanism of transcriptional repression and due to its simplicity of design, one might expect these promoter interactions to be used in biology as a basic method of regulating gene expression.

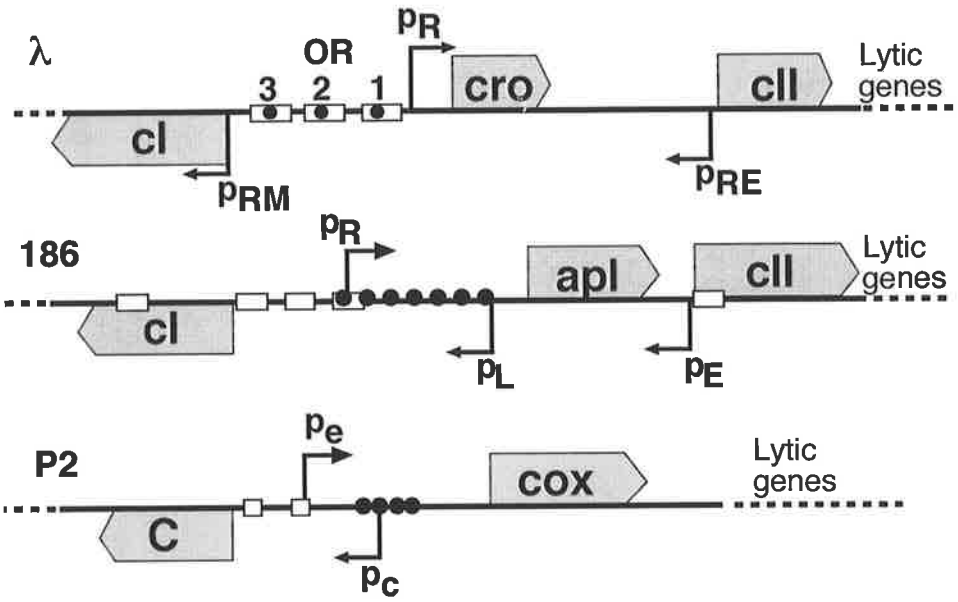


Figure 1.1. Lysis-lysogeny switches of l, 186, and P2.

CI and C are immunity repressors, and Cro, Apl, and Cox are anti-immunity repressors. The diagrams are not to scale, but the relative distances between the repressor and anti-immunity repressor open reading frames for the three switches are accurate. The locations of the binding sites for the immunity (rectangles) and anti-immunity (filled circles) repressors are indicated. P_{RE} and *pE* are the CII-activated establishment promoters of λ and 186, respectively. Adapted from Neufing *et al.* 2001.

To gain a better appreciation of the significance of interference by convergent transcription in nature requires an understanding of how interference can regulate gene expression (reviewed in section 1.2), an analysis of the frequency of this type of promoter arrangement (section 1.3) and finally knowledge of the mechanism of interference. The aim of this project is to investigate the mechanism of interference operating in the natural examples of convergent transcription found in the developmental switches of bacteriophage 186 and P2 (fig 1.1). Current theories and previous mechanistic studies are reviewed in section 1.4.

1.2 Gene regulation by convergent transcription.

The precise regulation of gene expression is essential for an organism's ability to change, be it a cellular response to environmental conditions, or a developmental choice such as sex determination or the bacteriophage lysis or lysogeny decision. Consequently, an extremely diverse array of mechanisms have evolved to control gene expression. The interference caused by a face-to-face arrangement of aggressive and sensitive promoters has been shown to be the underlying mechanism for examples of: (1) low expression of constitutive genes, (2) serendipitous reduction in gene expression caused by the genomic insertion of promoters, and (3) controlled regulation of gene expression. These three types of gene regulation are discussed further using examples from the literature.

1.2.1. Constitutive gene expression

Transcriptional interference can be a general mechanism for low expression of constitutive genes by expressing the gene from a sensitive promoter that is convergent to an uncontrolled aggressive promoter. An example is the very inefficient expression of the *transposase* gene of the insertion sequence, IS10 of *Escherichia coli*, which contains convergent promoters *p*IN and *p*OUT separated by 36bp (Simons *et al.*, 1983). *p*IN is a weak promoter which transcribes the *transposase* gene (responsible for IS10 transposition) and *p*OUT is a strong promoter (4.4 fold stronger in *lacZ* reporter studies than *p*IN) directing transcription toward the nearby IS10 termini. The effects of promoter interference due to convergent transcription, examined *in vivo* by measuring β

-galactosidase reporter gene activity, showed that p_{IN} was repressed to nearly undetectable levels (an 18 fold reduction) when assayed in the face of active p_{OUT}. Interference was shown to act at both a transcriptional and translational level (Simons *et al.*, 1983; Simons and Kleckner, 1983). Transcription from p_{OUT} is effectively unregulated and its interference with p_{IN} ensures a very low constitutive level of transposase activity. Convergent transcription of *transposase* genes has also been found in other transposable elements including: *Pot2* from the rice blast fungus (Kimura and Yamaguchi, 1998), F elements of *Drosophilia melanogaster* (Contursi *et al.*, 1993), and the *Dictyostelium* repetitive element from *Dictostelium discoideum* (Shumann *et al.*, 1994).

1.2.2. Serendipitous reduction of gene expression

Transcriptional interference by convergent transcription can result in the serendipitous reduction in gene expression caused by natural or artificial insertions of promoters into the genome. For example genomic insertion of transgenes next to host promoters or between multiple adjacent transgenes can have the unfavourable outcome of reducing gene expression by convergent transcription. The potential for this reduction in gene expression has been shown in transgenic plants (Ingelbrecht *et al.*, 1991; Padidam and Cao, 2001), and in cultured mouse cells (Eszterhas *et al.*, 2002). This phenomenon has been implicated as one of the contributing sources of variation in expression frequently observed among independent transformants.

Transposable elements and retroviruses often contain very strong promoters at their genetic boundaries which, upon insertion into a host chromosome, could provide a natural example of directing interfering convergent transcription into the neighbouring genome. For example, the 5'-untranslated region of human L1 retrotransposon has an antisense promoter driving transcription into adjacent cellular sequences and yielding chimeric transcripts (Nigumann *et al.*, 2002). Insertion of this retrotransposon adjacent to a genomic promoter in an opposing orientation could lead to repression of that promoter by transcriptional interference as a result of convergent transcription (Nigumann *et al.*, 2002). It has been proposed that this type of repression could influence the expression of a very large number of mammalian genes, leading to

phenotypic variation among mammals (Whitelaw and Martin, 2001). Transposable B2 short interspersed elements (SINEs) are highly abundant components of mammalian genomes that have been propagated by retrotransposition. These elements have been shown to provide mobile RNA polymerase II promoters as well as RNA polymerase III promoters (both promoters are convergent to each other within the same element), which can regulate the expression of neighbouring genes (Ferrigno *et al.*, 2001). Analysis of the mouse transcriptome showed that among cDNAs with an annotated coding sequence (CDS), 14.2% carry one or more repeats (mostly repeats of the SINE (short interspersed nucleotide element), simple repeats, LINEs (long interspersed nucleotide elements) and LTR (long terminal repeats) classes) that overlap with the CDS (FANTOM Consortium, 2002). The potential for gene regulation by transcriptional interference is thus very large. Outward promoters capable of convergent transcription with genomic promoters have been found in other transposable elements including the pOUT promoters of IS10 (Simons *et al.*, 1983) and IS3 (Charlier *et al.*, 1982) in *E. coli* and the F elements of *D. melanogaster* (Minchiotti *et al.*, 1994).

1.2.3. Controlled regulation of gene expression

Transcriptional interference can be used as an indirect method of regulation of gene expression from a sensitive promoter by having the aggressive promoter controlled by some regulatory factors such as a transcription factor, methylation sensitivity or supercoiling. If both promoters transcribe different genes then regulation of two different transcriptional states is possible. Alternatively, if only one promoter transcribes a functional gene and transcripts produced from the other promoter are both non-coding and non-functional, then regulation of only one functional, transcriptional state is possible. The potential for this alternative type of regulation is very poorly recognised in the literature despite the evidence that many examples of this type of convergent transcription exists (see section 1.3).

Transcriptional activity from the aggressive promoter may serve only to reduce the activity of the sensitive promoter, with only the sensitive promoter responsible for transcription of a functional gene. An example is *E. coli* regulation of the *fis* operon (Nasser *et al.*, 2002). Transcription of the operon is controlled by a cluster of five promoters within a stretch of 85 bp; four of these are arranged in tandem directing

transcription of the *fis* operon, and the fifth *Pdiv* is located between the two middle promoters in a convergent orientation and does not direct transcription of any known gene. The consequences of the many different promoter interactions are complex, but it has been shown *in vivo* and *in vitro* that the convergent promoter interferes with expression of the promoter located 7 bp downstream (the other downstream promoter was not considered due to its low level of transcription). The activity of the convergent promoter is regulated by the binding of IHF, which then changes the balance of individual promoter interactions, and thus the overall level of *fis* operon expression (Nasser *et al.*, 2002). Transcription of the *fis* operon is also regulated by supercoiling and this may also be a factor influencing transcription from the convergent promoter, and thus the balance of promoter interactions.

The potential for this type of regulation in eukaryotes is largely ignored, however a number of well studied regulatory systems could be adequately explained by transcriptional interference such as the control of mammalian X-chromosome inactivation. In mammals, either of the two X-chromosomes in females is inactivated to compensate for dosage. Silencing involves the accumulation of Xist RNA from one X-chromosome. Xist action is repressed by transcription of an antisense transcript, Tsix, which originates 12kb downstream of Xist and traverses beyond the Xist promoter region in mice. This is an example of convergent transcription. Both transcripts are non-coding, but whereas the Xist transcript has been shown to be responsible for chromosome silencing, a functional role for the Tsix transcript is still unclear although splicing variants have been found (Shibata and Lee, 2003). The repression of Xist accumulation by Tsix is proposed to be caused by either the activity of the antisense RNA product or antisense transcriptional movement (transcriptional interference). However, the cis-limited action of Tsix would suggest some additional role for the overlapping nature of convergent transcription. In the development of extraembryonic cells control of Tsix expression is by differential maternal and paternal imprinting of a Tsix CpG-rich region, probably by methylation. There are a number of examples of gene expression controlled by differential imprinting, many of which involve cis-limited interfering transcription of a non-coding antisense RNA, for example the Igf2r/Air locus in mammals (reviewed in (Rougeulle and Heard, 2002)). The potential for transcriptional interference to explain these regulatory phenomenon in eukaryotes is

poorly recognised in the literature, as much of the debate about how these antisense promoters cause decreased expression has focussed on the potential activity of the antisense transcript rather than the act of transcription itself. However the orientation of the promoters and the cis-limited activity of repression in these examples is consistent with a mechanism of transcriptional interference to account for the observed repression.

If transcription from the aggressive promoter also encodes a functional gene then interference may be used to ensure that the transcription of two opposing functions is coordinated. Three different examples from bacterial plasmids are presented. The colicinogenic plasmid pColA in bacteria uses convergent transcription to express genes for two opposing functions. These genes code for the antibiotic protein colicin A (Caa), and the immunity protein (Cai) (Lloubes *et al.*, 1986). Expression from the *caa* promoter is very strong and transcriptionally interferes with expression of the weak *cai* promoter. Interference is normally absent due to the presence of a host protein, LexA, which binds to the aggressive *caa* promoter and represses its activity. This causes immunity in the absence of strong colicin A activity. Under conditions which inactivate LexA, such as DNA damage, interference is induced causing colicin A production in the absence of immunity protein. A similar situation is seen in the switch between transcription of genes for the opposing functions of vegetative replication and conjugative transfer of the promiscuous plasmid RK2 (Jagura-Burdzy and Thomas, 1997). As with pColA, genes for the two opposing genetic functions are also transcribed from convergent promoters, however transcriptional interference is normally present rather than absent and caused by the aggressive promoter responsible for vegetative replication. Repression of the aggressive promoter is by proteins encoded elsewhere on the plasmid. Another example of this type of regulation is in the control region for plasmid copy number of pIP501(Brantl and Wagner, 1997). Here the two opposing functions are the transcription of genes that increase or decrease plasmid replication. The aggressive promoter (pII) transcribes *repR*, a gene required for replication, the sensitive promoter (pIII) transcribes an RNA that works as an attenuator of *repR* to decrease its expression. Control of pII expression is by a repressor protein Co*pR*, encoded elsewhere on the plasmid. Decreasing Co*pR* concentration triggers an increase in replication by relieving the repression of pII transcription and also increasing the translation of *repR* due to transcriptional interference with pIII activity. Note that the

replicon of pIP501 is a member of the IncFII-type plasmid family which all have a similar type of replication control mechanism involving convergent transcription (Brantl, 2002). Convergent transcription is also found in the replicons of ColE1-type plasmids (Brantl, 2002).

If, during this type of regulation, transcription from the sensitive promoter encodes a repressor of the aggressive promoter then an additional level of transcriptional control is possible by positive autoregulation of the sensitive promoter. An example of this is the lysis-lysogeny developmental switch of the non-lambdoid family of bacteriophage including phage Mu (van Rijn *et al.*, 1989), P2 (Saha *et al.*, 1987a), HP1 (Esposito *et al.*, 1997) and 186 (Dodd and Egan, 2002)). Convergent transcription in the switches of 186 and P2 is the subject of this study.

1.2.4. Convergent transcription as part of the genetic switch for temperate bacteriophage development.

Coliphages 186 and P2 are temperate bacteriophages that infect *E. coli*. Upon phage infection, a developmental decision is made between two life cycles: lysogenic or lytic development. These life cycles are mutually exclusive yet interchangeable. Lysogeny involves the maintenance of a stable prophage, a state which is established by the integration of the circular phage DNA into the host genome. A prophage can be induced resulting in excision of phage DNA and entry into the opposing lytic phase. Lytic development involves the replication of the phage genome combined with expression of lytic genes to eventually produce many phage particles. Finally the host cell is lysed, releasing into the extracellular environment mature phage which can subsequently infect surrounding (non-lysogenic) cells to begin the cycle again. The use of transcriptional interference between convergent promoters in the development of bacteriophage 186 and P2 is presented below.

i. The role of transcriptional interference in the development of phage 186

The convergent promoters in the developmental switch of 186 are the lytic promoter, *pR*, responsible for early lytic transcription, and the lysogenic promoter, *pL*, which transcribes all the genes necessary for establishing and maintaining lysogeny. The

starting points of transcription from each convergent promoter are located 62 bp apart (Dodd *et al.*, 1990) (Fig. 1.1). Reporter studies have shown that the activity of *pR* is intrinsically about 10-fold stronger than *pL* and that convergent transcription from the aggressive promoter, *pR,* inhibits transcription from the sensitive promoter, *pL,* some 6-20 fold (Dodd *et al.*, 1990; Reed *et al.*, 1997; Neufing *et al.*, 2001; and Dodd and Egan, 2002). Interference is important during lytic phage development to allow early lytic transcription from *pR* to occur in the absence of significant lysogenic transcription from *pL.* Entry into lysogeny is by expression of sufficient levels of the lysogenic repressor, CI, which strongly represses *pR* (Dodd *et al.*, 1990). However, *pL* activity alone can not establish lysogeny (Neufing, 1997) due to the continual interference with *pL* activity by *pR* preventing the expression of sufficient levels of CI. This ensures that the decision to enter lysis or lysogeny is not dependent on the chance transcription of *pL* but rather the activation of an alternative promoter for CI transcription, pE, found upstream of *pL.* Activation of pE is by the CII, protein produced from the *pR* transcript. pE activity produces a transient burst of CI expression sufficient to begin repressing *pR.* The subsequent relief of transcriptional interference from *pR* allows for positive autoregulation of *pL* activity in the absence of further pE activity. This leads to the maintenance of CI expression from *pL* activity alone, thus establishing lysogeny (Dodd and Egan, 2002). This method of activating the lysogenic promoter is in contrast to that of lambdoid phage which have divergent lytic and lysogenic promoters (eg. λ) and require the lysogenic maintenance promoter P_{RM} to be directly activated by binding of λ CI at the promoter (see fig 1.1) (Ptashne, 1992). The use of interference in 186 development is similar to that of previous examples (section 1.1.3.) of convergent promoters which ensure that transcription of two opposing genetic states is coordinated. However, since the sensitive promoter, *pL,* transcribes a repressor (CI) of the aggressive promoter, *pR,* an additional level of control is possible during the establishment of lysogeny as a result of interference.

ii. Convergent transcription in phage P2.

The convergent promoters of the developmental switch in P2 are the lytic promoter, *pe* and the lysogenic promoter, *pc* located about 30 nucleotides apart (fig 1.1). The lysogenic promoter *pc* controls synthesis of lysogenic proteins including C, a repressor

of the lytic promoter, whereas transcription from *pe* results in the synthesis of early lytic genes including Cox, a repressor of the lysogenic promoter. Regulation of promoter activity by convergent transcription was shown by plasmid based CAT promoter reporter assays (Saha *et al.*, 1987b). The intrinsic activity of *pc* was 5 fold less than *pe*, but 139 fold less when measured in the presence of converging *pe* transcription. Transcriptional interference is expected to be involved in regulating transcription of the genetic switch in a similar manner to 186. Specifically, during lytic development interference ensures that lytic transcription from *pe* occurs in the absence of significant transcription from *pc*, and secondly, the relief of interference via C protein mediated repression of *pe* ensures positive autoregulation of the lysogenic promoter *pc*, once the decision to proceed into lysogeny has been made. Additionally, because P2 does not have a CII/*pE* system of establishing lysogeny, the decision is instead thought to be made by the chance expression of *pc* over *pe* (Saha *et al.*, 1987b). Interference must therefore play an additional role in determining the frequency of lysogeny along with the intrinsic activities of the convergent promoters.

1.3 The frequency of convergent transcription in nature.

There are a number of known examples of convergent transcription in both eukaryotes and prokaryotes, some of which have been mentioned in section 1.2. The advent of large scale genome sequencing has revealed that transcriptional interference by convergent transcription may be responsible for a genome wide level of gene regulation that is additional to factor dependent regulation. In fact, it has been suggested that the use of promoter interactions to regulate gene expression could have co-evolved with factor dependent regulation, or alternatively there was a primordial RNA polymerase-dependent homeostatic regulation and a superimposed control by transcription factors has evolved during the course of evolution (Nasser *et al.*, 2002). Searching the literature for published examples of convergent transcription is difficult due to the varying terms used to describe convergent transcription. These include overlapping divergent promoters, antisense transcription, opposing promoters and promoters of divergent operons. Due to the lack of adequately defined promoter sequences in some organisms, it has also been difficult to accurately predict the presence of face-to-face promoters and

convergent transcription by *in silico* analysis. To date most genome analysis has focussed on the prediction of open reading frames rather than their promoters.

Examples of convergent transcription and its use in gene regulation appear frequently in the genomes of extrachomosomal elements such as bacteriophage (lambdoid and non-lambdoid), retroviruses , transposable elements, insertion sequences and plasmids (see section 1.2). It is suspected that the highly compact nature of these genomes facilitates the evolution of convergent transcription to co-regulate gene expression.

For the larger chromosomal genomes, I have classified convergent transcription into two types (represented in fig 1.2): 'Class I' face-to-face promoters, which both produce coding mRNA, and 'Class II' face-to-face promoters, of which only one produces a coding mRNA, whilst the other produces non-coding RNA. The first type can be located within a genome sequence from the locations of promoter sequences responsible for the transcription of adjacent but oppositely oriented coding sequences. These coding sequences will be divergent (eg. fig 1.2 Class I: type A or B) or overlapping (eg fig 1.2 Class I: type C or D). A review in 1988 referenced at least four examples in *E. coli* of face-to-face promoters used to express divergent operons (Beck and Warren, 1988). These include (i) the promoters for the DNA replication genes *dnaQ* and *rnh* where *dnaQ* is transcribed from two promoters, P1 and P2 which are separated from the convergent promoter *rnhp* by 107 and 24 bp respectively (Nomura *et al.*, 1985a), and (ii) promoters for genes of the arginine regulon (Cunin *et al.*, 1983). At least one other example has since been reported, being the promoters of the *mgtA* and *treR* genes (Yamamoto *et al.*, 2002). An analysis of sequence information from the *E. coli* genome project suggests many more examples are yet to be described. I have analysed a list of 4641 predicted *E. coli* promoters available from the RegulonDB database (Salgado *et al.*, 2001), for examples of face-to-face promoters. 173 pairs of promoters (7.4% of the total number of predicted promoters) were face-to-face and placed less than 200 bp apart, while 435 promoter pairs (or 18.7%) were spaced less than 2kb apart. Not all of these pairs are expected to direct convergent transcription, as some may not be expressed and others may contain termination sequences between them. Unfortunately, the methods employed to predict the promoters have not yet been published, so it is difficult to assess the accuracy and relevancy of the face-to-face promoters, or whether

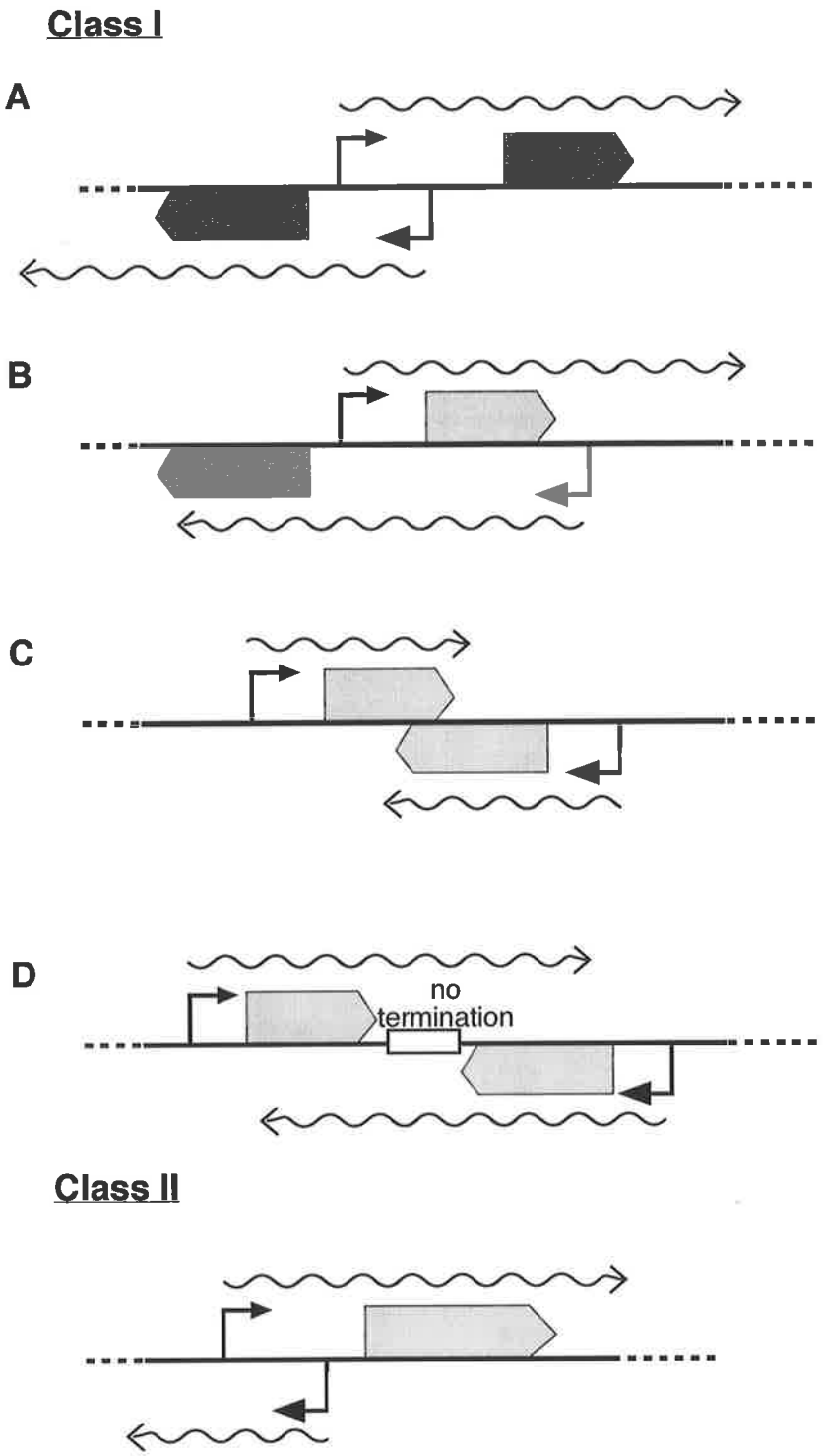


Figure 1.2. Types of convergent transcription.

Bent arrows indicate the start and direction of transcription from a promoter, and grey boxes indicate open reading frames. To simplify the diagram, examples of convergent transcription involving different arrangements of exons and introns are not included. Refer to section 1.3 for more details.

they are expected to direct transcription of divergent or overlapping open reading frames without further detailed analysis, but it would appear that many more are yet to be discovered.

The genomic sequence of the lower eukaryote *Saccharomyces cerevisiae* is highly compact and gene rich, leaving little space for noncoding DNA. Adjacent genes could direct convergent transcription as shown in fig 1.2. Genomic analysis has revealed a disproportionately large number of adjacently located genes are transcribed away from each other on opposite strands (Kruglyak and Tang, 2000). This suggests that there appears to be evolutionary pressure to select against Class I type B, C and D convergent transcription (fig 1.2) possibly to avoid interference by detrimental transcriptional collisions (Prescott and Proudfoot, 2002). On average divergent open reading frames are only 618 nt apart (Dujon, 1996), of these the promoters responsible for their transcription may be either back-to-back or face-to-face (eg. fig 1.2 class I type A). A number of these have been found to exhibit correlated expression patterns, which is indicative of back-to-back promoters sharing common regulatory mechanisms (Kruglyak and Tang, 2000). Face-to-face promoters that direct transcriptional interference would be expected to exhibit a reverse correlation of expression. A detailed analysis of the exact orientation of promoters responsible for transcription of close, adjacent divergently orientated ORFs in yeast is yet to be performed.

A few examples of closely divergent genes have been found in the genomes of higher eukaryotes some of which have been found to involve antisense transcription (ie. they are transcribed from face-to-face promoters) (Jain, 1996). The increase in intron size in higher eukaryotes allows for 'nested genes', that is genes that reside within the intronic regions of other genes (Portin, 2002). Convergent transcription is expected when the nested gene is on the opposite strand of the DNA. A number of examples of these types of genes have been reported in both *D. melanogaster* and humans (Portin, 2002).

The second type of convergent transcription which results in only a single gene product, is expected to be common as a result of transcription from repetitive elements and transposons placed near a host promoter. Such elements are highly repeated and dispersed throughout the genomes of many organisms. Examples are described in the

section 1.2.2. An additional example is the developmental control of expression of the human *ε-globin* gene which appears to involve convergent transcription from an Alu repetitive element that contains an active PolIII promoter (Wu *et al.*, 1990).

Convergent transcription that does not involve transcription of two functional genes is more difficult to predict from sequence analysis alone. The traditional concept of a gene involving transcription of large open reading frames would lead to questions about the relevance of predicting promoter locations that can not be attributed to transcription of an open reading frame. As a result the potential for convergent transcription involving non-coding RNA would be underestimated by a whole genome approach. Discovery of examples of convergent transcription would thus be limited to individual studies of transcription for particular genes. One example of this type of convergent transcription in *E. coli* is the cluster of promoters responsible for regulation of the *fis* operon (Nasser *et al.*, 2002).

Currently there is speculation that an alternative form of genetic regulation may be based on the abundant non-coding RNA transcripts (Eddy, 2002), particulary in eukaryotes (Mattick, 2001). A major family of these types of transcripts are "cis-antisense" noncoding RNAs which are transcribed from the opposite strand of protein-coding genes and are thus the result of convergent transcription. Although not widely recognised, many of these transcripts may not possess any regulatory functions themselves but could merely be the consequence of regulation by transcriptional interference from the activity of face-to-face promoters. Evidence for a large scale, genome wide potential of this type of convergent transcription is presented by an analysis of the transcripts produced from a genome, known as the 'transcriptome'. An analysis of the *E. coli* transcriptome grown under 13 different conditions has been performed using high density oligonucleotide probe arrays, which interrogated the sense strand of coding sequences (which is not expected to reveal convergent transcription) and both strands in the intergenic regions of the genome (Tjaden *et al.*, 2002). Using conservative criteria, a set of 1102 transcripts were identified in the intergenic regions of *E. coli*, of these 334 were classified as unidentified in that they were longer than 70 bp but could not be classed as operon elements or belonging to a 5' or 3' untranslated region of a downstream or upstream gene. Many of these transcripts could in fact be

antisense non-coding RNAs produced from convergent transcription. Comprehensive analysis of this data in combination with other evidence, such as locations of predicted promoters, would greatly advance speculation about the frequency of convergent transcription in *E. coli*. However this is outside the scope of this thesis.

The presence of large scale non-coding RNA in eukaryotes is suggested from recent data on mice and humans. (Kapranov *et al.*, 2002) identified active areas of transcription in human chromosomes 21 and 22, using oligonucleotide arrays with unique probes spaced along the entire chromosome. They found transcription was occurring at an order of magnitude more than can be accounted for by the predicted and characterised exons. This can be interpreted to mean that transcription is originating from many more locations than previously thought, many of these transcripts are suggested to be non-coding and could be the result of convergent transcription. An analysis of full-length mouse complementary DNA sequences (cDNAs) identified 33,409 transcriptional units (defined as a segment of DNA from which transcripts are generated) many of which were non-coding (Consortium. *et al.*, 2002). Moreover, 2,431 pairs of transcripts were found to be sense-antisense, overlapping in the exons of the sense gene by at least 20 bases. This indicates that a significant amount of convergent transcription must occur in the mouse genome, but whether this is associated with regulation of gene expression by interference has not been determined. Analysis of cDNAs limits detection of transcripts to cytoplasmic RNAs and will exclude other non-coding RNAs that are restricted to the nucleus. Antisense RNAs produced by convergent transcription from promoters that transcriptionally interfere may not be processed or transported to the cytoplasm and would therefore not be detected by analysis of cDNA.

In conclusion, a significant number of convergent promoters that use interference to regulate gene expression have been reported and continue to be reported in a variety of different organisms. Moreover, interference by convergent transcription as a common mechanism of gene expression is suggested by i) the widespread occurrence of repetitive elements and transposons that can contain active promoters, ii) analysis of the positions of predicted promoters from available genome sequences, and iii) large scale transcriptome analysis in *E. coli*, mice and humans.

1.4 The mechanism of interference.

The biological significance of interference will only be fully revealed once the molecular mechanisms of interference are known. Without this information it is difficult to predict the existence and extent of interference for particular examples of convergent transcription. Knowledge of the mechanism will provide an understanding of the intrinsic and extrinsic properties required for interference. A number of published studies have examined the mechanism operating in both eukaryotes and prokaryotes, using both natural and artificial examples. Findings from these studies are discussed initially in terms of what has been discovered about the properties of interference. The possible theoretical mechanisms of interference are then proposed and briefly explored, followed by details of studies that are relevant to particular mechanisms. Finally, the aims of this thesis are presented. To aid the discussion of how convergent transcription can lead to interference, a review of transcription from a single promoter is presented with reference to how transcription can be negatively regulated.

1.4.1. Transcription from a single promoter

To understand how convergent transcription might reduce the activity of a promoter it is important to be aware of the process of transcription and how it can be negatively regulated. As the work presented in this thesis examines interference in an *E. coli* system, this review is mainly concerned with *E. coli* transcription. For the purposes of discussing a more generalised mechanism of interference a brief presentation of transcription in eukaryotes is also presented.

Transcription initiation is a multistep process involving several classes of bimolecular complexes between promoter DNA and a RNA polymerase (RNAP) holoenzyme as illustrated in fig 1.3. A wealth of knowledge is available on this complex reaction incorporating many years of biochemical, genetic and structural studies, and a large number of reviews are available (eg. Hsu, 2002; Young *et al.*, 2002 and Record *et al.*, 1996). RNAP holoenzyme is defined as the 1:1 complex of the core polymerase (subunit composition, $\alpha_2\beta\beta'$) and the appropriate specificity (σ) subunit (the promoters used in this study are recognised by σ^{70}). Promoter DNA is the sequence from which

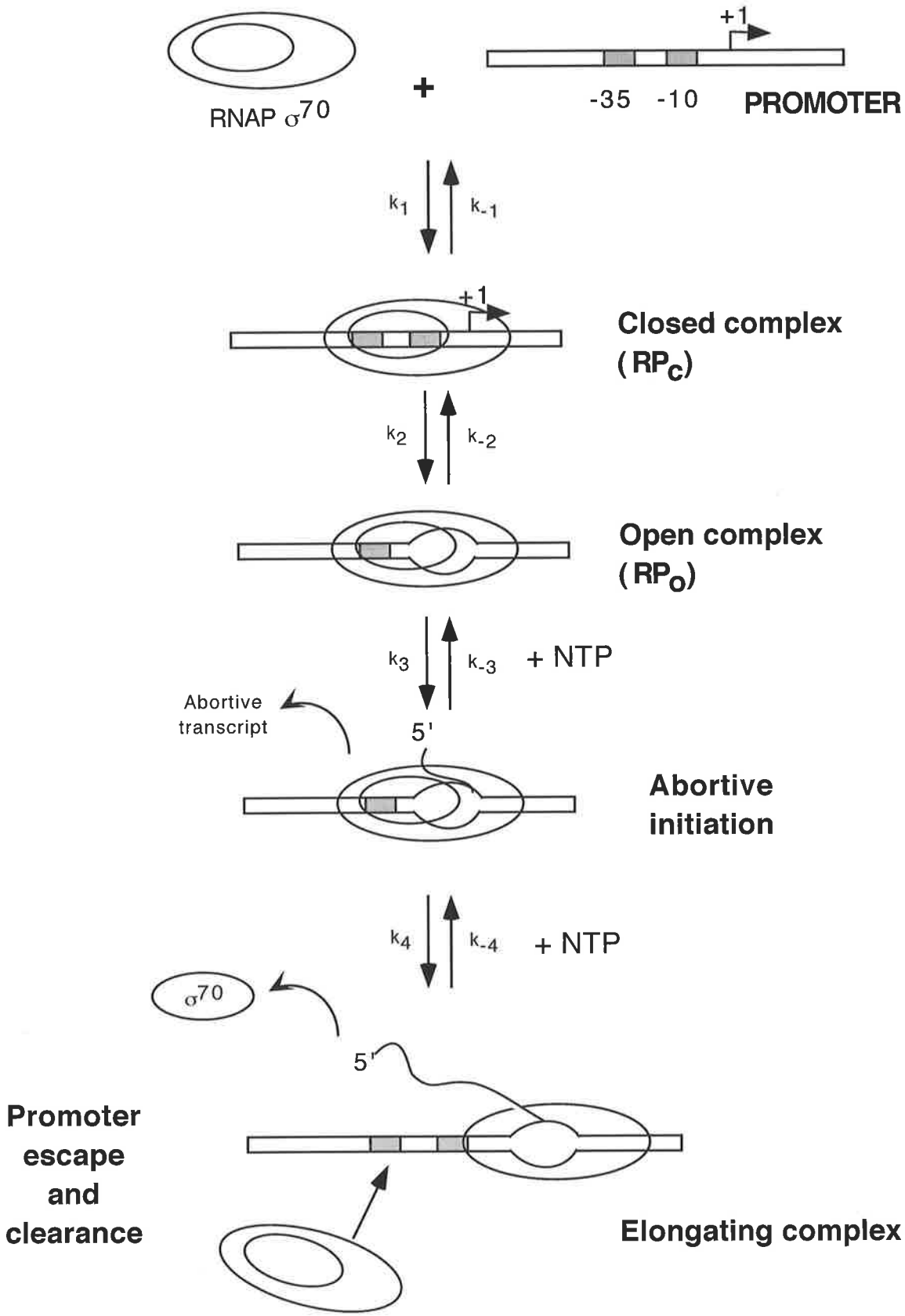


Figure 1.3. Events of initiation of transcription in *E. coli*.

Refer to section 1.4.1 for more details.

RNAP initiates transcription. Sequence analysis of ~300 naturally occurring *E. coli* σ^{70} promoters has defined a consensus promoter which contains three major control elements: two conserved 6-bp DNA sequences centered approximately 10 and 33 bp upstream from the transcription start site (+1), called the "-10" and "-35" hexamers (Hawley and McClure, 1983); and the length of DNA that separates them, called the "spacer" region, commonly ~17 bp in length.

On most *E. coli* RNAPσ^{70} promoters, RNAP holenzyme binds the promoter DNA to set up an active catalytic complex through two major conformational transitions. Upon first binding to the double-stranded promoter DNA, a complex lacking catalytic activity called the closed complex (RP_c) is formed. Contacts are made on one face of the helix between the -55 to +20 region. Kinetic investigations show that the closed complex formation is a rapid equilibrium governed by a binding constant K_B (i.e. k_1/k_{-1}). The next step involves the reversible unwinding ("opening") of a specific region of the promoter DNA from ~-10 to +2, forming a catalytically competent open complex (RP_O). Isomerisation to form the open complex can be described by forward and reverse constants k_2 and k_{-2}. On several promoters, kinetically significant intermediates within each major transition have been demonstrated (see Record *et al.*, 1996). Presented here is a simplified view of initiation. On many promoters the open complex is highly stable and its formation is essentially irreversible (i.e. k_{-2} is negligable).

Once the transcription start site is open, +1 complementary nucleoside triphosphate (NTP) binds to form the ternary initiated complex. Short RNA chains of up to 6-9 nucleotides (nt) in length are then synthesised and released by RNAP which is still bound at its upstream contacts. This process of RNAP idling or stuttering is called abortive initiation. Productive RNA chain synthesis will only occur once transcription has proceeded past position +7 to +12 and is characterised by the displacement of σ and loss of promoter contacts such that the amount of DNA occupied by RNAP is reduced from 70 bp to 40 bp. This process is called "promoter escape", and RNAP now forms a transcribing or elongating complex. Promoter escape may or may not involve sigma release (see Hsu, 2002). "Promoter clearance" will occur once the elongating complex has cleared the promoter region to allow for re-initiation to occur.

During elongation the moving polymerase protects about 30 bp along the DNA against nuclease digestion which includes a transiently open "transcription bubble" (~18 bp in length), flanked by double-stranded DNA. The process of elongation has been reviewed by Korzheva and Mustaev, (2001) and Erie, (2002). Elongation complexes (also called ternary complexes) are extremely resistant to dissociation with stabilising interactions thought to include those between RNAP and single and double stranded DNA and also the RNA-DNA hybrid (~9-12bp in length) (see the structure of elongating RNAP in fig 1.4). Elongation is a highly dynamic process with elongation complexes able to adopt a number of different conformational states. Elongation is highly regulated both by protein factors that bind to the DNA template, the RNA transcript, or the transcription complex as it moves along the template, and by specific sequence elements (expressed by DNA or RNA) that interact with the transcribing complex. During normal rapid elongation, RNAP exists primarily in a long lived activated state that is "cocked" and ready to catalyse nucleotide addition. At certain sites (such as pause and termination) or at positions where RNAP is halted by NTP deprivation or by physical roadblock, the elongation complex undergoes a conformational change to an 'unactivated state' or 'slow intermediate'. RNAP can catalyze RNA synthesis in this state, but at a much slower rate. Polymerase undergoing RNA synthesis exists in either a pre- or post-translocation state with or without NTPs bound. Once in the unactivated state, the ternary complex can enter states that induce cleavage of the transcript. In these states, RNAP can catalyze the hydrolysis of the RNA transcript, rapidly releasing the 3'-terminal fragment (which can be as large as 17 nt), and resuming synthesis. Alternatively RNAP can translocate backwards along the DNA template, extruding the 3' end of RNA, and eventually forming an arrested state. RNAP in this state cannot resume elongation without the action of an accessory protein GreB, which activates the cleavage activity of RNAP. Finally, at positions where a hairpin can form in the nascent transcript, complexes in the unactivated state can undergo a hairpin-induced transition to a hypertranslocated state, in which RNAP slips forward along the DNA template dislodging the 3' terminus from the catalytic site.

In recent years, the resolution of X-ray crystal structures of the RNAP core in *Thermus aquaticus* (Zhang *et al.*, 1999), the RNAP holoenzyme in *Thermus thermophilus* (Vassylyev *et al.*, 2002) and the yeast *S. cerevisiae* RNAP II (Cramer *et al.*, 2001; Gnatt

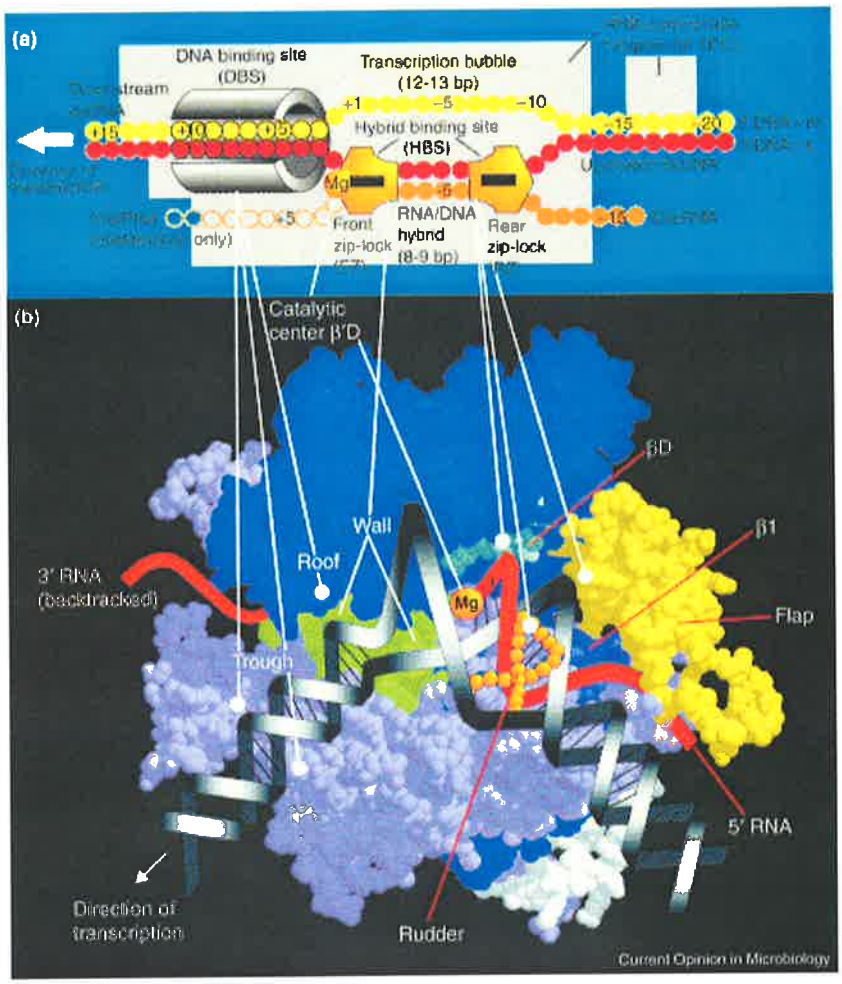


Figure. 1.4. Structure-function model of the transcription elongation complex (TEC).

(a) Structure of the nucleic acid scaffold of TEC based on contemporary knowledge of its configuration and interactions with RNAP protein. Five parts are distinguished: the upstream DNA duplex; the downstream DNA duplex; the 8-9 bp RNA-DNA hybrid; the single-stranded region of DNA in the transcription bubble; and the emergent ssRNA upstream of the hybrid. In the case of the 'backtracked' complex, an additional segment of ssRNA appears downstream of the bubble. Filled circles represent DNA (template strand in red; nontemplate strand in yellow). The DNA binding site is represented by a sliding clamp that encloses 9 bp of the downstream DNA duplex. The hybrid binding site that accommodates the RNA-DNA heteroduplex is represented by two zip-locks that hold onto the edges of the heteroduplex and either zip or unzip the hybrid, maintaining its constant size during lateral movement of RNAP. The gray area represents the RNAP footprint.

(b) Schematic model of TEC showing its main features. White lines show the correspondence between the functional and structural features of TEC.

Reproduced from figure 2, Korzheva and Mustaev, 2001.

et al., 2001) has provided structural models of core and holoenzyme, open complex with DNA and elongating bacterial RNAP (see fig 1.5) (reviewed in Korzheva and Mustaev, 2001; Young *et al.*, 2002; and Hsu, 2002). Bacterial core RNAP can crudely be described as a crab claw whose active site is positioned at the base of its two pincers. During elongation, downstream DNA is located in an internal channel formed between the pincers (or jaws) and then separates into its two strands near the active site. Template and non-template strands track different paths through the polymerase and reanneal to form upstream duplex, which is at a right angle to the downstream DNA. Nascent RNA follows the template strand for about 9 bases and then exits the polymerase underneath a flap that juts out from the bottom of the pincers (known as the RNA exit channel). The DNA kink is located near the catalytic site, where a chelated Mg^{2+} ion is positioned between the *i* and *i+1* nucleotide substrate binding sites. The *i* site is occupied by the 3'-OH nucleotide of the nascent transcript, while the *i+1* site accommodates the successive incoming NTP. Behind the active site Mg^{2+}, a secondary channel opens out to the rear surface of the enzyme and through this channel, NTP substrates can reach the active site. During abortive initiation the released short transcripts can diffuse through this channel.

In the holoenzyme, the three domains of σ are spread across one face of core. The process of promoter recognition involves these domains mediating binding to the different promoter elements. Sigma binding to core RNAP leads to large conformation changes and closing of the two pincers, to a distance that is too narrow to accommodate double-stranded DNA. The pincers must seemingly need to open and close again during open complex formation. Two of the σ domains are separated by a large distance which is traversed by a long polypeptide linker that occupies the entire length of the RNA exit channel through which a nascent transcript longer than 8 nt must also traverse. This positioning has lead to the speculation that during promoter clearance, nascent RNAs must successfully compete with this σ linker to be retained in elongating polymerase. When RNA transcripts lose the competition, they are ejected as abortive transcripts; when they win, the linker region is ejected and the transcript is successfully elongated. Release of the linker region could weaken the σ/core interface and cause promoter clearance. Open complex formation involves substantial conformational changes leading to a complete closure of the active site channel (the pincers). Within this

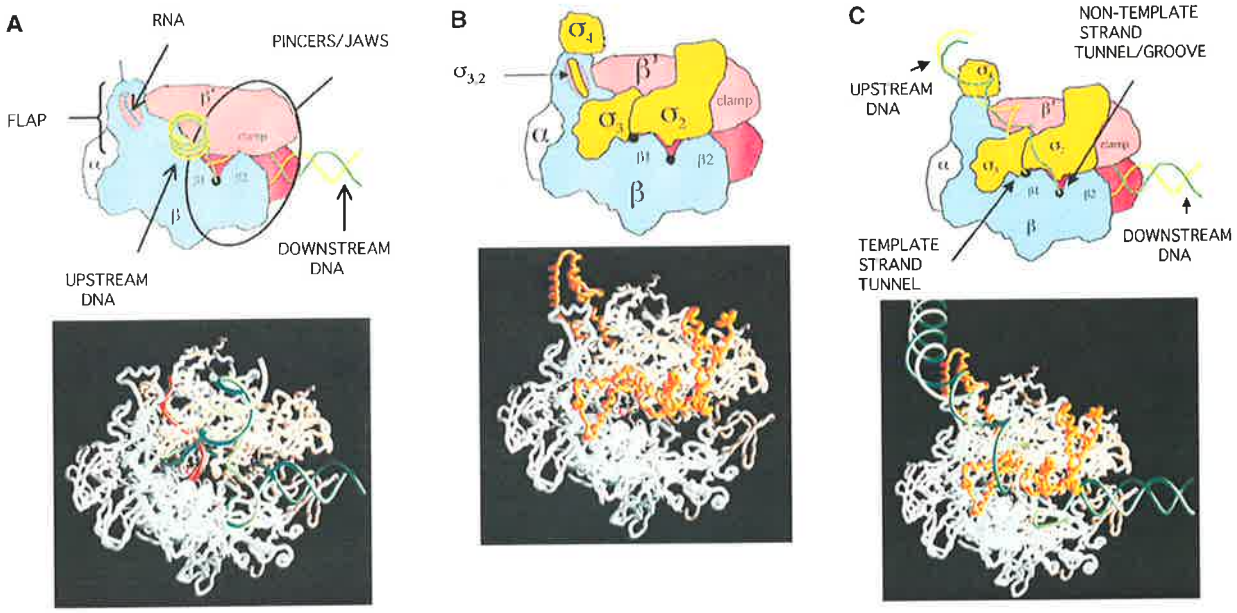


Figure 1.5. RNA polymerase structures.

(**a**) Elongating core. (**b**) Holoenzyme.(**c**) Open complex.
Ribbon diagrams derived from crystal structures are given below and cartoons showing RNAP subunits, DNA and relevant features are given above. Crystal structures show the placement of the Mg^{2+} ion at the active site in pink, the sigma subunit in gold, template DNA in Dark green, non-template DNA in light green and RNA in red.
Reproduced from figure 2, Young *et al.* (2002).

enclosed structure, the two strands of DNA in the transcription bubble are buried in different tunnels, which are separated by a highly conserved rudder structure.

The topology of DNA during transcription has also been well studied and it appears that during open complex formation promoter DNA is wrapped around RNAP with two thirds of the DNA upstream of the transcription start site and one third downstream (Rivetti *et al.*, 1999). During elongation RNAP is thought to be placed at the top of an apex of twisted DNA, feeding through the template DNA as a right handed helix and also bending the DNA at the site of contact (Heggeler-Bordier *et al.*, 1992). As elongating polymerase is thought to be fixed in space, the event of feeding and unwinding DNA as it translocates will produce domains of overwound DNA upstream and underwound DNA downstream of the polymerase (Liu and Wang, 1987).

The strength of a promoter depends on the combined efficiency of each of the individual steps of promoter initiation, so that the least efficient step becomes rate-limiting, acting as a bottleneck. All of these steps are potential targets for regulation, either by accelerating or slowing down the transition of one complex to the next. Thus, activity from a given promoter can be reduced by a large variety of mechanisms. Regulatory factors have been shown to act on virtually every step in the process of transcription in bacteria, from polymerase binding to elongation (reviewed in Rojo, 2001). Interference of promoter activity caused by convergent transcription is expected to work in a similar way, reducing the efficiency of one or more steps in initiation.

Eukaryotic transcription is complicated by the presence of histones and a much larger initiation complex but the principal mechanisms of RNA polymerisation are the same as prokaryotic transcription due to the high degree of structural similarity between the prokaryotic and eukaryotic RNAPs (reviewed by Ebright, 2000). Therefore although the mechanisms of transcriptional interference considered in this thesis are examined using a prokaryotic example of convergent transcription, conclusions may also apply to eukaryotic systems and are further discussed in Chapter 6.

1.4.2. Properties of interference based on previous studies.

Based on previous studies a number of general properties of interference by convergent transcription can be deduced.

Ward and Murray (1979) used the *trp* promoter of *E. coli,* placed convergent to the very strong P_L promoter of phage λ, to show *in vivo* that convergent promoters lead to blocked *trp* expression and partially inhibited *pL* expression. In this example promoter interference acted over a distance of several kilobases, was shown to be dependent on strong converging transcription (weak transcription from the *trp* promoter did not interfere with P_L activity), and required continuous transcription of the promoters (repression of P_L activity by λ CI repressor led to a derepression of the *trp* promoter). The repressive effects of convergent transcription demonstrated a dependence on promoter strength, with the stronger promoter repressing transcription from the weaker.

These properties were also observed using another artificial example of convergent transcription where a strong promoter was placed convergent to transcription from the weakly expressed *aadA* gene, at a distance of about 1.5 kilobases (Elledge and Davis, 1989). *aadA* encodes aminoglycoside adenyltransferase, which provides bacterial resistance to aminoglycoside antibiotics such as spectinomycin. Interference by convergent transcription was measured semi-quantatively by the level of spectinomycin sensitivity of cells carrying the convergent promoter construct. Interference was shown to be dependent on a convergently oriented promoter (that is reversing the orientation of the strong convergent promoter gave no effect) and by continuous transcription (repressing strong convergent transcription lead to a loss of interference). The influence of relative convergent promoter strengths on interference was also investigated. Altering the strength of the aggressive promoter demonstrated that interference was dependent on strong opposing transcription. Unfortunately, it is difficult to determine the precise nature of the relationship between aggressive promoter strength and the extent of interference as comparisons of the *in vivo* convergent promoter strengths by reporter gene constructs were not performed. However, generally a decrease of convergent transcription below a defined range of activity led to a loss in *aadA* promoter interference.

A recent study of interference in cultured mouse cells used artificially constructed convergent promoters that were recombined into the cells genome (Eszterhas *et al.*, 2002). Promoters were spaced approximately 2kb apart and each expressed a different reporter gene. Reporter gene expression was shown to be lowest for promoters arranged in a convergent orientation, rather than divergently or in tandem, indicating that interference in a eukaryotic system is also dependent on its convergent nature. A dependence on overlapping transcription for the display of interference between promoters spaced 2-4kb apart has been demonstrated in plant and yeast cell systems, by the observation that interference is lost when transcription from both promoters is terminated midway between promoters (Ingelbrecht *et al.*, 1991; Padidam and Cao, 2001; and Prescott and Proudfoot, 2002).

The influence of spacing between promoters on the extent of interference has only been examined for promoters at close range (Jagura-Burdzy and Thomas, 1997). The native convergent promoters of plasmid RK2 were used and the interference of promoters spaced at a range of distances from the wild type distance of 49 bp down to 20 bp, was measured. The activity of the sensitive promoter decreased as the promoters were brought closer together. Comparisons of promoters spaced at larger distances have not been made.

1.4.3. The possible mechanisms of interference.

The possible molecular mechanisms of interference presented here are based on the properties of interference gleaned from published studies, and the discussions of previous reports of interference in both bacterial and eukaryotic systems (Brantl and Wagner, 1997; Elledge and Davis, 1989; Horowitz and Platt, 1982; Jagura-Burdzy and Thomas, 1997; Ward and Murray, 1979; Eszterhas *et al.*, 2002; Padidam and Cao, 2001; Prescott and Proudfoot, 2002). The possible molecular mechanisms of interference (fig. 1.6) can be divided into two components: the action of the aggressive promoter, and the response of the sensitive promoter. The aggressive promoter may cause interference by: 1) binding RNAP, 2) producing RNAP complexes elongating towards, over or past the sensitive promoter, 3) producing anti-sense transcripts, 4) inducing changes in local DNA topology, or 5) restructuring of architectural DNA proteins that affect convergent promoter activity. The effect on the sensitive promoter may be due to 1) interference

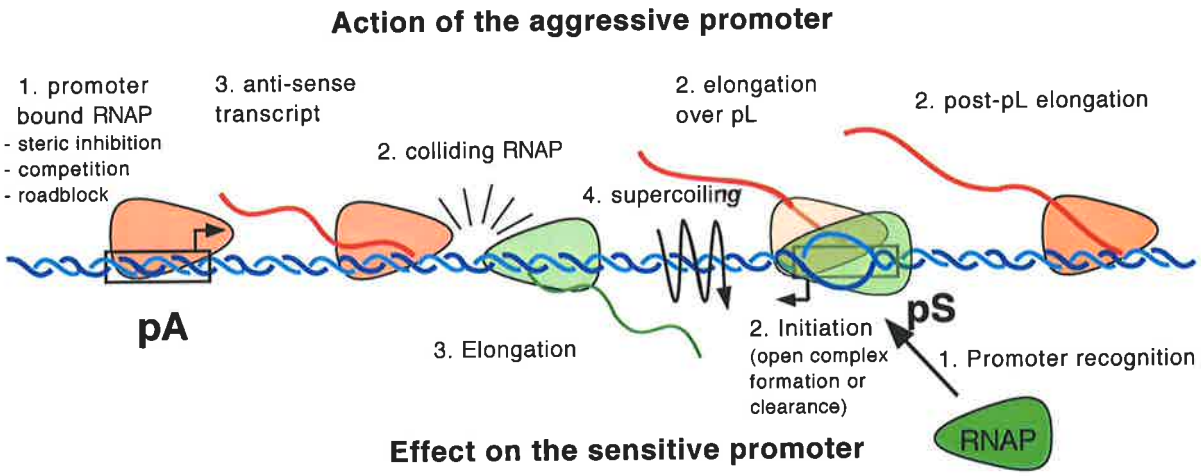


Figure 1.6. Cartoon of possible interference mechanisms.

pA is the aggressive promoter, pS is the sensitive promoter. See text for further details.

with promoter recognition, 2) a disruption of open complex formation or other promoter initiation steps or 3) disruption of elongation from the sensitive promoter either by pausing or premature termination of its transcript. Combining these two components gives a wide variety of possible mechanisms, and more than one mechanism may be operating for any given set of convergent promoters. A number of previous studies have begun to examine the likelihood of some of these mechanisms for the specific example of convergent transcription being studied. As determination of the relevance of each mechanism operating in a specific natural example is the subject of this study, the exact nature of each mechanism is elaborated further in the results chapters.

1.4.4. Specific mechanisms examined by previous studies.

Previous studies have examined some of these specific mechanisms of interference, including those that involve steric hindrance from polymerase bound at the aggressive promoter, the effect of antisense transcription, collisions between polymerases and changes in DNA topology.

Steric hindrance by promoter bound polymerase

RNAP bound at the aggressive promoter could influence activity at the sensitive promoter in a fashion similar to 'classical' repression (eg. *lexA* repressor binding over the –35 region of the *uvrA* promoter, (Bertrand-Burggraf *et al.*, 1987)) where the binding of the repressor molecule blocks the access of RNAP to the repressed promoter. However, unlike most classical repressors that bind operators which overlap the sites of RNAP recognition, binding of the repressor molecule (RNAP) occurs downstream of the repressed promoter. Dnase I and hydroxyradical DNA footprinting studies of the σ^{70} RNAP open complex formed at a promoter, showed DNA protection from approximately –55bp to +20bp from the start site of transcription (Record *et al.*, 1996). The importance of spacing between promoters in enabling this type of interference was shown by Jagura-Burdzy and Thomas, (1997). Simultaneous occupancy of convergent promoters by *E. coli* RNAP was measured using mobility shift assays of PCR products of convergent promoters spaced 49, 37, 32, 26 and 20 bp apart with RNAP. Results showed that RNAP can simultaneously occupy promoters spaced 37 and 49 bp apart but not closer.

Interference at the level of polymerase binding is suggested for the convergent promoters of the *E. coli* genes *dnaQ* and *rnh* (Nomura *et al.*, 1985b). *dnaQ* is transcribed from two promoters, P1 and P2 which are separated from the convergent promoter *rnhp* by about 107 and 24 bp respectively (Nomura *et al.*, 1985a). Linear templates containing these convergent promoters were used to analyse the amount of convergent transcription from each promoter *in vitro* in single round transcription assays. Under conditions of a molar excess of polymerase to template concentration, the amount of transcription was similar for *dnaQ*-P1 and *dnaQ*-P2, both about twice as strong as *rnhP*. However, when DNA is at molar excess, far more transcription was observed from *dnaQ*-P2 than *dnaQ*-P1 or *rnhP,* suggesting differential utilisation of each promoter at limiting polymerase concentrations. When the promoters were separated under these conditions, *dnaQ*-P2 was still dominant, however *rnhP* activity was greater than *dnaQ*-P1 suggesting promoter interference as a consequence of convergent transcription, but only at low polymerase concentrations. To observe interference it was necessary to utilise low polymerase concentrations; the authors suggest this interference could well be due to steric occlusion occurring between the closely spaced *dnaQ*-P2 and *rnhP*. This conclusion is consistent with that of Jagura-Burdzy and Thomas, (1997). A mechanism of interference at the level of open complex formation is also suggested for the convergent promoters of the *fis* operon. The promoters are separated by 7 bp and open complex at one promoter was shown by $KMnO_4$ footprinting to reduce that of the opposing promoter (Nasser *et al.*, 2002).

The potential for interference at the level of mutual open complex formation was examined in the convergent promoters *p*IN and *p*OUT of the *E. coli* insertion sequence, IS10, separated by 36bp (Simons *et al.*, 1983). The *in vitro* steady state rate of abortive initiation from each promoter was analysed on linear template fragments carrying either both or one of the promoters. No major differences between the activities of convergent or intrinsic promoters were observed, and respective *p*IN and *p*OUT activities correlated to those found *in vivo*. This demonstrated that on a linear template open complex formation at the two promoters is not mutually exclusive. It was suggested that independent open complex formation was possible because the sites of polymerase interaction lie on opposite sides of the template helix. Hence, interference for promoters

spaced 36bp apart or more does not appear to involve mutual inhibition of open complex formation. It should be noted that co-binding of RNAP has only been studied *in vitro* using linear templates. The conclusions from these studies may not be relevant during promoter activity in an *in vivo* environment which includes transcription from naturally supercoiled promoter DNA.

Interference by antisense transcription

Antisense RNA effects have been documented in both eukaryotic and prokaryotic systems (Brantl, 2002). It is generally considered that a mechanism involving the hybridisation of antisense transcripts impeding either transcription or translation, could act *in trans*. In the case of convergent transcription, supply of the interfering transcript generated *in trans* (eg. from a separate plasmid during *in vivo* experiments) should therefore be able to restore the interference of a promoter normally affected by convergent transcription. *In trans* experiments using plasmids have been undertaken using a number of convergent promoters. Using the convergent lytic and lysogenic promoter system of coliphage P2, an *in trans* supply of lytic transcript reduced lysogenic transcription only by 3-fold (Saha *et al.*, 1987b). This was deemed minor compared to the large ~30-fold reduction caused by *in cis* lytic transcription. Elledge and Davis (1989) reported no change in *aadA* transcription when antisense *aadA* transcript was supplied *in trans*. A subsequent study using the same *aadA* convergent transcription system investigated the repressive effects of DNA triplex formation and found that the *in trans* supply of small RNAs complementary to the 5' end of the anti-*aadA* transcript actually relieved the inhibition of *aadA* expression caused by strong converging transcription (Soukup and Maher, 1998). The effect of *in trans* expression of an antisense transcript was also examined in yeast cells and again no interference was observed (Prescott and Proudfoot, 2002).

It should be noted that the use of an *in trans* supply of antisense transcript from plasmids may not be equivalent to that supplied from a convergent promoter, both in location and amount. Additionally, convergent transcription will produce antisense transcripts in combination with overlapping transcription. Although the antisense transcript itself may not cause interference, the hybridisation of the 5' ends of

convergent transcripts may act to "tether" RNA polymerases elongating in opposite direction on the same DNA molecule, such that the combination of overlapping and antisense transcripts produce the interference.

Interference due to the head-on collisions between converging polymerases.

The simplest mechanism of interference is that described by Ward and Murray (1979) where two RNAP molecules collide, transcription stops and RNAP is released from the DNA. This type of interference would be expected to reduce the aggressive promoter by the same amount as that of the sensitive promoter. However in the case of strong transcription (eg. 1000 units of activity), convergent to weak transcription (eg. 100 units of activity), an 80 unit reduction in activity would greatly reduce the weak promoter (eg. 5 fold) but be insignificant to the strong promoter.

The likelihood of a collision mechanism of interference would depend on the outcome of such collisions. No interference would be expected if RNA polymerases could pass each other unhindered. However interference may be possible if collisions cause both polymerases to become unstable and terminate transcription, or cause only one of the two polymerase molecules to become unstable (with instability dependent on the surrounding DNA sequence), or if polymerases could pass but only at a slow rate. Different outcomes might be expected from the collision of an elongating polymerase and a promoter bound complex. A pausing effect of collisions between elongating RNAP and promoter bound polymerase is suggested from single round kinetic *in vitro* transcription experiments, performed on linear templates containing convergent *E. coli lac* and *trp* operon promoters separated by 163bp (Horowitz and Platt, 1982). Transcription from pre-formed open complexes was initiated by the addition of magnesium and transcript accumulation was measured with time. Studies from non-convergent templates had shown that full length transcript production from the *lac* promoter was much slower than from the *trp* promoter. *In vitro* transcription experiments using convergent promoters were expected to induce collisions between polymerase molecules elongating from the *trp* promoter and open complexes formed at the *lac* promoter. Initially, only a small transcript of a size expected for elongation from the *trp* promoter which had halted at the *lac* promoter was observed. At the same time

that full length *lac* promoter transcript was observed, the paused transcript began to disappear and, eventually, only full length *trp* and *lac* transcripts were recorded. This experiment was interpreted to indicate that the progress of an elongating polymerase involved in head-on collisions with a promoter bound polymerase is initially blocked, but once the promoter bound polymerase initiates transcription then the two molecules pass through each other to produce full length transcripts. If this type of pausing was to occur *in vivo*, then it would be expected to affect activity of the aggressive promoter as much as the sensitive one. Interference of the *in vivo* activity of these convergent promoter was examined in a previous study and found that there was little or no effect on transcription from opposing promoters (Miller *et al.*, 1970).

No other *in vitro* study has directly examined the outcome of collisions between elongating polymerases, however predictions can be made based on evidence about the stability and structure of elongating RNAP. Elongating polymerase is extremely stable to dissociation with an approximate net stability of –18 kcal mol^{-1} at an average template position (Hippel, 1994). Due to the double strandedness of DNA it seems plausible that converging polymerases may be able to pass through each other, keeping contact with the DNA through interaction with its single stranded complementary template. However, from the X-ray crystal structures (see section 1.4.1.) of elongating bacterial polymerase it seems likely that head-on collisions will inevitably involve steric clashes between polymerase domains. Thus, passing polymerases could not occur without major conformational changes, which would suggest inherent flexibility of polymerase structures not apparent from the current models. Clues about the stability of elongating polymerase to collisions with other polymerase molecules are given by *in vitro* experiments performed to observe the outcome of collisions between an elongating RNAP and a replication fork. Studies have examined the outcome of either T4 DNA polymerase or ϕ29 DNA polymerase colliding both head-on or behind a elongating RNAP stalled downstream of the promoter by omitting one NTP (Liu *et al.*, 1993; Liu and Alberts, 1995; Elias-Arnanz and Salas, 1997; Elias-Arnanz and Salas, 1999). Surprisingly a T4 replication fork was found to pass through a codirectionally located transcription complex without displacing it, and upon the addition of all four NTPs, elongation was able to resume (Liu *et al.*, 1993). Similar results were obtained using moving RNAP (Liu *et al.*, 1994). Comparable studies involving head-on collisions

found that the replication fork was again able to pass RNAP but only after a pause of a few seconds (Liu and Alberts, 1995). Using the bacteriophage ϕ29 DNA polymerase, codirectional collisions caused a complete block in replication, both polymerases remained associated with the DNA, and once RNAP translocation was allowed, DNA polymerase resumed replication (Elias-Arnanz and Salas, 1997). Head-on collisions also caused stalling of the replication fork, but again both remained bound to the template and, when the halted RNAP was allowed to move, both polymerases resumed normal elongation (Elias-Arnanz and Salas, 1999). These experiments have led to the suggestion of a resolving mechanism of RNAP to survive polymerase collisions. If an elongating RNAP can survive the progress of a converging replication fork, it therefore seems plausible that converging RNA polymerases can also pass through each other, which could make this mechanism of promoter interference obsolete unless such collisions caused substantial pausing.

Despite the *in vitro* evidence suggesting that collisions are unlikely to have a significant role in interference, the effect of transcriptional collisions between RNAP II molecules was the explanation given for the *in vivo* interference observed between convergent genes in budding yeast (Prescott and Proudfoot, 2002). A transcription 'run-on' (TRO) analysis was used to probe the amount of transcript produced from each convergent promoter at different distances from the start sites of transcription. In the absence of convergent transcription the amount of transcript produced from each promoter generally remained constant for all probes used. During convergent transcription, the amount of transcript produced from each promoter directly after initiation was equivalent to that measured in the absence of convergent transcription, indicating that the rate of initiation from each promoter was unchanged. However the amount of transcript decreased almost linearly with distance downstream from the point of initiation, until the point elongation reached the convergent promoter, almost background levels of transcript were seen. This was interpreted to indicate that the interference of convergent transcription does not reduce initiation rates of either promoter, but does reduce elongation, most likely by head-on collisions. It is important to note that this study was in a eukaryotic system and a similar study in prokaryotes is yet to be performed. Additionally, the convergent promoters used in this study were spaced nearly 3kb apart, and as the frequency of collisions is expected to be dependent

on the distance between promoters (see section 3.3.1), this mechanism may not account for the interference of more closely spaced promoters.

As discussed by Prescott and Proudfoot, (2002), a collision effect could be caused by direct physical impediment to the transcriptional machinery or an indirect effect caused by supercoiling changes to the DNA template during transcription. In the latter scenario, a transcription bubble progresses along a torsionally constrained template, creating positive supercoils ahead and negative supercoils behind (Liu and Wang, 1987). If these positive supercoils were generated from both strands simultaneously, it would create a region of hyper-supercoiling that would be predicted to prevent further advancement in either direction on the template. Studies that have so far examined collisions *in vitro* have not used torsionally constrained templates, which may explain the apparent differences in their conclusions.

Topology

Active transcription of topologically constrained DNA, such as chromosomal DNA, is known to cause changes in the local DNA supercoiling, with waves of positive supercoils induced ahead of transcribing polymerase and negative supercoils behind (Liu and Wang, 1987). As changes in DNA topology have been known to inhibit or promote the expression of a number of promoters (Opel *et al.*, 2001), it is conceivable that the changes in topology caused by transcription of the aggressive promoter could be responsible for interference of a convergent sensitive promoter. The reverse of this is that transcription from closely spaced divergent promoters should influence each other positively by transcription induced negative supercoiling, leading to transcriptional coupling of promoters. This has been shown in a natural biological system for divergently transcribed promoters of the *ilvYC* operon of *E. coli* (Opel *et al.*, 2001). It is conceivable that repeated waves of positive supercoils resulting from initiation at a strong promoter would interfere with RNAP attempting to bind at a nearby downstream convergent promoter.

To test whether interference is affected by supercoiling, Northern blot analysis of the amount of transcripts produced from the convergent promoters pII and pIII in the

replication control region of pIP501 in *Bacillus subtilis*, was performed on cells treated with novobiocin (an inhibitor of subunit B of DNA gyrase) (Brantl and Wagner, 1997). pII was shown to interfere with pIII transcription. With novobiocin treatment, activities of both promoters were shown to be sensitive to supercoiling, with activity increasing with time after treatment. Interference also increased, prompting the suggestion that transcriptional induced changes in supercoiling are involved in the mechanism of interference. As novobiocin treatment indicated that pIII is supercoiling sensitive, it was suggested that transcription-induced changes in supercoiling may be responsible for interference in pIP501. It is also suggested that supercoiling-induced changes in DNA topology, such as cruciform extrusion at one of several inverted repeat sequences in or near the pIII promoter region, may also be responsible. It is difficult to determine from this experiment the precise involvement of supercoiling during interference as the increase in interference may be due to changes in the cellular ability to diffuse supercoils or simply an indirect result of novobiocin treatment changing the intrinsic properties of the convergent promoters, rendering them more sensitive to mechanisms of interference that do not involve supercoiling. I am not aware of any other published studies which have specifically examined interference mechanisms involving changes in DNA topology. The status of this type of interference mechanism is discussed further in Chapter 6.

1.5 Aims of this study

To better understand the significance of convergent transcription as a method of regulating gene expression in biology, the aim of this study was to understand the interference mechanism operating in the convergent promoters of the developmental switch of bacteriophage 186, using the switch promoters of bacteriophage P2 to confirm any predictions made. A natural example of convergent transcription has been chosen because a biological role of interference has already been determined, and an understanding of the interference mechanism will contribute to an understanding of a known means of controlled gene expression. A study of the interference between the lytic and lysogenic promoters of 186 uses an easily manipulated prokaryotic system and involves a strong aggressive promoter and a weak sensitive promoter, spaced only 62 bp apart. It was intended that understanding the mechanisms operating for one specific

example would lead to a more generalised understanding of interference. Analysis of another example of convergent transcription in P2 contributes to this aim.

To begin a study of interference between the switch promoters of 186, an *in vivo* promoter reporter system was established to assay promoter activities and interference. This work is detailed in Chapter 2. This reporter system was then used in Chapter 3 to investigate what property of the aggressive promoter is responsible for interference. This was achieved by an analysis of interference after alterations in the arrangement of the aggressive and sensitive promoters. Placement of a transcriptional terminator between the promoters showed that the majority of interference *in vivo* was due to RNAP elongation over the sensitive promoter. The nature of the interference caused by the passage of RNAP is reported in Chapter 4 where *in vitro* transcription assays showed that *pL* forms open complexes which are slow to clear, and that the activity of these open complexes is reduced by RNAP transcription from *pR*. Using the lysogenic promoter from the related bacteriophage P2, it is shown in Chapter 5 that a promoter's sensitivity to interference correlates with the speed of promoter clearance. Thus, it is proposed that interference in the 186 switch system occurs because RNAP complexes at *pL* that are waiting to clear are sensitive to passing elongating polymerase from *pR*. This 'sitting duck' mechanism is likely to be important for promoters that are close together. Finally, a general mechanism for transcriptional interference by convergent promoters is discussed in Chapter 6. (The materials and methods used in this study including the construction of all relevant plasmids are reported in the final chapter 7.)

CHAPTER TWO

Measuring transcriptional interference between 186 *pR* and *pL in vivo*

Chapter 2

Measuring transcriptional interference between 186 pR and pL in vivo

2.1 Introduction

Previous studies of *in vivo* promoter activity have shown that the activity of *pR* is intrinsically 10-fold stronger than *pL* and that convergent transcription from *pR* inhibits *pL* transcription some 6-23 fold (Dodd *et al.*, 1990; Reed *et al.*, 1997; Neufing *et al.*, 2001; Dodd and Egan, 2002). These results, together with the types of promoter assays used in these studies, are summarised in fig. 2.3 and discussed in section 2.3 with respect to the results of this study. Due to the variation in interference values observed between practitioners, it was evident that in order to pursue a definitive study of transcriptional interference, it was necessary to establish a method of determining promoter strength that was reliable, internally consistent, allowed for an adequate statistical analysis of interference and could easily allow comparisons of the effect of different promoter constructs on *pL* activity (such as those used in chapter 3 to examine how *pR* activity causes interference). This chapter details the establishment of this system.

2.2 Rationale for establishing a system to measure 186 interference.

To measure interference *in vivo* it was necessary to establish 1) the fragment of promoter DNA to be assayed, 2) the type of assay to be performed, 3) the type of promoter vector system to be used, and 4) the host strain to be assayed. These elements are detailed here.

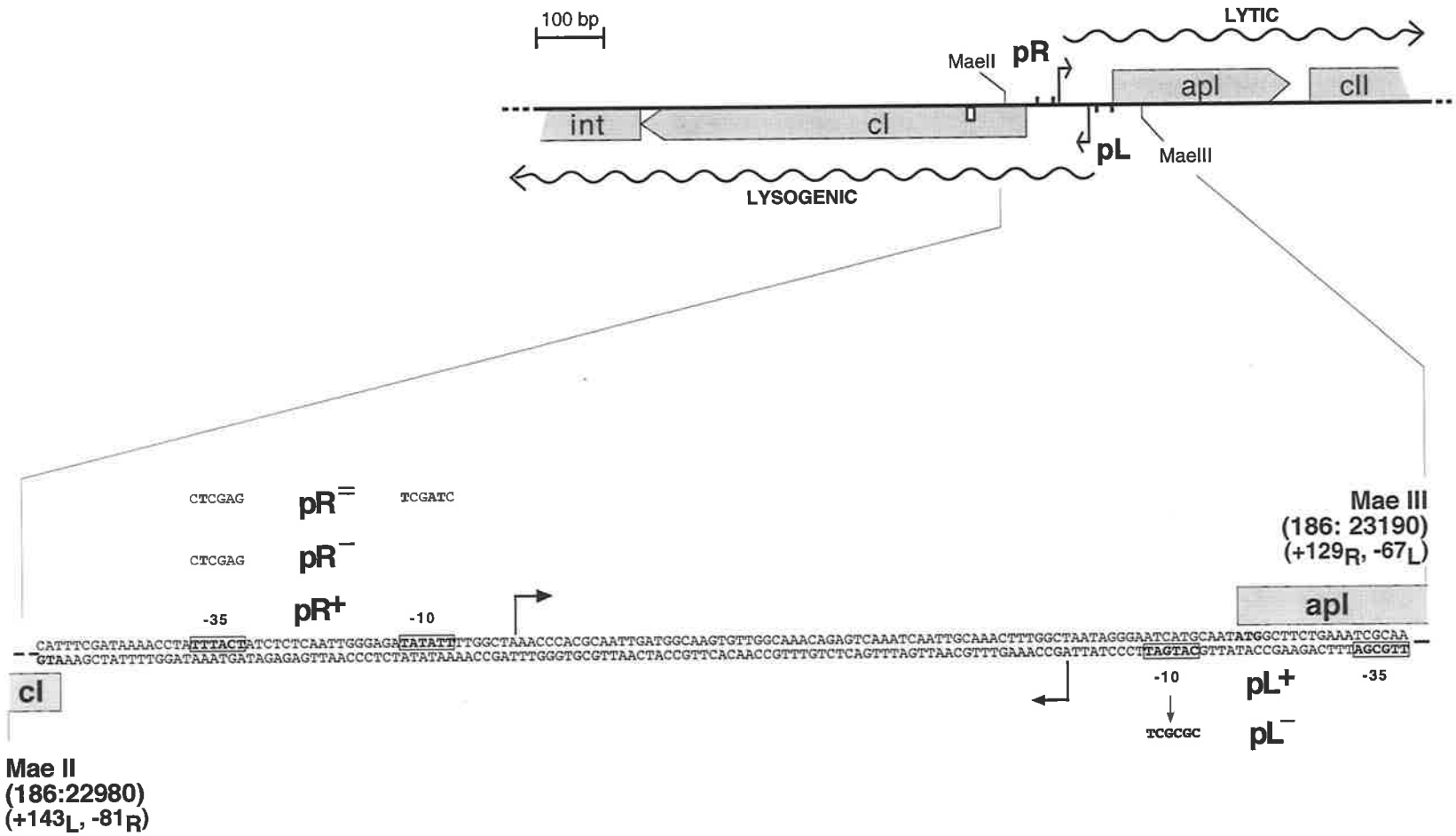
2.2.1 Promoter fragment.

The region of DNA chosen for assaying convergent promoter activity was the 210bp of 186 (coordinates 22980 to 23190, Genbank reference: U32222), which includes the Mae II and Mae III sites and DNA from 81bp upstream of *pR* to 67bp upstream of *pL* (fig 2.1). This region contains the near minimal fragment that includes both *pR* and *pL*

Figure 2.1

Convergent promoters of the 186 developmental switch.

Developmental switch region of 186 (above) and the DNA sequence (below) of convergent promoters *pR* and *pL* of the fragment used in this study. The switch region is drawn to scale. Filled boxes indicate positions of –10 or –35 hexamers for promoters *pR* or *pL*, the white box indicates the predicted IHF binding sequence. On the sequence, 186 genome coordinates are those found in GenBank. Bent arrows indicate the start sites of transcription, predicted –10 and –35 hexamers are shown in bold and boxed and the promoter mutations used also indicated. $\#_R$ and $\#_L$ indicate the distance from the start of *pR* and *pL* transcription respectively.


100 bp
LYTIC
MaeII
pR
apI
cII
int
cI
pL
MaeIII
LYSOGENIC
CTCGAG
pR=
TCGATC
CTCGAG
pR−
pR+
-35
-10
Mae III
(186: 23190)
(+129R, -67L)
apI
CATTTCGATAAAACCTATTTTACTATCTCTCAATTGGGAGATATATTTTGGCTAAACCCACGCAATTGATGGCAAGTGTTGGCAAACAGAGTCAAATCAATTGCAAACTTTGGCTAATAGGGAATCATGCAATATGGCTTCTGAAATCGCAA
GTAAAGCTATTTTGGATAAATGATAGAGAGTTAACCCTCTATATAAAACCGATTTGGGTGCGTTAACTACCGTTCACAACCGTTTGTCTCAGTTTAGTTAACGTTTGAAACCGATTATCCCTTAGTACGTTATACCGAAGACTTTAGCGTT
cI
-10
pL+
-35
TCGCGC
pL−
Mae II
(186:22980)
(+143L, -81R)

sequences, as it is known that in closed or open complexes with RNAP, promoter DNA is protected from DNase I digestion from approximately –55 to +20 from the start of transcription (Record *et al.*, 1996). This short DNA fragment containing *pR* and *pL* was used so that any interference measured could be attributed to the minimal DNA region containing both promoters, rather than to any flanking sequences. Moreover, during anticipated *in vitro* transcription experiments performed with this promoter region, short DNA fragments would allow for differential separation of *pR* and *pL* transcripts on a 6% polyacrylamide gel. Note that a weak IHF binding site is present within the *cI* gene upstream of *pL* (see fig 2.1), however preliminary experiments attributed no role to the IHF site in normal 186 phage development (D. Reynolds, unpublished). The IHF site was not included in these constructs.

Measuring interference requires assaying the sensitive promoter activity in the presence (repressed activity) and absence (intrinsic activity) of the active interfering promoter. This can be achieved by either deletion of the convergent promoter or by inactivation by mutation. Assaying *pL* from the 210bp 186 fragment carrying mutations in *pR* (fig 2.1) was the preferred method, as it maintains the similarity in DNA context and thereby minimises differences in the stability and translation of the reporter gene used to assay promoter strengths. A mutation in the –35 hexamer of *pR* (mutated from TTTACT to **CTCGAG**, (Reed *et al.*, 1997)) was used to measure leftward transcription in the absence of *pR* activity, and a mutation in the-10 hexamer of *pL* (mutated from CATGAT to **CGCGCT**, (Neufing *et al.*, 2001)) was used to measure rightward transcription in the absence of *pL* activity. The activities of these mutants are described in section 2.3.5.

2.2.2. Type of promoter assay.

Promoter activity was assayed by determining the amount of β-galactosidase produced from strains carrying the promoter of interest fused upstream of a promoterless *lacZ* gene. The advantage of *lacZ* gene fusions is that *lacZ* mRNA and protein are stable in the cell and assays for β-galactosidase activity are simple and widely used.

Lac Z activity of reporter strains was assayed using a semi-automated microtitre plate system, as reported by Dodd *et al.*, (2001). The advantages of this system over

traditional methods (Miller, 1972) are that firstly, the 96-well format quickly and easily enables a large number of assays to be performed. Activities of one strain were repeated 10 to 12 times in one plate and plates were often repeated 2 to 3 times. Secondly, the assays are kinetic, measuring the rate of ONPG degradation over 1 hour and are therefore more accurate than traditional methods which use a single time point to assay activity (Miller, 1972).

Initial experiments (data not shown) established that there was a small, but significant, correlation between lacZ activity and A600 of the culture assayed. There was also some variation between the same strain assayed on different days. To accommodate this variation in a non-biased fashion, interference was calculated from the lacZ activities as follows: i.) the lacZ activities for the individual cultures (usually 10 to 12) of a particular clone were first ranked according to their A600 values, ii) ranked units for the intrinsic clone were then divided by equivalently ranked units for the interfering clone, which had been assayed on the same 96-well plate, giving individual interference measurements, iii) the $\log_{10}$ of each of these ratios was calculated and the average, standard error, and 95% confidence limits (based on the students *t* distribution) of these logs of ratios was calculated. (The log of the ratios was used rather than the numerical ratios, as logs of ratios are more likely to follow a normal distribution than the ratios themselves (Ryder and Robakiewicz, 1998).) iv) The antilogs of these averages and confidence limits were then taken, giving the average interference ratio and the 95% confidence limits of that ratio. LacZ units for a particular strain are the averages and 95% confidence limits, based on the students *t* distribution, from experiments performed on at least two separate days.

2.2.3. Type of promoter vector system.

The λ-based, chromosomal *lacZ* operon fusion system of Simons *et al.* (1987), modified as described in Dodd *et al.* (2001) was used. This is a single copy system where the promoter of interest is cloned in front of a promoterless *lacZ* gene contained on a plasmid. The insert and functional *lacZ* gene sequences are then transferred to a modified bacteriophage λ by recombination, and bacterial λ lysogens containing only a single chromosomal copy of this recombinant phage are then created and assayed for lacZ activity (see Materials and Methods Chapter 7). The advantage of single copy

systems over plasmid based multi-copy systems is the elimination of variations in lacZ activity resulting from variation in plasmid copy number, which is discussed further in section 5.5.1.

The *lac Z* assay vector initially chosen was pMRR9 (fig 2.2) used by Reed *et al.*, (1997); Neufing *et al.*, (2001) and Dodd and Egan, (2002). This vector features a pUC19 multiple cloning site (MCS) upstream of a promoterless *lacZ* gene, with four tandem repeats of the strong transcriptional terminator T1 upstream of the site of promoter insertion. This isolates the transcriptional activity of the promoter insert from transcription upstream of the insert. These features result in a vector that has a low background lacZ activity, which is helpful for the accurate measurement of low promoter activities eg.*pL*. This vector also features three stop codons downstream of the MCS which prevents the continued translation into *lacZ* (in all three frames) initiated from upstream inserted sequences.

The modified λ initially chosen was λRS45 (fig 2.2). This phage was constructed by Simons *et al.*, 1987 and has been used in combination with pMRR9 to generate single copy chromosomal *lacZ* reporter fusions in this laboratory (Reed *et al.*, 1997; Neufing *et al.*, 2001; and Dodd and Egan, 2002)).

To analyse the activity of a particular construct, in general at least two different single lysogens of the same construct were assayed and the data pooled. Lysogens were either different single lysogens from the same recombinant phage or single lysogens from different recombinant phage. For reasons that are not understood, there was some variations in activity between different lysogens that were within the 95% confidence limits given for each construct.

2.2.4. Host strain.

The Δ*lacIZYA E. coli* strain NK7049 was initially chosen as the host strain for lacZ assays because it is a host for 186, and has been used in similar types of experiments that measure interference (Dodd and Egan, 2002). The use of different host strains is explored further in section 2.3.3.

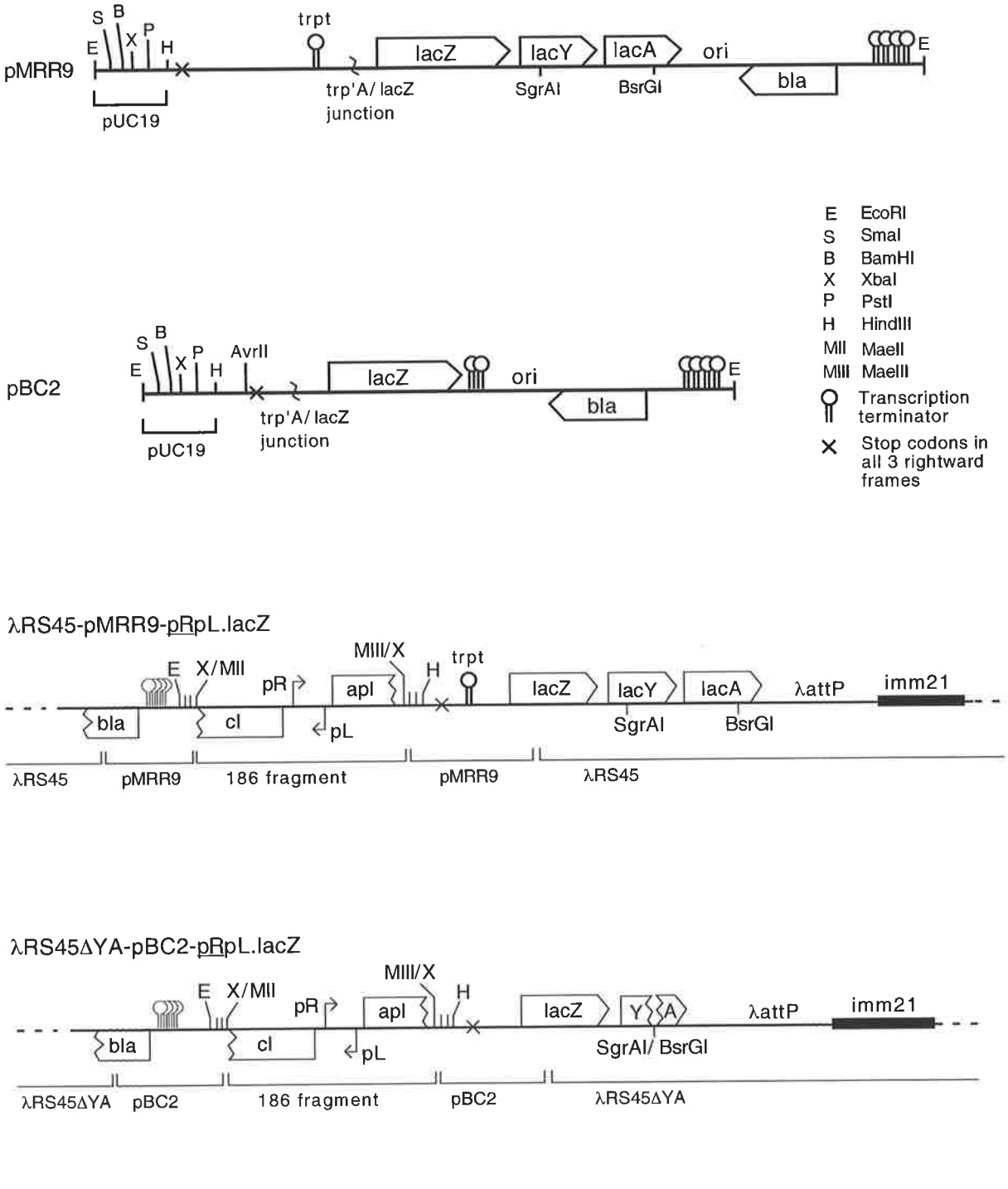


Figure 2.2 *LacZ* reporter vectors and λ phage reporters.

Diagram showing the relevant details of the *lacZ* containing plasmids used to measure *in vivo* promoter activities in this study and also a representation of the recombinant λ phage used to produce single copy chromosomal versions of the reporters, (adapted from Dodd and Egan, 2002).

2.3. Results

2.3.1. Short fragments assayed using pMRR9 gave 7 fold interference.

pMRR9-*pRpL*, pMRR9-*pRpL* and pMRR9-*pR⁻pL* constructs were made using the short 186 DNA fragments and single copy NK7049(λRS45) lysogens made from these clones. These were assayed for *β-galactosidase* activity as discussed above. Results are shown in (fig 2.3(a)). The nomenclature used for constructs is that the promoter being assayed is underlined, and promoters are active unless otherwise indicated. Results showed that intrinsic *pL* activity was 42 ± 6 units, which was about 20-fold weaker than *pRpL* activity, found to be 748 ± 59 units. In the presence of active *pR* transcription, *pL* activity was reduced to 6 ± 0.6 units, with interference calculated to be 7 ± 0.4 fold. This level of interference was close to the ~6 fold interference published by Dodd *et al.*, (2002).

2.3.2. Repeating the assays in an improved vector system gave 5.3 fold interference.

Many of the *lacZ* reporter assay systems, including pMRR9, are derivatives of the original W205 *trp-lac* fusion (Linn and Pierre, 1990). In 1998, Liang *et al.* reported that this fusion retains a weak temperature dependent terminator (*trpt*) near the junction, and reduces transcription *in vivo* by approximately 40% (Liang *et al.*, 1998). It was possible that the presence of this weak terminator could contribute to variation between lacZ activities of different strains or between individual assays, as termination efficiency could differ between different promoter constructs and as a result of slight variations in temperature. Furthermore, measuring interference requires accurate measurement of low lacZ units from weakened promoters, and a weak terminator would further reduce low levels of transcription. Accurate measurements of lacZ activity are more difficult at low expression levels. These problems were eliminated by the use of a new *lacZ* reporter vector, pBC2 (fig 2.2), which deletes *trpt*. pBC2 was constructed such that it contains the same multi-cloning site as pMRR9 and exhibits a low level of background lacZ activity. Additionally, strong transcriptional terminators are placed downstream of the *lacZ* gene to completely isolate the insert/*lacZ* gene region from outside transcription. The *lacY* and *lacA* genes were also deleted which may assist in the cloning of strong promoters and is useful for other studies involving IPTG induced

	Reporter plasmid	Reporter phage	Host strain	Fold Interference	Leftward Activity (Units)	Construct	Rightward Activity (units)
(a)	pMRR9	λRS45	NK7049	**7** (±0.4)	6 (±0.6)	pR ⟶ pL	748 (±59)
					42 (±6)	pR⁻ ⟶ pL	ND
(b)	pBC2	λRS45	NK7049	**5.3** (±0.2)	15 (±0.6)	pR ⟶ pL	998 (± 56)
					81 (±4)	pR⁼ ⟶ pL	ND
(c)	pBC2	λRS45ΔYA	MC1061.5	**5.6** (± 0.2)	26 (± 1)	pR ⟶ pL	1340 (± 113)
					142(± 3)	pR⁼ ⟶ pL	ND
(d)	pBC2	λRS45ΔYA	MC1061.5		ND	pR ⟶ ✕ pL⁻	1433 (± 161)
(e)	pBC2	λRS45ΔYA	MC1061.5		5(± 0.6)	pR⁼ ⟶ ✕ pL⁻	1(± 0.4)

Figure 2.3 Activities of *pL* and *pR* and calculated interference for different reporter systems.

(a) to **(c)** compare promoter activities and interference from *pRpL* using different reporter systems. **(d)** examines the effect on *pR* activity of inactivating *pL*. **(e)** examines the activities of mutated *pR* and *pL*. Activities of all constructs are the lacZ activities measured from the single copy lysogens carrying the constructs indicated which had been inserted into the XbaI site of pMRR9 or pBC2. Leftward activity indicates transcription measured by placing the constructs indicated into the *lacZ* reporter such that *pL* directs transcription of the *lacZ* gene. Rightward activity indicates equivalent constructs placed into the *lacZ* reporter such that *pR* directs transcription of the *lacZ* gene. Fold interference was measured as described in the text using ratios of pR^{-} (or $pR^{=}$) over pR^{+} leftward *lacZ* units. All lacZ assays were performed at least 12 times and errors show 95% confidence limits. ND indicates not determined.

expression of regulatory factors (not used here) (Linn and Pierre, 1990). A derivative of λRS45 was used that also has the *lacYA* deletions, ie. λRS45ΔYA (fig 2.2) (Dodd *et al.*, 2001).

Before repeating lacZ assays using the new vector it was decided to measure intrinsic *pL* activity in the presence of double mutations in *pR* ($pR^{=}$); with mutations in the -10 (mutated from TATATT to TCGATC) as well as mutations in the -35 (Reed *et al.*, 1997) (fig 2.1). This was included because *in vitro* experiments detailed in Chapter 4 had revealed significant *in vitro* transcription activity of pR^{-}, despite giving a minimal background activity *in vivo*. No *in vitro* activity was seen for $pR^{=}$. Intrinsic *pL* activity was therefore measured in the presence of *pR* mutations which inactivated its activity both *in vivo* and *in vitro*. To ensure a consistency between promoter constructs in this project the double mutant was used in all subsequent promoter constructs that required an inactive *pR* region.

The 210 bp *pRpL* fragments were cloned into pBC2, single copy reporter strains were generated in NK7049 and lacZ activities assayed (fig 2.3(b)). Absolute values for intrinsic *pL* activity when using pBC2 and $pR^{=}$ were nearly double those using pMRR9 and pR^{-}, 81 ± 4 versus 42 ± 6 units. *pL* lacZ activity in the presence of active *pR* more than doubled from 6 ± 0.6 units using pMRR9 to 15 ± 0.6 units using pBC2. However the increase in *pR* lacZ activity was less than 2 fold, from 739 ± 59 to 998 ± 56 units, giving the difference in activities between *pRpL* and $pR^{=}pL$ closer to 10 fold. Interference of *pL* activity was calculated to be 5.3 ± 0.2 fold, which was lower than the previous value of 7 ± 0.4. The increase in intrinsic *pL* activity could be due to the use of $pR^{=}$ but is more likely to be caused by the removal of *trpt,* as an increase in activity is consistent with the removal of a weak transcriptional terminator expected to be about 40% efficient. Additionally, an equivalently large increase was observed for *pRpL*. The less than double increase in *pR* activity due the removal of *trpt* was unexpected and suggests that termination at *trpt* is inversely proportional to promoter strength. Accordingly, the decrease in interference measured using pBC2 compared to that measured using pMRR9 would reflect the increased termination of the weaker promoter by *trpt.*

2.3.3. Interference does not change between different host strains.

The possibility that different bacterial strains confer different levels of interference was suggested by the lower level of interference obtained in NK7049 (fig 2.3(a) and (b), and (Dodd and Egan, 2002)) compared to assays performed in MC1061.5 (Neufing *et al.*, 2001; and Reed *et al.*, 1997) (see fig 2.4). The effect of different host strains was examined by comparing the lacZ activities and interference of the single copy *pRpL* reporter from section 2.3.2. using either MC1061.5 or NK7049 as the host strain (fig 2.3(c) versus (b)). The activities for all promoters increased when measured in MC1061.5 compared to NK7049. The increase in *pL* activity was nearly 2 fold; intrinsic *pL* activity increased from 81 ± 4 to 142 ± 3 units and *pL* activity in the presence of active *pR* increased from 15 ± 0.6 to 26 ± 1 units. The increase in *pR* activity was not as great, from 998 ± 56 to 1340 ± 113 units. Interference of *pL* activity using MC1061.5 was calculated to be 5.6 ± 0.2 fold, which is very similar to the 5.3 ± 0.2 fold observed using NK7049. The choice of strain to measure promoter activities was therefore unimportant in terms of its effect on interference, however the use of MC1061.5 was preferable, as the accuracy of determining differences in activities between weak promoters is increased by the measurement of higher lacZ units compared to the same construct in NK7049. The increase in the activities of all promoters could be due to differences in *lacZ* transcription (possibly by differences in RNAP concentration), translation (including *lacZ* mRNA degradation rates) or activity of lacZ protein. MC1061.5 was used as the host strain for all subsequent lacZ assays (unless otherwise indicated).

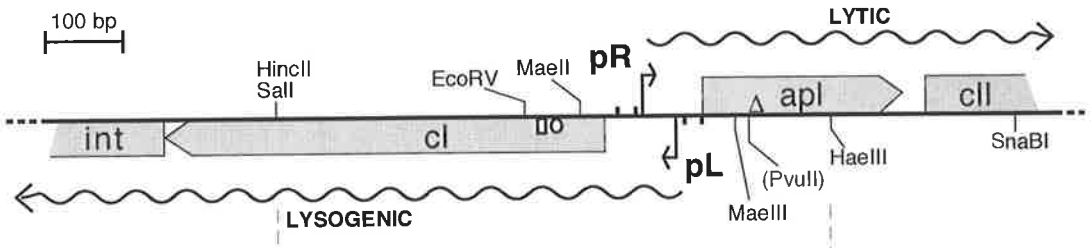
2.3.4. *pL* does not interfere with *pR* activity.

Previous studies reporting interference between *pR* and *pL* have not determined the possible influence of *pL* activity on *pR* activity. For a more complete analysis of the level of interference between *pR* and *pL*, rightward transcription was assayed in the presence of mutations in the -10 hexamer of *pL* (mutated from CATGAT to C**GC**G**CT** (Neufing *et al.*, 2001)) using the pBC2 vector and in strain MC1061.5 (fig 2.3(d)). *pR* activity increased from 1340 ± 113 units in the presence of active *pL*, to 1433 ± 161 units in the absence of active *pL*. This increase in activity was within the 95%

Figure 2.4

***pL* and *pR* activities and interference as measured by previous studies using different reporter systems.**

Details of the 186 genetic switch and the 186 fragments assayed in each study are drawn to scale. The legend for the switch region is as for fig 2.1, the Δ sign indicates the site of a 15 bp deletion within the *apl* gene which inactivates *apl* gene function. The circle represents the *CIts* mutation. In the table of 186 fragments assayed, numbers above the line indicate the position from the +1 site of *pR* and numbers below the line indicate the position from the +1 of *pL*. Fold interference is calculated as the ratio of leftward units in the absence and presence of *pR* activity. The fragments used for studying promoter strengths in my study are shown for comparison, for activities see fig 2.3.



Study Reporter system	Fold Interference	Leftward Units	186 Fragment assayed	Rightward Units
1. Dodd et al. (1990) galK plasmid based C600	6.6 **12.4** 82		SalI −432 pR pL HaeIII −190 EaeI +35 pL HaeIII −190	566
2. Reed et al. (1997) lacZ pMRR9 single copy λRS45 lysogens MC1061.5	9.8 **23** 230		EcoRV −155 pR pL PvuII −80 EcoRV −155 × pL PvuII −80	1600
3. Neufing et al. (2001) lacZ pMRR9 single copy λRS45 lysogens MC1061.5	8 **14** 115		HincII −432 pR Δ pL SnaBI −430 HincII −432 × Δ pL SnaBI −430	2275
4. Dodd and Egan (2002) lacZ pMRR9 single copy λRS45ΔYA lysogens NK7049	7 **~6** 41		SalI −432 pR Δ oo pL SnaBI −430 SalI −432 × Δ oo pL SnaBI −430	400
5. This study			MaeII −81 pR pL MaeIII −67 MaeII −81 × × pL MaeIII −67	

confidence limits of both *pR* activities, and indicates that the increase was not significant. Convergent *pL* activity does not cause reciprocal interference of *pR* activity.

2.3.5. Activities of $pR^{=}$ and pL^{-}

The promoter assay system utilising pBC2 and MC1061.5 as the host strain was used to determine the activities of the mutated promoters $pR^{=}$ and pL^{-}. To remove any potential interference from convergent transcription, promoter mutant activity was determined in the presence of a mutated convergent promoter. pBC2-$\underline{pR}^{=}pL^{-}$ and pBC2-$pR^{=}\underline{pL}^{-}$ were constructed and single copy lysogens assayed (fig 2.3(e)). Double mutations in *pR* (–10 and –35 region) reduced its activity over 1000 fold from 1433 ± 161 units to 1 ± 0.4 units, and single mutations in *pL* (–10 region) reduced its activity at least 50 fold from 142 ± 3 units to 5 ± 0.6 units. This indicates that the promoter mutations used in this study effectively inactivate rightward or leftward transcription from the 210 bp *pRpL* construct. Additionally, this shows that the only sequences involved in the *in vivo* transcriptional activity from this construct are *pR* and *pL*.

2.4 DISCUSSION

2.4.1 Measuring interference *in vivo*.

In this study interference of *pL* activity by *pR* transcription was calculated to be between 5.3 and 7 fold (fig 2.3) depending on the type of reporter system used. Previous published data using a variety of *in vivo* methods (summarised in fig 2.4) have reported measurements of interference between 6 and 23 fold. This discussion analyses the differences between the methods and attempts to find an explanation for this variability.

Comparisons with previous studies. Study 1 (fig 2.4) (Dodd *et al.*, 1990) measured the intrinsic activity of *pL* from a 225bp *EaeI-HaeIII* 186 fragment (from $+35_L$ to -190_L) which deletes the *pR* promoter, and compared this to the activity of *pL* from a larger 732bp *SalI-HaeIII* 186 fragment (from -482_R to -190_L) which contains active *pR*. (Note that $-\#_R$ is the number of bp upstream of *pR* +1 and $-\#_L$ is the number of bp upstream of *pL*.) DNA fragments were cloned into a vector containing a downstream promoterless *galK* gene. *E. coli* C600 transformed with the vectors were assayed for galactokinase

activity to quantitate the level of transcription originating from these DNA fragments. To account for variations in plasmid copy number, units were normalised to the plasmid DNA content of each strain assayed. A 12.4 fold reduction of *pL* transcription was seen when *pR* was present (82 versus 6.6 galK units). The use of a multi-copy vector to assay promoter activity is not ideal, nor is deletion of *pR* to measure intrinsic *pL* activity, as it does not account for the differences in the upstream galK sequence, leading to possible context affects. The use of single copy chromosomal reporters and mutations in *pR*, such as that employed in subsequent studies, is a more reliable method of assaying promoter activity.

While both the next two studies used single copy lacZ assays and *pR* mutations to measure intrinsic *pL* activity, both gave higher values of interference than that measured here. Study 2 (fig 2.4) (Reed *et al.*, 1997) assayed a 295 bp EcoRV-PvuII 186 fragment (from -155_R to -80_L) cloned upstream of the promoterless *lacZ* in the vector pMRR9. LacZ activity was assayed using a modified version of the λ-based single copy, chromosomal *lacZ* operon fusion system of Simons *et al.* (1987), with *E. coli* MC1061.5 as the host strain. A mutation in the –35 hexamer of *pR* (mutated from TTTACT to **CTCGAG**) was used to measure leftward transcription in the absence of *pR* activity. A 22.5 fold reduction in *pL* activity caused by active convergent *pR* transcription (250 versus 10 Miller units) was measured. Study 3 (Neufing *et al.*, 2001) used a similar single copy *lacZ* reporter system and the same *pR* mutant, but assayed a larger 972 bp Sal I-SnaB I 186 fragment (from -482_R to -430_L). A 14 fold reduction in *pL* activity caused by active convergent *pR* transcription (114 versus 8 Miller units) was measured.

Study 4 (Dodd and Egan, 2002), used the *lacZ* reporter plasmid pMRR9 and the same 186 fragment and *pR* mutant as Neufing *et al*, 2001 except that the λRS45 used to create single copy lysogens had a BsrGI/SgrAI deletion in the *lacY* and *lacA* genes and the host strain was *E. coli* NK7049. A kinetic lacZ assay in 96-well microtitre plates was used whereas the studies of both Reed *et al.* (1997) and Neufing *et al* (2001) used the single time point lacZ assay described by Miller (1972). Inactivation of *pR* was shown to improve *pL* transcription ~6 fold.

The present study (study 5 fig 2.4) has assayed the 210bp region of 186 DNA (from -81_R to -67_L) from single copy lacZ reporters using the semi-automated kinetic lacZ assay described in study 4. When using the reporter vector pMRR9, λRS45 and NK7049 as the host strain, inactivation of *pR* by mutation of the –35 region, was shown to improve *pL* transcription 7 fold (from 6 to 42 units). However, when using the reporter vector pBC2, λRS45ΔYA and either NK7049 or MC1061.5 as the host strain, inactivation of *pR* by both mutation of the –35 and the –10 regions was shown to improve *pL* transcription 5.3 or 5.6 fold respectively (from 15 to 81 units or from 26 to 142 units). Interference measured here is closer to that measured in study 4 but much less than that measured in studies 2 and 3. Variations in the fragment assayed, the type of promoter reporter system used and the type of assay performed are considered as possible explanations for these differences in interference.

The size of 186 DNA fragment assayed can not account fully for these differences, as the same sized fragment was used for study 3 and study 4. The same reporter vector, pMRR9, was used in all four studies, and therefore can not account for this variation. Although different *E. coli* host strains were used, this also can not account for variations as the present study has shown that interference is not affected by the host strain (fig 2.3(b) and (c)). One variation in the studies is the presence or absence of intact *lacY* (which encodes a permease gene) and *lacA* genes in the phage used to generate single copy chromosomal reporter. Increased interference is observed when λRS45 containing *lacYA*$^+$ was used (Study 2 and 3), whereas a decrease in interference was observed when using the modified λRS45ΔYA (study 4). Consistent with this observation, a decrease in interference was observed in the present study when λRS45ΔYA was used instead of λRS45 (7fold compared to 5.3 fold, fig 2.3(a) versus (b)). However the variation in interference observed here is better explained by differences in either the reporter vector used (pMRR9 versus pBC2) or the type of *pR* mutation used to measure intrinsic *pL* (*pR*$^-$ versus *pR*$^=$).

A major difference between the studies that report high levels of interference and those reporting lower levels is the type of lacZ assay performed. Higher levels of interference were calculated using an end point lacZ assay described by Miller (1972) to measure

promoter strengths (studies 2 and 3), whereas lower levels of interference were calculated using the 96-well format kinetic assay (Dodd *et al.*, 2001) (study 4 and this study). This latter method allows many repeats of the same experiment and the assay of lacZ activity is kinetic (30 readings at 2 minute intervals). Additionally, the current method uses polymixin B to lyse the cells whereas the older method uses chloroform. To test the possibility that the type of lacZ assay used could account for large variations in interference, I repeated the lacZ assays from study 3 using the same constructs but assayed using the 96-well kinetic assay (results not shown). Although interference was slightly reduced, from 14 fold to 11 fold (intrinsic *pL* activity was reduced from 78 units to 7 units in the presence of active *pR*), it was still higher than that of my study or study 4. Based on comparisons of the factors involved in measuring interference, it is not clear why there are large variations in the reported levels of interference. However it is also possible that the clones used in studies 2 and 3 were not completely confirmed by sequencing.

The system to be used here. From these studies it is difficult to assess the correct and 'real' level of transcriptional interference that occurs. It is possible that other techniques which directly examine the efficiency of promoter initiation, such as RNA protection assays, may achieve a more 'realistic' assessment of interference. Although it is difficult to assess the 'real' level of interference occurring between *pR* and *pL* by the methods presented here, this study does establish a method of determining promoter strength that is reliable, internally consistent, allows for an adequate statistical analysis of interference and can easily allow comparisons of the affect of different promoter constructs on *pL* activity.

2.4.2 Conclusions about interference between *pR* and *pL* activity.

The data presented in this chapter and in previous studies suggests a number of properties regarding the nature of interference between *pR* and *pL* that need to be accounted for when investigating the mechanism of interference between convergent promoters.

1. Interference must be at the level of transcription. This is because promoter reporter studies use transcriptional (or operon) fusions rather than translational (or gene) fusions

to study interference. The possibility of additional translational interference was investigated by Dodd and Egan (2002), using CI:lacZ protein fusions where both the transcription and translation of *lacZ* are dependent on signals from *pL*. Since interference was not substantially different to that seen with the operon fusions, it was concluded that the effect of *pR* on *pL* activity is at the level of transcription only and does not affect translation.

2. Only the sequences between the MaeII and MaeIII sites of 186 are responsible for interference. No additional interference is conferred when a larger promoter fragment which includes the weak IHF site found in the CI gene is used- compare results of fig 2.3 with those of study 4 (fig 2.4). Furthermore, as interference is relieved by mutations in the regions of *pR* that are responsible for promoter recognition by σ^{70}RNAP, then the negative effect of the wildtype Mae II to MaeIII region on *pL* activity must be dependent on the action of polymerase activity at *pR*.

3. Interference of *pL* activity is incomplete. Fig 2.3(c) shows that approximately ~18% of the transcripts normally produced from *pL* escape interference.

4. The interference of *pL* activity by *pR* is not reciprocal. Fig 2.3(d) shows that *pR* activity was unaffected by the presence of active *pL*. This result also indicates that the mechanism of interference affects *pL* activity without disturbing the transcriptional activity of the aggressive promoter. Elledge and Davis (1989) demonstrated a dependence on strong transcription from the aggressive promoter to achieve interference, therefore knowing that *pL* is about 10 fold weaker than *pR* activity, the non-reciprocal nature of interference was not surprising.

2.5 SUMMARY

A reliable *in vivo* system to measure transcriptional interference between *pR* and *pL* was established. This system assays the 210 bp 186 fragment (from –81R to –67L) using a semi-automated kinetic lacZ assay, the pBC2 vector, λRS45ΔYA to generate single copy *lacZ* operon fusions, MC1061.5 as the host strain, and a double *pR* mutation to measure intrinsic *pL* activity. Using this system, inactivation of *pR* was shown to improve *pL* transcription 5.6 fold. Four properties of this interference were determined

that will need to be incorporated into any proposed mechanism for how this interference occurs. The next two chapters further investigate the cause and mechanism of this interference.

CHAPTER THREE

What action of *pR* transcription causes interference?

Chapter 3

What action of pR transcription causes interference?

3.1 Introduction

Transcriptional interference is a mechanism of repression. Conventional repression involves the binding of a repressor molecule or complex in static equilibrium with its operator and thus affecting the activity of the promoter located within the sphere of influence of the repressor binding site (usually nearby). For 186 interference the 'operator' is a convergently orientated strong promoter sequence placed 62bp downstream from the repressed promoter, *pL*, and the repressor molecule is some unknown consequence of the activity of that 'operator'. In Chapter 2 it was shown that this interference can be abolished by mutations which severely reduce σ^{70}RNAP binding to *pR*. This implicates RNAP as a component of any 'repressor complex'. This chapter aims to detail precisely what aspect of RNAP activity at *pR* causes interference. Interference is expected to be a consequence of RNAP either binding to *pR*, or elongating from *pR*. Any proposed mechanism of *pR* interference will need to account for the non-reciprocal nature of interference (chapter 2) and the expected dependence on the strong RNAP activity of *pR* (Elledge and Davis, 1989). In this chapter alterations in the arrangement of *pR* and *pL* and the use of a transcriptional terminator between *pR* and *pL* in lacZ reporters show that the majority of interference *in vivo* was due to RNAP elongation over *pL*. Construction details of all vectors used in this chapter are given in the Materials and Methods, (Chapter 7).

3.2 Interference by *pR* bound RNAP

Three mechanisms are proposed whereby *pR* bound RNAP could reduce *pL* activity; 1) steric hindrance of RNAP binding at *pL*, 2) competitive inhibition of RNAP binding to *pL*, and 3) a roadblock to transcription from *pL*. These mechanisms are discussed here in detail.

Steric hindrance.

pR bound RNAP could influence *pL* activity by a mechanism similar to 'classical' repression, for example the lexA repressor binding over the –35 region of the uvrA promoter, (Bertrand-Burggraf *et al.*, 1987) where the binding of the repressor molecule to the promoter blocks the access of RNAP. Dnase I and hydroxlyradical DNA footprinting studies of the σ^{70}RNAP open complex formed at a promoter, measured *in vitro* on linear DNA, showed a DNA protection from approximately –55bp to +20bp from the start of transcription (Record *et al.*, 1996). Jagura-Burdzy and Thomas, (1997) have shown *in vitro* that RNAP can simultaneously occupy two convergent promoters spaced 37 bp apart but not closer (see the Introduction 1.4.4). Given a spacing of 62 bp between the +1 positions of *pR* and *pL* it was considered unlikely that the interference observed was due to the RNAP bound at *pR* precluding RNAP binding to *pL*. However, differences in DNA protection may arise *in vivo* in an environment of supercoiled DNA and molecular crowding. For such a mechanism the strength of *pR* was considered to be important, as open complexes would need to be formed very efficiently and continuously at the expense of open complex formation at *pL*.

Local sink mechanism.

A theoretical possibility is that the presence of the strong promoter *pR* might competitively inhibit nearby *pL* activation by acting as a local sink, sequestering RNAP molecules that might otherwise interact with *pL*. This type of inhibition is similar to other competitive enzyme inhibitors in that the presence of the inhibitor (*pR* DNA) acts to reduce the active free enzyme concentration. This type of promoter competition is dependent upon limiting concentrations of free cellular holoenzyme, as has been reported by (Shepherd *et al.*, 2001). Competition will also be influenced by how freely RNAP can diffuse between competing promoter sites. The need for a close spatial *in cis* location of the polymerase sink may be understood by a description for promoter recognition that involves 'sliding' along the DNA rather than simple diffusion. Evidence of sliding or one-dimensional diffusion is presented by a direct observation of the dynamics of nonspecific and specific RNAP-DNA complexes using scanning force microscopy (Guthold *et al.*, 1999).

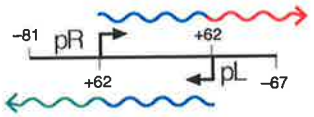
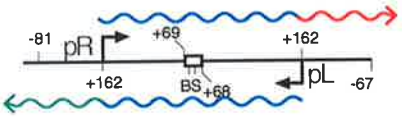
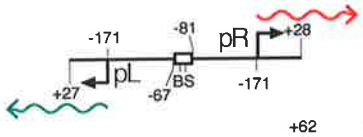
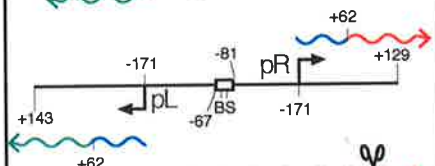
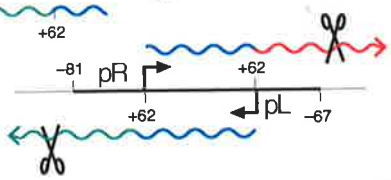
Roadblock mechanism.

pR bound polymerase could act to inhibit *pL* elongation by a 'roadblock' mechanism. Roadblocks to transcription were first described by (Deutschle *et al.*, 1986) who found that lac repressor bound downstream of a promoter blocks transcribing RNAP and terminates transcription. Hypothetically, if *pR* is consistently occupied by a bound RNAP molecule, than any polymerase initiating from *pL* would have to contend with this downstream obstacle which could impede elongation, in a similar way to the lac repressor, and so cause interference.

3.2.1. Steric interference by *pR* bound RNAP

To test whether steric hindrance contributes to the *in vivo* interference of *pL* activity, the level of *pL* interference was determined when the distance between the start points of *pL* and *pR* was increased from 62 bp to 162 bp. This distance is at least twice the length of an *in vitro* RNAP footprint (Record *et al.*, 1996), and should easily allow the co-binding of RNAP at *pR* and *pL* without steric interference. Any loss in interference would indicate a contribution of steric hindrance between the promoters when spaced the native 62 bp apart.

To gain this increased spacing, but preserve the original DNA context of each promoter, the *pR* fragment –81 to +69 was cloned in front of the *pL* fragment +68 to –67. Including a 25 bp spacer containing restriction sites, the resulting construct, *pL*(+100)*pR* (fig 3.1(b)), separates *pR* and *pL* by an additional 100 bp. To minimise any potential variation in *pL* lacZ activity, caused by differences in RNA translatability and decay, the intrinsic lacZ activity of *pL* was measured using an identical construct carrying the $pR^{=}$ mutation, *pL*(+100)$pR^{=}$. As shown in fig 3.1 the insertion of a further 100 bp between *pR* and *pL* had little impact on the intrinsic strength of *pL*, measured at 138 ± 10 units, compared to the 142± 3 units previously measured. Convergent *pR* activity decreased *pL* expression to 20 ± 1 units, and interference with *pL* activity was calculated to be 7.2 ± 0.2 fold. This was a significant increase from the interference measured for promoters spaced 62 bp apart, 5.6 ± 0.2 fold. Therefore, increasing the spacing between promoters did not decrease interference, indicating that interference was not due to steric hindrance between RNAP molecules.

Entry	Name	Fold Interference	Leftward $pR^{=}$	Leftward pR^{+}	Construct	Rightward
(a)	**pR pL**	**5.6** (± 0.2)	142(± 3)	26 (± 1)		1340 (± 113)
(b)	**pR (+100) pL**	**7.2** (± 0.3)	138 (± 10)	20 (± 1)		n.d.
(c)	**pL (div) pR**	**1.0** (± 0.06)	128 (± 5)	125 (± 4)		2403 (± 373)
(d)	**pL (div MM) pR**	**1.0** (± 0.1)	173 (± 6)	175 (± 14)		n.d.
(e)	**pBC3-pR pL**	**4.8** (± 0.3)	101(± 5)	21 (± 1)		1433 (± 161)

Legend

- RNaseIII Cleavage site
- vector DNA
- pL transcript
- pR transcript
- complementary transcript

3.2.2. *pR* acting as a local sink for unbound RNAP

If RNAP binding at *pR* acting as a competitive inhibitor for polymerase binding at *pL* is the major source of interference, then this mechanism should operate for 186 promoters arranged in a similar proximity but back-to-back. With regards to promoter concentrations, activity, spatial distance between promoters (in the nucleoid), orientation (with respect to the direction of transcription) and promoter strengths, a close divergent orientation of *pR* and *pL* will ensure that the extent of competitive inhibition is equivalent to that of convergent *pRpL* promoters. Divergent promoters (*pL*(div)*pR* and *pL*(div)$pR^{=}$, fig. 3.1(c)) were constructed to conserve the original DNA context of each promoter. However there was a concern that if divergent *pR* and *pL* promoters were constructed such that they each transcribe beyond +31 bp of wildtype DNA downstream of their start sites, then the transcripts produced would be complementary at their 5' end (the first 62 nt of transcripts produced from convergent *pR* and *pL* are naturally complementary). To account for possible interference caused by antisense transcription, divergent promoters were constructed such that the possibility of complementary transcription was avoided. To construct these back-to-back promoters, *pL* and *pR* promoter 'modules' were used. These modules contain the same upstream regions as previous constructs (ie. just outside the MaeII and MaeIII sites), but the downstream regions extend to the midway point between *pR* and *pL* (from the –10 regions), giving +28 and +27 from the start sites of transcription from *pR* and *pL*, respectively. The resulting divergent promoters obtained from these promoter fragments are separated by 171bp from the start sites of transcription.

The activity of the *pL* module without *pR* activity (*pL*(div)$pR^{=}$) was 128 ± 5 units (fig 3.1(c)), which was slightly lower than the values obtained using previous constructs, 142 ± 3 units (fig 3.1(a)). In contrast, the activity of *pR* increased nearly two-fold when the downstream *pL* region was deleted, from 1340 ± 113 to 2403 ± 373 units. This may reflect differences in 5' mRNA decay and translation of *pR-lacZ* transcripts, or a slightly negative effect of downstream DNA on *pR* transcription. For the purposes of this experiment, the transcriptional activity of *pR* in this divergent orientation would therefore be expected to retain the ability to act as a local sink. The activity of *pL* in the presence of active divergent *pR* transcription was 125 ± 4 units, which was within error of the measured intrinsic activity, 128 ± 5 units (fig3.1(c)). The influence of divergent

pR on *pL* was calculated to be 1.0 ± 0.06 fold, indicating that the activity of a strong, close, divergent promoter without a complementary transcript does not affect the activity of the weak divergent promoter. This suggests no role for a local sink mechanism of interference.

3.2.3. RNAP bound to *pR* as a roadblock for *pL* transcription.

If the roadblock mechanism caused the interference by *pR* of *pL*, then it would be independent of transcript elongation from *pR*. The role of elongation from *pR* was investigated in the next section where I show that *pR* elongation is necessary for interference. Any role for a roadblock mechanism contributing to the decrease in *pL* activity must therefore be minor.

3.3. Interference by *pR* elongation

The results of section 3.2. show that interference is not caused by RNAP bound at *pR* and must therefore be dependent on the movement of RNAP from *pR*. I reasoned that possible mechanisms of interference caused by *pR* elongation are: a) detrimental head-on collisions between elongating polymerase from both *pR* and *pL*, b) an antisense affect caused by hybridisation of the overlapping transcripts, and c) the consequence of the passage of converging RNAP across *pL*. This section describes the elucidation of which mechanism involving *pR* elongation is operating.

3.3.1. Head-on collisions.

If terminal head-on collisions between elongating polymerase molecules initiating from *pR* and *pL* were the major cause for interference then the extent of *pL* repression should be directly proportional to the distance between promoters. Increasing the inter-promoter distance should cause a proportional increase in the elongation time between promoters, and thus increasing the chance that an elongating polymerase from *pL* would encounter an elongating polymerase from *pR*. Increasing the separation of *pR* and *pL* from 62bp to 162bp, yielded a small increase in interference from 5.6 to 7.2 fold (fig. 3.1(a) versus (b)), which is less than the 14.6 fold expected if interference were completely caused by collisions $\left(5.6 \times \frac{162}{62}\right)$. Therefore, while head-on collisions were

expected to play some role, it was not the major contribution to interference with the closely spaced convergent *pRpL*.

3.3.2. Antisense RNA.

To test the possible inhibitory effect of the antisense *pR* transcript, *pL* activity was measured with and without the presence of its antisense transcript. Antisense *pL* was supplied from divergently placed *pR* which initially transcribes the natural 62 bp *pL* complementary transcript, (see constructs *pL*(divMM)*pR* and *p L*(divMM)*pR*$^{-}$; fig 3.1.(d)). *pL*(divMM)*pR* was constructed as follows: the 210bp Mae II- MaeIII fragment from *pR*$^{=}$*pL* was ligated to the Mae II-MaeIII *pRpL*$^{-}$ fragment to form *pR*$^{=}$*pL.pRpL*$^{-}$. The *pL* and *pR* transcripts from this construct diverge and do not overlap, but the first 62 bp of their respective transcripts are complementary. Supply of antisense transcription from an *in cis* divergent *pR* was expected to be at a level found for natural convergent *pRpL* transcription (ie. at least ~ 10 fold excess *pR* transcript to *pL*), and is also in close proximity to *pL* to allow ready access to its complementary sequence. The control was the construct of fig3.1(c), that is *pR*(div)*pL*, as the *pL* and *pR* transcripts from these constructs carry no complementary sequence. The control constructs previously showed that divergent, non-antisense, *pR* transcription does not affect *pL* activity. The intrinsic activity of *pL* in the presence of an inactive upstream *pRpL* fragment increased slightly to 173± 6 units, from previous measurements (142 ± 3 units, fig 3.1(d) versus (a)). In the presence of divergent, antisense, *pR* transcription the activity of *pL* was 175 ± 14 units (fig 3.1(d)). No interference was evident. Therefore, it was concluded that the presence of antisense transcripts alone, produced by divergent *pR*, do not cause the interference observed between convergently arranged promoters.

Interference could require the use of an overlapping transcript such that an elongating polymerase become 'tethered' to the DNA as a result of overlapping transcripts. A description of a potential 'tethering mechanism' follows. In the event of simultaneous elongation from *pR* and *pL*, the close proximity of the nascent 3'antisense transcripts could lead to the tethering of one RNAP to the other via the formation of a 62 bp double stranded RNA. The DNA template could no longer be threaded through each RNAP, as this would require the DNA template to be twisted in opposite directions. Instead each RNAP must wind around the template, which would lead to the restricted progress of

each polymerase. The stress created could be enough to break the contacts at the 3' end of the transcript with RNAP and terminate transcription. It should be pointed out that with the native *pRpL* *lacZ* constructs, the rightward transcript is terminated after approximately 330 bases, giving only a limited opportunity for this type of interference.

The influence of a tethering mechanism was tested by observing interference from constructs that contain a RNaseIII site downstream of *pL* but before the *lacZ* gene. Clipping of the transcript by RnaseIII at this position should release the *pL* transcript attached to RNAP from any tether formed by overlapping *pR* transcription. A lacZ reporter vector, pBC3, was constructed which is similar to the vector pBC2 except that it contains an RNaseIII cleavage site downstream of the MCS and upstream of the promoterless *lacZ* gene. The MaeII/MaeIII *pRpL* fragment was cloned into pBC3, similar to the construction of pBC2-*pRpL*, to produce pBC3-*pRpL*. The lacZ activity of single copy lysogens was assayed (fig 3.1(e)). *pR* activity in the presence of the RNaseIII site was assayed to be 1433 ± 161 units which was within the 95% confidence limits of previously measured *pR* activity, 1340 ± 113 (fig 3.1(e) versus (a). Measurements of the intrinsic level of *pL* activity however decreased from 142 ± 3 units in the absence of the RNaseIII site to 101 ± 5 units in the presence of the RNaseIII site (fig 2.1(a) versus (e)). In the presence of active *pR* transcription, *pL* activity was reduced to 21 ± 1 units with the inclusion of an RnaseIII cleavage site and interference was measured at 4.8 ± 0.3 fold. This small reduction in interference compared to the 5.6 ± 0.2 fold interference of *pL* activity in the absence of a RNaseIII cleavage site, indicates that a mechanism involving overlapping and tethered RNA makes only a minor contribution to the overall interference of *pL* activity. only contributes to a small proportion of the overall reduction in *pL* activity. Theoretically a tethering mechanism should also be possible for *pR*(divMM)*pL* constructs, the fact that no interference was observed suggests that overlapping transcription is also required for this mechanism.

3.3.3. Elongation over *pL*

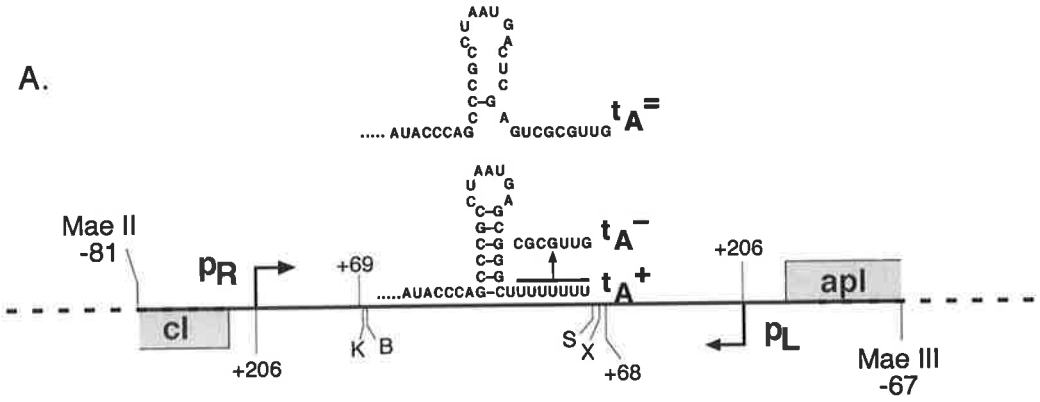
If interference occurs due to the progress of polymerase molecules from *pR* passing over *pL*, then blocking the progress of RNAP from *pR* before it reaches *pL* should prevent interference. Alternatively, if no change in interference is observed, then RNAP from *pR* must act before it reaches the vicinity of *pL*. To test this proposal an element

such as a transcriptional terminator was required that stops elongation from *pR* prior to transcription over *pL*, but has a minimal affect on *pL* activity, and is therefore assymetric.

The intrinsic terminator from the *trp* attenuator, (*trpA* or *t*A), was chosen. *trp*A has a common structure of a 6 bp G+C stem, 7 nt loop and an tract of 8 uracils immediately 3', with termination occurring at the last 2 Us (fig 3.2). Terminator efficiency was measured as 71% efficient *in vitro* and 85% efficient *in vivo* (Reynolds *et al.*, 1992). The bi-directionality of *trpA* has not previously been investigated, however *in vitro* termination of the structurally similar *thrA* terminator, was 78% efficient in its normal orientation but only 7% in its reverse orientation (Yang *et al.*, 1995). This indicates, at least *in vitro*, that termination from this type of sequence is essentially uni-directional. To generate the promoter constructs, a 49 bp sequence carrying tA was inserted centrally into *pR*(+100)*pL*, to yield *pR*(tA)*pL*. With at least 70 bp between the start site of either promoter to the point of termination, it was anticipated that termination of *pR* elongation or pausing of *pL* elongation would not interfere with initiation events at *pL* or *pR*. As increasing the distance between converging promoters by 100 bp led to a minimal increase in interference (fig 3.1(b)), the introduction of a further 49 bp of *trpA* DNA between *pR* and *pL* was not expected to significantly influence interference by simple spacing differences. As a control for the influence of the introduction of tA sequence, *pR*(tA^{-})*pL* constructs were made where the U tract of tA was mutated from UUUUUUUU3' to UCGCGUGU3' which was expected to prevent termination (Reynolds *et al.*, 1992) (Steiner and Malke, 1997). The termination efficiency of *pR* transcription by tA was assessed using constructs measuring *pR* activity placed into vectors containing a RNaseIII site between the promoter cloning site and the start of the *lacZ* gene. This site reduces context affects (Linn and Pierre, 1990), allowing a better comparison between different *pR* clones.

The introduction of mutated tA (tA^{-}) had a minimal impact on the activity of intrinsic promoter activity. *pR* activity was 1374 ± 137 units (fig 3.2B(a)) and is best compared with the activity of *pRpL* in the presence of the RNaseIII site, 1433 ± 161 units (fig 3.1(e)).This indicates that no significant termination is occurring. Intrinsic *pL* activity in the presence of the stem loop was 150 ± 6 units (fig 3.2B(a)), which is equivalent to

A.



B.

Entry	Fold Interference	Leftward pR$^=$	Leftward pR$^+$	Construct	Rightward pBC3-pR$^+$(tA?)pL$^+$
(a)	**9.2** (± 0.5)	150 (± 6)	16 (± 1)	pR ⌜(t_A^-)⌟pL	1374 (± 137)
(b)	**9.8** (± 0.8)	134 (± 6)	14 (± 1)	pR ⌜($t_A^=$)⌟pL	1508 (± 139)
(c)	**3.3** (± 0.1)	114 (± 5)	35 (± 1)	pR ⌜(t_A^+)⌟pL	408 (± 27)

previous measurements of intrinsic *pL* (fig 3.1). Despite these relatively unchanged intrinsic activities, *pR* decreased *pL* activity to 16 ± 1 units, much lower than previously observed. Interference was calculated to be 9.2 ± 0.5 fold, which was larger than that observed for *pL*(+100)*pR* constructs, 7.2 ± 0.3 fold (fig 3.1(b)).

It was possible that this unexpected increase in interference was a consequence of either the increase in the distance between promoters, or the introduction of stem loop between *pR* and *pL* creating additional interference. The transcribed stem-loop of tA and tA^- is similar in structure from either direction and, once transcribed, it primarily functions to slow down the rate of addition of the next nucleotide. The possibility exists that hairpin formation leads to the transcriptional pausing of *pL* initiated complexes, (although it appears that U-tract transcription may also be necessary for slowing polymerase and providing additional time for hairpin formation within the transcription complex (Gusarov and Nudler, 1999)). Pausing could create an artificial level of interference where elongation complexes from *pL* pause, then dissociate from the DNA after collisions with an oncoming polymerase from *pR*. Interference could also be created by the accumulation of *pR*-derived paused and terminating structures hampering the progress of polymerase from the direction of *pL*. To determine whether this increase in interference was a consequence of pausing at the tA stem loop or simply due to the increased spacing, the control experiments were repeated using constructs containing a double mutation of tA, *pR*($tA^=$)*pL*, where both the U-tract sequences and the stem-loop sequences have been altered (fig 3.2 A). Promoter activities were measured and are summarised in fig 3.2B(b). The mutation of the stem loop increased *pR* activity from 1374 ± 137 units to 1508 ± 139 units, indicating that a small amount of pausing by the stem loop may be occurring. Intrinsic *pL* activity was 134 ± 6 units indicating that the stem loop slightly increases the lacZ activity of *pL*. The activity of *pL* in the face of active *pR* was still low at 14 ± 1 units, and interference was calculated to be 9.8 ± 0.8 units. This indicates that spacing rather than the stem loop was responsible for the increased interference. The best control for the affect of active termination on interference was construct containing the single tA mutation, as this more closely resembles constructs containing an active tA terminator.

The presence of active tA reduced rightward transcription by 73% (or ~3.7 fold) from 1508 ± 139 units to 408 ± 27 units (fig 3.2(b) versus (c)), indicating that termination of rightward transcription at tA^+ was most likely occurring. The measurement of intrinsic *pL* activity was also reduced in presence of active tA from 134 ± 6 to 114 ± 5 units. This reduction may be a consequence of tA reducing the amount of *pL* elongation or alternatively a consequence of potential differences in the translation or stability between *lacZ* transcripts. Active convergent *pR* transcription in the presence of active tA termination decreased *pL* activity to 35 ± 1 units, and interference was calculated to be 3.3 ± 0.1 fold (fig 3.2B(c)). Therefore, reducing rightward elongation over *pL* by 3.7 fold decreased interference of *pL* activity from 9.2 fold to 3.3 fold, or nearly 3 fold (fig. 3.2B compare (b) with (c)). The 3.3 fold interference is compatible with a 27% readthrough of the terminator region from *pR* elongation causing 27% of the original 9.2 fold interference ie 2.5 fold interference. This implies that elongation from *pR* must pass over *pL* to generate a majority of interference, and that this type of mechanism is likely to be the major contributor to overall interference. Furthermore, this experiment confirms that there is minimal interference generated by a mechanism that involves *pR* bound polymerase as either an inhibitor of *pL* binding or as a roadblock for *pL* transcription, or by mechanisms that involve elongation for the 70 bp prior to termination.

3.4. Discussion

This chapter aimed to determine what aspect of RNAP activity at *pR* causes interference. The proposed model for interference is that elongation from *pR* over *pL* is the major contributor to the 5.6 fold interference of *pL* activity. Consistent with this model were the following observations.

Increasing the spacing between the promoters an extra 100 bp brought no loss in interference but rather a slight increase. This is inconsistent with a steric hindrance model of interference where RNAP bound at *pR* inhibits RNAP binding at *pL*. Placing the promoters in a close divergent orientation completely abolished interference, which is contrary to a competitive inhibition model of interference. Divergent promoters that also actively transcribe the 62 nt of 5' antisense transcript did not restore interference,

contrary to a model of interference that involves antisense transcription. Clipping of the *pL* transcript by RNaseIII after transcription of the *pR* promoter region and prior to transcription of the lacZ gene, demonstrated only a small decrease in interference. This is contrary to a major role of an antisense 'tethering' mechanism that requires antisense as well as overlapping transcription. It is possible that interference by antisense RNA requires that the complementary sequences in fact directly overlap (ie. the antisense transcripts are transcribed from the same DNA location). However in the *pR*(tA$^+$)*pL* construct (described below), all the *pR* transcripts, including those terminated at tA, show at least 70 bases of antisense and overlapping sequence with the *pL* transcripts that have extended past tA. If overlapping antisense transcript was required for interference then interference in the *pR*(tA$^+$)*pL* construct (3.3 fold) should have been as high as the native *pRpL* construct (5.6 fold), which has 62 bp of overlapping and antisense sequence. The significant reduction in interference caused by the placement of a unidirectional intrinsic terminator between *pR* and *pL*, is also contrary to both a mechanism of interference where RNAP bound at *pR* acts as a roadblock to elongation from *pL*, and also a mechanism of interference caused by the changes in DNA topology when RNAP polymerase binds to a promoter and begins elongating. This is also inconsistent with a major role for collisions between elongating RNAP. It is however consistent with models of interference which involve elongation over the sensitive promoter. The introduction of a transcriptional terminator between convergent promoters has also been successfully used in a number of eukaryotic systems to prevent interference (Ingelbrecht *et al.*, 1991; Padidam and Cao, 2001; and Prescott and Proudfoot, 2002).

The residual 3.3 fold interference in the *pR*(tA$^+$)*pL* constructs would not be expected if all the interference of *pL* activity was due to elongation from *pR* passing over *pL*, unless termination of *pR* transcription is not 100% efficient. Termination was however shown to be inefficient, such that 408 ± 27 units or 27% of *pR* transcripts are still able to pass over *pL* and effect interference. If it were true that all of the interference was caused by elongation over *pL* and that the extent of interference was directly proportional to the amount of transcription over *pL*; then as each unit of *pR* activity gives 9.8/1508 fold interference, 408 units of *pR* activity should give 2.7 fold interference. This is close to

the 3.3 fold interference measured, however the assumption of a linear relationship between promoter strength and interference has not been experimentally tested. An explanation of inefficient termination as the source of this residual interference could be tested by the use of a more efficient or second terminator between *pR* and *pL*. If interference is not further reduced, this would suggest a minor contribution of other types of interference, presumably head-on collisions between elongating polymerase, collisions with terminating complexes at tA, a roadblock mechanism, a tethering mechanism, and/or supercoiling effects. Head-on collisions are still able to occur between *pL* elongating polymerase and either uninhibited *pR* transcription occurring 70bp prior to termination or reduced *pR* transcription occurring 70bp after termination. Some of these additional mechanisms are unique to the *pR*(tA)*pL* clone which has promoters spaced 206bp apart and are unlikely to have a large impact for the natural *pRpL* convergent promoters spaced 62 bp apart. These include collisions with terminating complexes at tA, and additional supercoiling effects. If transcribing complex from *pR* produces waves of positive supercoiling ahead and negative supercoiling behind, termination of these complexes prior to transcription over *pL* would lead to an increase in positive supercoils reaching *pL* which would not normally be present during naturally spaced convergent *pRpL*.

Although elongation over *pL* is an adequate model for the major mechanism of interference for naturally spaced *pRpL*, this does not explain the increase in interference observed when the spacing between promoters was increased, without increasing *pR* transcription (fig 3.1(a), (b) and fig 3.2B(b)). Measurements of intrinsic *pL* activity remained basically unchanged for promoters spaced 62 bp apart (*pRpL*), 162bp apart (*pR*(+100)*pL*) and 206bp apart (*pR*(tA$^{=}$)*pL*) (142 units, 138 units and 134 units respectively), however the action of *pR* reduced *pL* activity proportionately with increased spacing (now 26 units, 20 units and 14 units respectively). The amount of interference was calculated to be 5.5 fold for *pR*-62bp-*pL*, 7.2 fold for *pR*-162bp-*pL* and 9.8 fold for *pR*-206bp-*pL*. This suggests that as the spacing between convergent promoters increases, so does the influence of other interference mechanisms that are dependent on spacing. For widely spaced promoters these other mechanisms are expected to become the dominant source of interference superseding that of elongation over the sensitive promoter. Possible mechanisms that would be dependent on spacing

include head-on collisions between elongating polymerases, a tethering mechanism, and supercoiling affects. The influence of head-on collisions for widely spaced promoters is further discussed in the general mechanism section of the final discussion (Chapter 6). For interference between *pR* and *pL* spaced 62 bp apart, these mechanisms contribute only a minor role.

3.5. Summary

This chapter has focussed on one half of the interference story, namely what causes the *pR* dependent interference of *pL*. Elongation from *pR* over *pL* was been shown to be the major contributor to the 5.6 fold interference of *pL* activity. The next chapter investigates the nature of interference caused by elongation over *pL*.

CHAPTER FOUR

Nature of the interference caused by the passage of RNAP across *pL*

Chapter 4

Nature of the interference caused by the passage of RNAP across pL

4.1 Introduction

4.1.1. Possible explanations for the reduction of *pL* activity caused by the passage of polymerase over this region.

Chapter 3 showed that elongation over *pL* was necessary to display interference. This chapter details the nature of this type of interference. Two possible explanations exist, either the passage of RNAP molecules initiated at *pR* restricted the access of other RNAP molecules attempting to bind *pL* (termed 'promoter occlusion' (Adhya and Gottesman, 1982)) or, converging, elongating RNAP molecules from *pR* collide with and disrupt initiation intermediates formed at *pL* (termed here the 'sitting duck' mechanism) (see fig 4.1). The blocking of access to the promoter during promoter occlusion can result from direct steric hindrance or from distortion of DNA structure. The influence of this mechanism during *pR* interference of *pL* activity, is dismissed on theoretical grounds in chapter 6. The discussion also includes alternative explanations of interference involving transcriptional induced changes in DNA topology or the binding of other host proteins that can affect promoter activity. This chapter is concerned with finding evidence that bears on a sitting duck mechanism of interference.

4.1.2. The sitting duck mechanism of interference.

If occlusion was the only mechanism of interference, it would imply that any polymerase molecules that manage to bind *pL* between waves of *pR* transcription would be uninhibited by the next converging polymerase and go on to produce full-length *pL* transcripts. This would happen if either the speed of *pL* initiation was fast than the time between waves of blocking polymerase or if initiation steps of *pL*, subsequent to promoter binding, were unaffected by elongation from *pR*. Failing this leaves the second proposed model on interference, termed the 'sitting duck' mechanism.

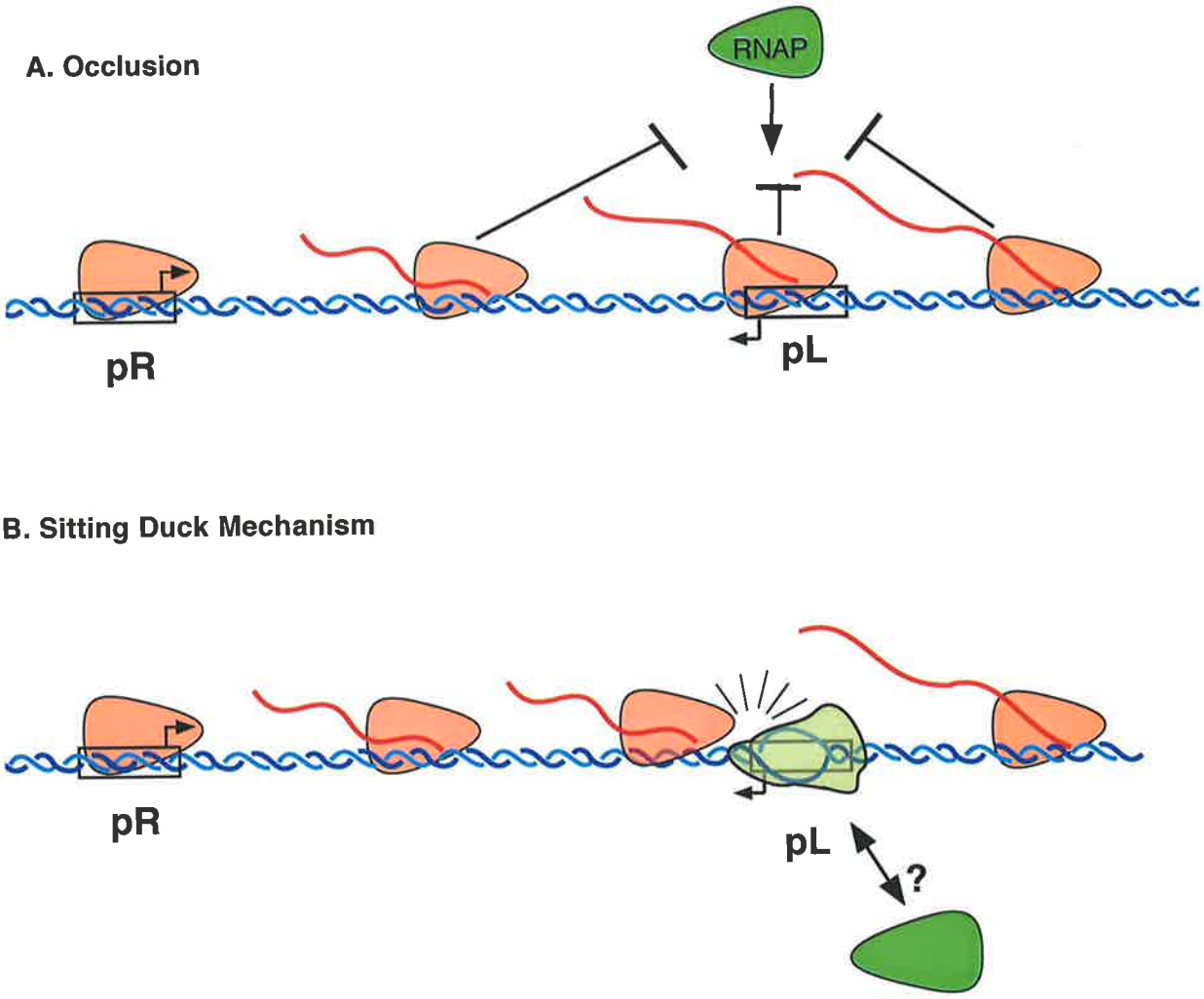


Figure 4.1 Cartoon of possible interference mechanisms involving elongation over *pL*.

A. During occlusion access of free RNAP to *pL* is blocked by transcription from pR.
B. During a sitting duck mechanism initiation intermediates formed at *pL* are involved in detrimental collisions with elongation from *pR*. These collisions may or may not lead to dissociation from the DNA.
For more information see text.

Simplistically one can think of the events of promoter initiation as: (i) promoter recognition where free holoenzyme and promoter DNA bind to form closed complexes, (ii) isomerisation where closed complexes open the DNA around the –10 region to form open complexes, and (iii) promoter clearance where open complexes go through the process of promoter escape to eventually elongate away from the promoter releasing the DNA ready for re-initiation (see section 1.4.1). A weak promoter such as *pL*, can be characterised as having reduced promoter recognition (weak binding), slow isomerisation and/or slow clearance. The model is that an RNAP slow to progress through initiation is a 'sitting duck' for interference from a convergently transcribing RNAP.

The likelihood of a sitting duck mechanism of interference to describe interference at *pL* was examined using an *in vitro* transcription assay to determine, (a) whether the initiation properties of *pL* should allow for the accumulation of promoter initiation intermediates ('sitting ducks') which can be 'hit' by elongating polymerase from *pR*, (b), whether collisions between these sitting ducks and converging RNAP from *pR* could reduce *pL* activity, and (c) whether multiple rounds of collision events could account for a 5.6 fold level of interference.

4.2 Results

4.2.1. Establishing an *in vitro* transcription assay system.

An *in vitro* transcription assay was developed which could measure comparative, overall initiation kinetics under conditions which would best simulate the *in vivo* environment used to measure a 5.6 fold interference of *pL* activity. The assay system required an appropriate DNA template, specific reaction conditions (including RNAP concentration) and methods of manipulation such that rates of different initiation steps could be measured. This section describes the establishment of this system.

DNA template.

To be consistent with the previous lacZ reporter studies, the DNA fragments to be assayed were the same 210bp (MaeII to MaeIII) *pRpL* fragments and contained the same promoter mutations, where appropriate. The promoter fragments were to be

inserted into a plasmid upstream of a transcriptional terminator. Plasmid DNA purified from cells is expected to be naturally, negatively supercoiled, and this negative supercoiling has been shown to be important in the activities of a number of promoters (eg. Dorman, (1995) and references therein). Assaying promoter activity from naturally supercoiled plasmid DNA would give a closer representation of the cellular promoter behaviour of *pL* responsible for its sensitivity to interference.

An *in vitro* transcription vector was required which had intrinsic transcriptional terminators placed either side of a multi-cloning site, so that when the *pRpL* DNA was inserted into the site, the transcripts from both *pR* and *pL* would be terminated at sizes that could be easily distinguished on a denaturing 6% polyacrylamide gel. It was important that transcription from these promoter inserts would be insulated from transcription occurring elsewhere in the vector which might interfere with the promoter activity of the insert. For this purpose the *in vitro* transcription vector, pBC1, was designed and constructed (fig 4.2, described in Chapter 7 'Materials and Methods'). pBC1 is a modified pUC19 based high-copy number plasmid. The lacZ promoter region has been removed to reduce the number of transcripts produced from the vector. The pUC19 multiple-cloning site is flanked by tandem repeats of the *rrnB* terminator region carrying the strong intrinsic terminators, T1 and T2. This terminator region was expected to be efficient at terminating transcription *in vitro* (Reynolds *et al.*, 1992) and would allow for a transcriptionally isolated cloning region available for insertion of promoter sequences. The distances between the middle of the *XbaI* site (used for cloning promoter inserts) and the site of terminations were 142bp and 111bp. This asymmetry was such that insertion of the 210bp *pRpL* region into the *XbaI* site of pBC1 in a particular orientation (checked by PCR and shown in fig 4.2) produced transcripts from *pR* and *pL* of sizes 243nt and 288nt respectively, which could easily be distinguished by electrophoresis on a denaturing 6% polyacrylamide gel (fig 4.2).

pR⁻pL, *pRpL*, and *pRpL⁻* regions were cloned into pBC1 (fig 4.2). To assess the ability of these pBC1 clones to report on *pL* and *pR in vitro* activity and also to assign the identity of transcript bands, a simple single round *in vitro* transcription assay was performed on each supercoiled template, as well as pBC1 alone, and transcript patterns obtained (fig 4.3). Controls reactions containing no DNA or no RNAP were included

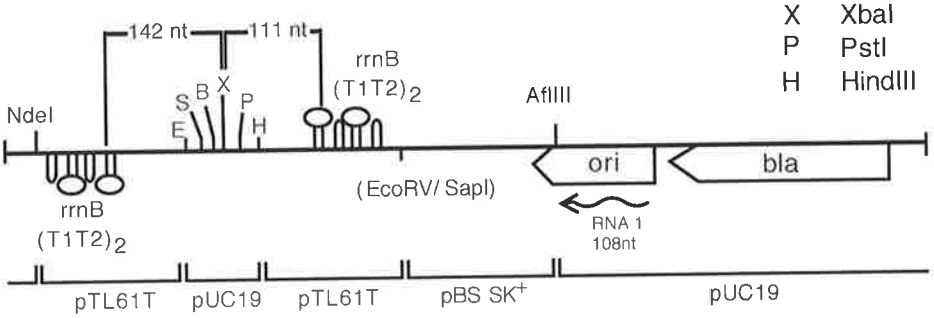
Figure 4.2

Templates for *in vitro* transcription.

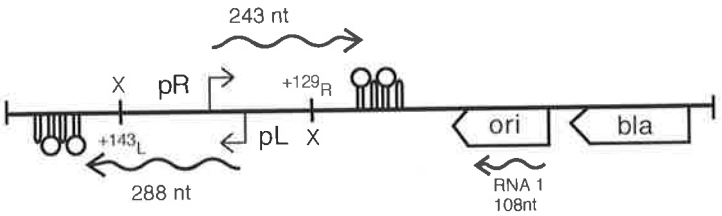
A diagram of the supercoiled pBC1 templates used *in vitro*. Details about the construction of pBC1 are shown and described further in Materials and Methods. The different distances between the centre of the *XbaI* cloning site (the site used for cloning *pRpL* DNA) and the expected point of termination at the flanking T1 terminators are shown. The pBC1-*pRpL* templates are also shown, giving the orientations and expected transcript sizes for each inserted promoter. RNA1 is also shown.

pBC1 (4006bp)

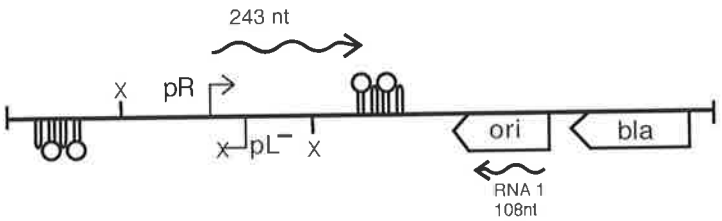
E EcoRI
S SmaI
B BamHI
X XbaI
P PstI
H HindIII



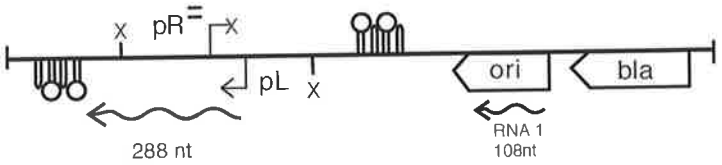
pBC1-pRpL



pBC1-pRpL⁻



pBC1-pR⁼pL



pBC1-pR⁻pL

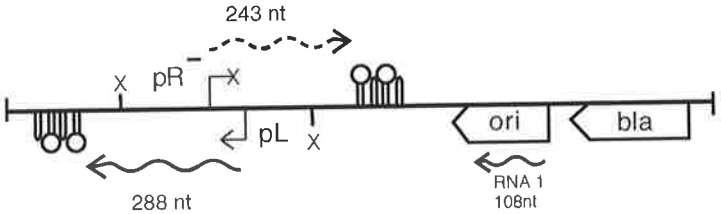
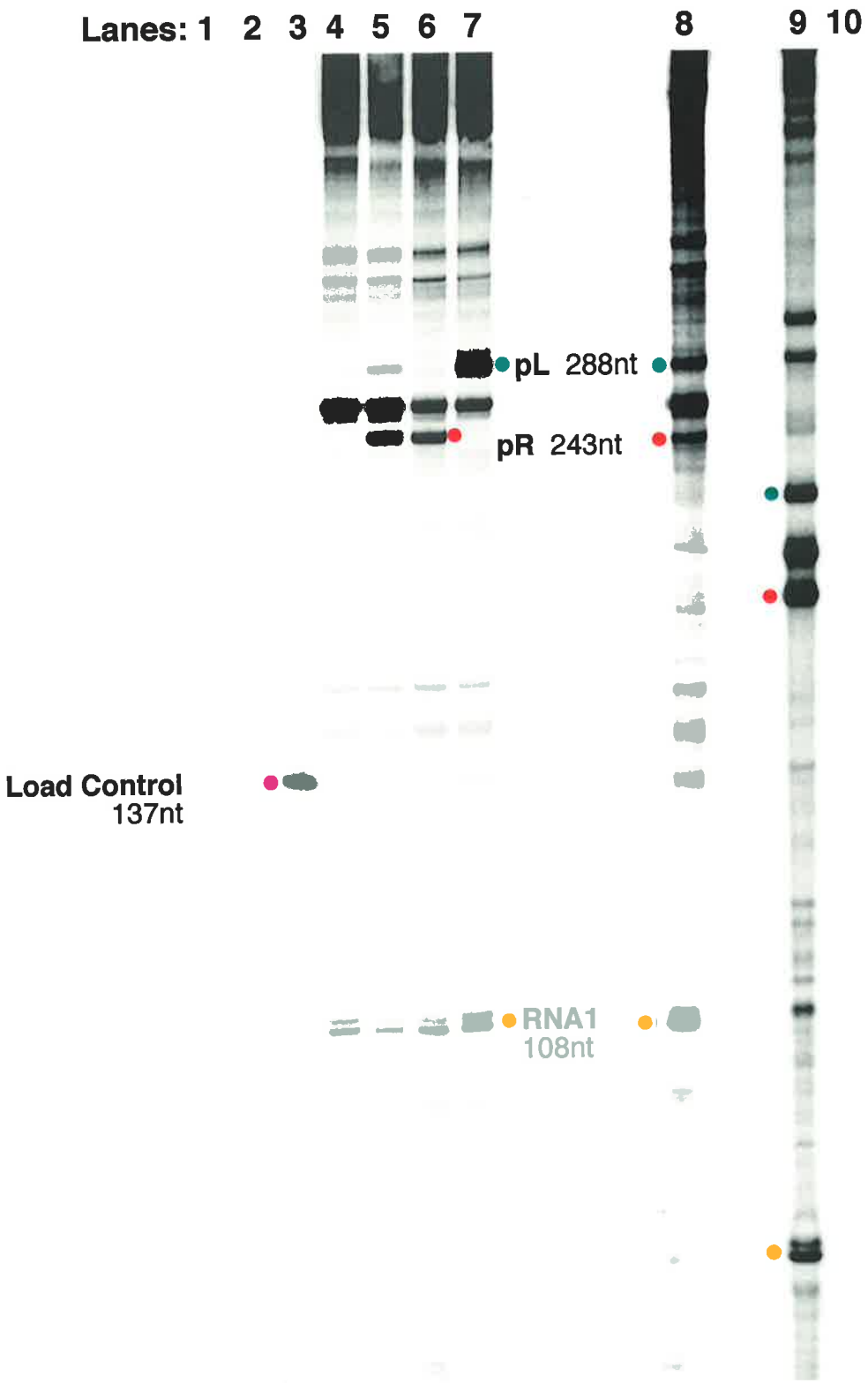


Figure 4.3

***In vitro* transcription of pBC1 templates.**

In vitro transcript pattern obtained for templates pBC1 only, pBC1-*pRpL*, pBC1-*pRpL*$^{-}$, pBC1-*pR*$^{=}$*pL* and pBC1-*pR*$^{-}$*pL* (lanes 4 to 8 respectively). Transcript patterns are from single round *in vitro* transcription assays, and were performed in 10 μl volumes for each template, using the conditions described in Materials and Methods. RNAP and DNA was pre-incubated for 30 minutes to allow maximum open complex formation at all relevant promoters, NTP/heparin solution was then added and elongation allowed to continue for 60 minutes to allow complete firing and elongation from all relevant promoters before termination with stop/load buffer. Lanes 1 to 3 are controls. Lane 1 is a control reaction with no DNA, lane 2 is a reaction with no RNAP and lane 3 is the 137nt labelled PCR product used as a loading control for normalising between reactions in subsequent experiments. Lanes 9 and 10 are the *in vitro* transcription patterns from reactions containing pBC1-*pRpL* and testing the ability of heparin to prevent transcription. Standard reactions were set up but reactions were initiated by adding DNA and heparin (final concentration 5mg/ml) to RNAP and NTPs in transcription buffer. Reactions were stopped after 60 mins. Lane 9 was a control containing no heparin. Grouped lanes (ie 1 to 7, 8 and 9 to 10) were from experiments performed on the same day.


Lanes: 1 2 3 4 5 6 7 8 9 10
pL 288nt
pR 243nt
Load Control
137nt
RNA1
108nt

and gave no background bands. Transcript sizes were initially determined with the aid of labelled pUC19 HpaII DNA markers which were run alongside the transcription reactions (not shown). pBC1 alone showed a large number of background transcripts, with the most significant transcripts being the 108 nt RNA1 transcript (seen as a doublet) produced from the plasmid origin of replication which is often used as an internal control during *in vitro* transcription (as it is here) (Richet and Raibaud, 1991). Other transcript include a longer ~260 nt transcript, and a number of larger transcripts some of which presumably include those initiated elsewhere in the replicon and from the *bla* gene promoter. Unique bands of the expected size for full-length transcripts from *pR* and *pL* were observed for the transcription pattern using the pBC1-*pRpL* template. The origins of these transcripts were confirmed by the absence of the *pL* band in the pattern for pBC1-*pRpL*$^{-}$ and the absence of the *pR* band in the pattern for pBC1-*pR*$^{=}$*pL*. Following electrophoresis of the transcripts on a 6% polyacrylamide gel, the bands for full-length *pL* and *pR* transcripts are adequately separated from each other and from background transcripts to allow quantitation of *pL* and *pR* transcripts.

Initially it was intended that pBC1-*pR*$^{-}$*pL* be used to measure intrinsic *pL* activity, which contains a mutation in just the –35 region of *pR*. However, despite the low activity of *pR*$^{-}$ *in vivo* (Reed *et al.*, 1997), *in vitro pR*$^{-}$ was found to produce nearly as many full-length transcripts as *pL* (fig 4.3). To avoid the possibility of *in vitro* interference by *pR*$^{-}$ activity on 'intrinsic' *pL* activity, *pR*$^{-}$ was further mutated by changing the –10 region from TATATT to TCGATC to create *pR*$^{=}$ (as described in Materials and Methods). The *in vitro* activity of *pR*$^{=}$ was undetectable (fig 4.3), indicating that pBC1-*pR*$^{=}$*pL* could be used as a template to measure intrinsic *in vitro* activities of *pL*. The *in vivo* activity of *pR*$^{=}$ was reported in chapter 2 and most constructs measuring *in vivo* interference in this study have contained *pR*$^{=}$.

A qualitative inspection of the transcript patterns for pBC2-*pRpL* and pBC2-*pR*$^{=}$*pL* from fig 4.3 showed that in this single round assay, the amount of full-length *pL* transcripts was reduced in the presence of active *pR* (lane 5 versus 7). This indicated

some level of interference occurring *in vitro*, which was investigated further in section 4.2.4.

Reaction conditions.

Reactions were performed essentially as described by (Ryu *et al.*, 1994), using a transcription buffer containing ion concentration that closely resemble that found *in vivo* (Leirmo *et al.*, 1987). Reactions were always performed at 37°C, the temperature used to grow the *E. coli* strains being assayed for lacZ activity. Attempts were made to find RNAP concentrations that were saturating, such that when attempting to simulate interference a maximum occupancy of *pR* and *pL* would ensure that the maximum amount of collisions occur. The activity of pBC1-*pRpL* was analysed over a range of polymerase concentrations. Unfortunately, at high levels of polymerase, specific promoter activity appeared to decrease as non-specific background activity increased (results not shown), a phenomenon observed by others (Shanblatt and Revzin, 1984). Therefore the concentration of RNAP that gave maximal specific activity for *pR* and *pL* was used. This was found to be a 25 molar excess of RNAP to DNA template. Although this may be a near saturating level for *pR* and *pL* it may not be saturating for all promoters.

Using the transcription assay to measure promoter kinetics.

In order to manipulate the *in vitro* system such that kinetic promoter properties could be measured, the transcription reaction was restricted to a single round by the addition of heparin salt. Heparin is a poly-anion that mimics DNA and presumably inhibits transcription by interfering with the formation of a specific RNAP-DNA complex at a promoter, inhibiting the formation of open complexes from free RNAP or from closed complexes, but not the activity of formed open complexes (Schlax *et al.*, 1995). During single round transcription involving heparin, RNAP is pre-incubated with the DNA template in transcription buffer without NTPs such that elongation can not occur. With the simultaneous addition of heparin and NTPs to this transcription reaction, all open complexes that have formed (heparin resistant complexes) will begin elongating and produce transcripts. The presence of heparin will prevent any closed complexes producing transcripts and any new rounds of promoter initiation. The overall rate of full-length transcript production for one round of transcription can thus be separated

into two events: the rate of formation of heparin resistant complexes at a promoter (which includes the events of DNA recognition plus closed and open complex formation) and the rate of production of full-length transcript from heparin resistant complexes (this is equivalent to the rate of escape, clearance and elongation up to the first T1 terminator). Measuring these two rates would indicate the potential for *pL* to accumulate 'sitting ducks' in the form of either open complexes waiting to clear or in the form of initiation intermediates prior to the formation of open complexes (possibly closed complexes waiting to isomerise).

Before measuring these rates of open complex formation and promoter clearance it was necessary to first find a concentration of heparin that rapidly prevented open complex formation at *pL* and *pR*. A concentration of heparin of 5 mg/ml was chosen. To test whether heparin quickly prevents open complex formation the DNA template pBC1-*pRpL* was incubated with heparin and added to a mixture of RNAP and NTPs in transcription buffer, and elongation allowed to occur for 60 minutes. A no heparin control was also performed. Results are shown in lanes 9 and 10 of fig 4.3. The transcript pattern showed no transcripts from *pL* or any background transcripts from pBC1, indicating that 5 mg/ml of heparin was sufficient to rapidly prevent open complex formation at *pL* and any pBC1 promoters. A very faint band corresponding to full-length *pR* transcript was observed; the band was still present when a heparin concentration of 10 mg/ml was used (results not shown). The *pR* band is very faint compared to that seen without heparin, indicating that although open complex formation at *pR* is efficient enough to outcompete some heparin activity most of the activity at *pR* activity is rapidly prevented.

As the promoter rate assays would be relying on accurate quantitation of the intensity of bands between different lanes, a loading control was included that could be used to normalise the amount of transcription reaction loaded in each lane. The loading control was a 137 bp labelled PCR product generated using primers KS and RSP and the template pBluescriptKS+, which was added to the stock of stop/loading buffer used to stop the transcription reaction prior to loading on the gel. Fig 4.3 lane 3 shows the band produced by the loading control.

4.2.2. The clearance rates at *pL* and *pR*.

The rate of full-length transcript production from heparin resistant open complexes was determined using the transcription assay developed in the previous section. Full-length transcript production involves not only the processes of promoter escape but also that of elongation. Given the size of *pR* and *pL* transcripts are about 250 nt and the expected elongation speed *in vitro* is about 40-50 nt/sec (Richardson and Greenblatt, 1996), it was assumed that the rate of elongation was negligible (unless pause sites exist or the rate of clearance is quicker than 5 seconds) such that the rate of full-length transcript measured in these experiments would be equivalent to the rate of promoter escape (or clearance). To measure this rate of clearance, heparin resistant complexes of RNAP and DNA (open complexes) were allowed to form for 30 minutes (to ensure maximum open complex formation), and upon the addition of rNTPs and heparin, the rate of appearance of full-length transcripts was determined by taking aliquots of the elongation reaction at different times then analysing and quantitating the transcript pattern for each aliquot. The clearance rate of the weak *pL* promoter was measured using pBC1-*pR*=*pL* and compared to the clearance rate of the strong *pR* promoter measured using pBC1-*pRpL*$^{-}$. Examples of the transcript patterns obtained for each template in these clearance assays are shown in fig 4.4. Bands for full-length transcripts were quantitated and normalised to the loading control for each lane (as described in the Materials and Methods, Chapter 7), then plotted against time. The rate of accumulation for each promoter was found to fit the first order exponential:

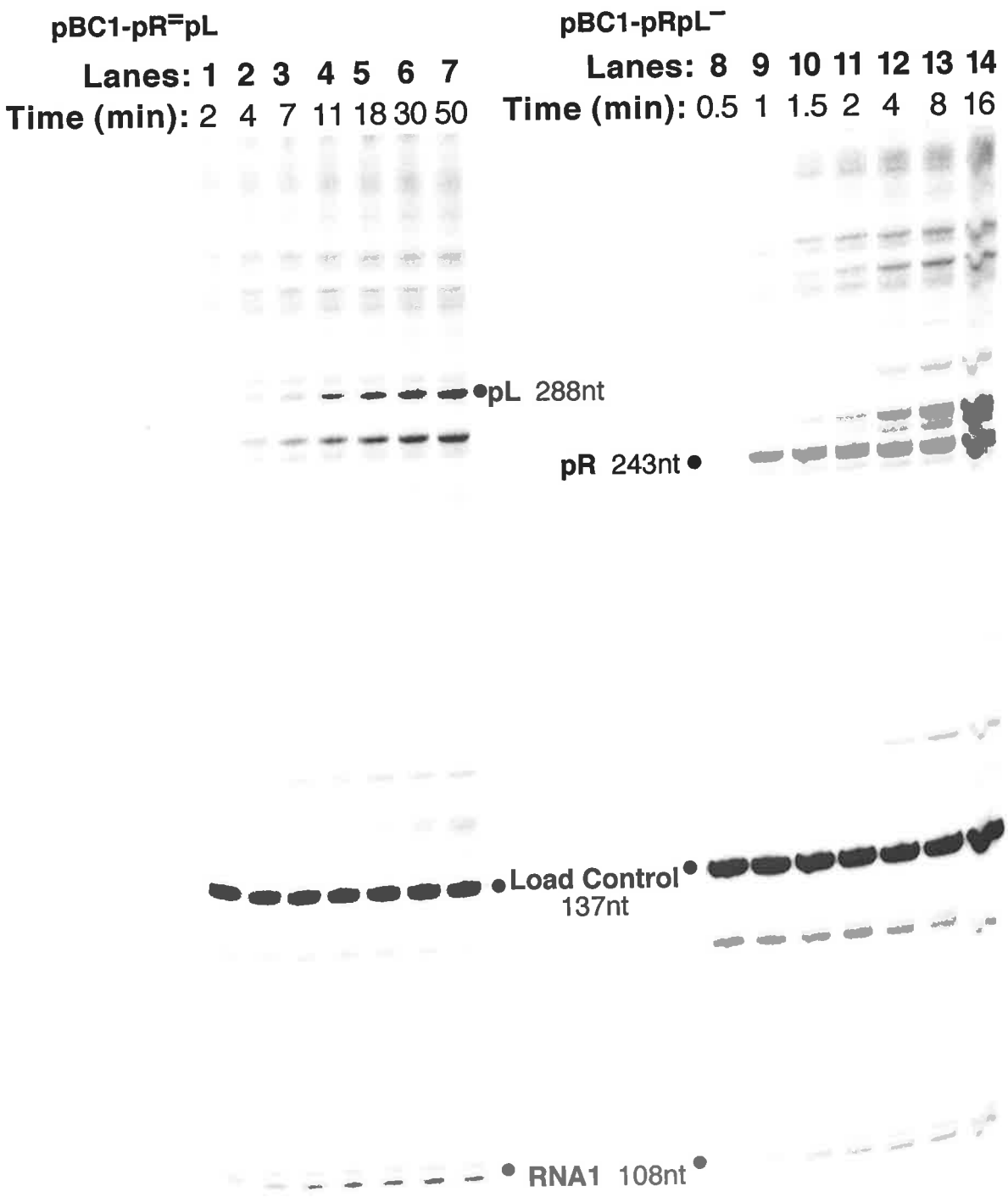
$$\%\,\mathbf{max} = a(1 - e^{-bt}) \qquad \textbf{Equation 4.1}$$

where %max is the amount of full-length transcript expressed as a percentage of the maximum, a is the maximum amount of transcript accumulated, b is the rate constant for clearance from the promoter assayed and t is the clearance time. For each individual experiment the value of a was determined by curve fitting (SigmaPlot v4), the data was recalculated, setting a=100% and plotted against clearance time (fig 4.5). Data from two independent experiments were combined and then fitted to equation 4.1, using a=100%, to define b for that promoter. From this value, the time for half of the maximum transcript production ($t_{50\%}$) was calculated and used to compare the clearance rates of

Figure 4.4

Transcript pattern showing the intrinsic rate of full-length transcript accumulation from *pL* and *pR* open complexes *in vitro*.

An example of the transcript pattern used to quantitate the rate of accumulation of the transcripts *pL*($pR^{=}$) (lanes 1 to 7) and *pR*(pL^{-}) (lanes 8 to 14) from heparin resistant complexes formed at templates pBC1-$pR^{=}pL$ and pBC1-$pRpL^{-}$ respectively. *In vitro* transcription escape assays were performed in 60 μl volumes for each template. DNA and RNAP were pre-incubated for 30 minutes to ensure complete/stable open complex formation at all relevant promoters, then elongation was initiated by the addition of NTP/heparin mix at time zero. At various times after the addition of NTPs, 7 μl aliquots of the reaction were added to 7 μl stop/load buffer (which also contained the loading control). 10 μl of each time point were then loaded onto a 6% polyacrylamide gel. Lanes 1 to 7 correspond to a transcription reaction terminated 2, 4, 7, 11, 18, 30 and 50 minutes after the addition of NTPs, and lanes 8 to 14 are from a separate transcription reaction terminated 0.5, 1, 1.5, 2, 4, 8 and 16 minutes after the addition of NTPs, respectively. These are examples of the patterns from one of two experiments used for quantitation in fig 4.5.


pBC1-pR=pL
Lanes: 1 2 3 4 5 6 7
Time (min): 2 4 7 11 18 30 50
pBC1-pRpL−
Lanes: 8 9 10 11 12 13 14
Time (min): 0.5 1 1.5 2 4 8 16
pL 288nt
pR 243nt
Load Control
137nt
RNA1 108nt

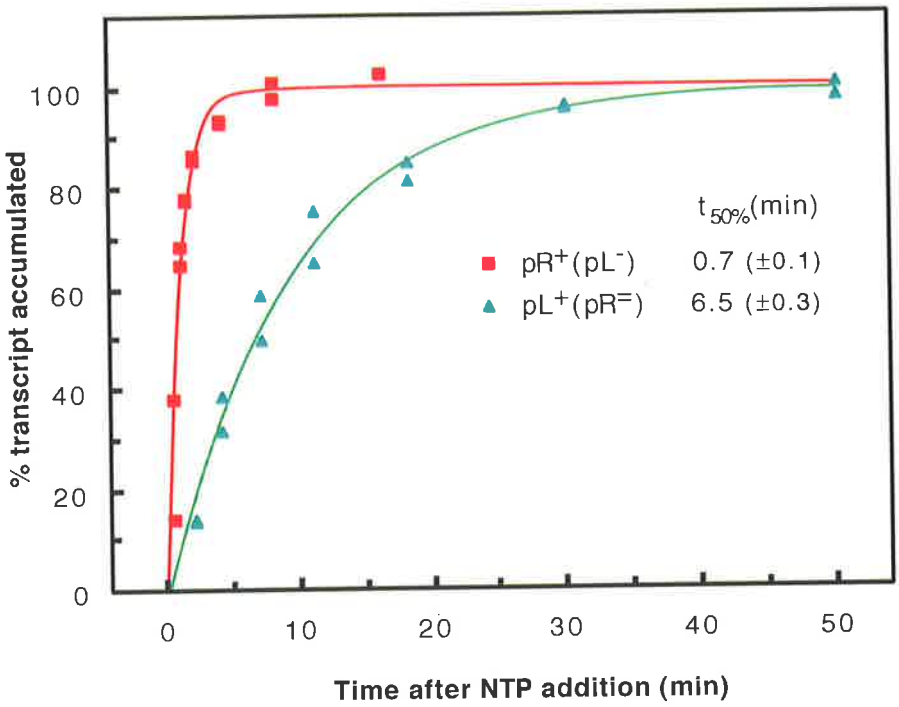


Figure 4.5. Graph showing the intrinsic rate of full length transcript accumulation from *pL* and *pR* open complexes in vitro.

Bands for full length transcripts $pR(pL^{-})$ (red squares), and $pL(pR^{=})$ (green triangles) from the transcript patterns of clearance assays described in fig 4.4 were quantitated and plotted against elongation time. The amount of each full length transcript accumulated was quantitated as described in Materials and Methods, normalised to the amount of loading control in each lane and then plotted as a percentage of the maximum level of specific full length transcript accumulated in that reaction (as described in the text, section 4.2.2). Values for $t_{50\%}$ are shown. Examples of the transcript patterns from one of these reactions were shown in figure 4.4.

each promoter (shown in fig 4.5). Uncertainties in $t_{50\%}$ were calculated from the standard errors in *b*.

The $t_{50\%}$ for *pL* was 6.5 ± 0.3 minutes and for *pR* was 0.7 ± 0.1 minutes. The rate for RNA1 production was closer to that of *pR* with a $t_{50\%}$ of ~ 1 minute (not shown). Thus the rate at which heparin resistant complexes produced full-length transcripts was over 6 fold slower for *pL* than other promoters. This reflects a much slower release of open complexes formed at *pL* compared to those formed at *pR* or *pRNA1*. These results suggest *pL* has a high potential to form 'sitting ducks', in the form of open complexes waiting to clear.

4.2.3.The rate of open complex formation at *pL* and *pR*.

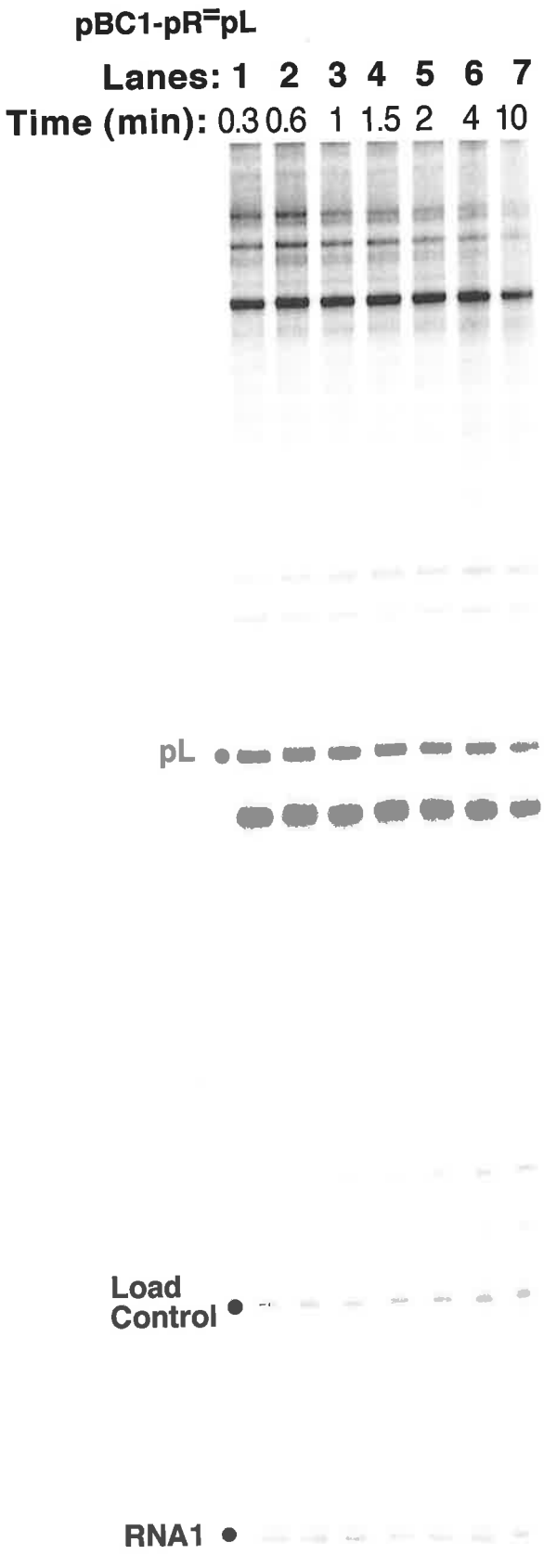
The rate of formation of active open complexes at *pL* from free DNA and RNAP was measured in this system, to know whether 'sitting duck' type interference at *pL* was more likely to occur at the level of collisions with accumulated initiation intermediates formed prior to the formation of active open complexes rather than after. To assess the rate at which heparin resistant complexes (open complexes) formed under these conditions, reactions containing RNAP and template DNA (pBC1-*pR*=*pL*) were incubated at 37°C and aliquots of the reaction were taken at various times and added to the NTP/heparin mix and allowed to elongate for 60 minutes (a time which would allow all active heparin resistant complexes formed to produce full-length transcripts, based on the results in fig 4.5). A transcription pattern was obtained for each reaction (fig 4.6) and bands for full-length *pL* and RNA1 were normalised to the amount of loading control in each individual lane. Normalised units were expressed as a % of the maximum level of transcript accumulated for that promoter and plotted against the time of DNA:RNAP incubation (fig 4.7).

At the earliest time point (20 seconds), over 80% of the maximum amount of active open complexes at *pL* had formed and over 90% at the RNA1 promoter. The maximum levels were reached at about 1 minute and this rise to a maximum. Unexpectedly, after 1 minute the amount of 'active' open complexes decreased, decaying exponentially until, at 10 minutes, just under 60% of the maximum level of open complexes at *pL* were now active and just under 80% at *p*RNA1. This may be due to either dissociation of open

Figure 4.6

Transcript pattern showing the rate of formation of heparin resistant complexes (or open complex formation) that are able to produce transcripts from pBC1-*pR*=*pL*.

Transcript patterns are from a (final volume) 90 μl transcription reaction containing transcription buffer and template DNA, pBC1-*pR*=*pL* which was warmed to 37°C. At time zero, RNAP was added to a final concentration of 50nM and open complex formation was allowed to proceed. At various time points the amount of open complex formation was assayed by taking aliquots of the reaction which were added to a NTP/heparin mix giving a final volume of 10μl. Elongation was allowed to proceed for 60 minutes prior to the addition of 10 μl stop/load buffer containing loading control. 10μl of the elongation reactions were loaded onto a 6% polyacrylamide gel and the transcript pattern was analysed. Incubation times of DNA and RNAP prior to the addition of heparin and NTPs (pre-incubation time) were 20 secs, 40 secs, 1 min, 1.5 min, 2 min, 4 min and 10 minutes (lanes 1 to 7, respectively).


pBC1-pR=pL
Lanes: 1 2 3 4 5 6 7
Time (min): 0.3 0.6 1 1.5 2 4 10
pL
Load Control
RNA1

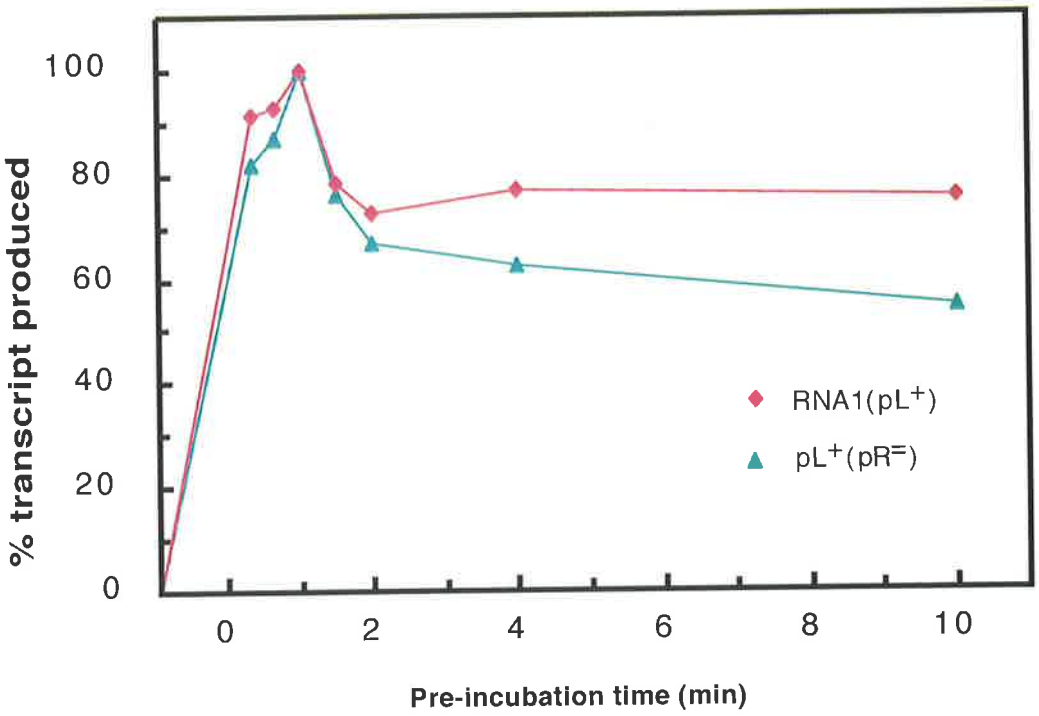


Figure 4.7 Graph showing the rate of formation of heparin resistant complexes (or open complex formation) that are able to produce transcripts from pBC1-*pR*$^{=}$*pL*.

Bands for full length transcripts for *pL*(*pR*$^{=}$) and RNA1 from the transcription pattern shown in fig 4.6 were quantitated and plotted against pre-incubation time. Band intensities were normalised to the amount of loading control in each lane and is expressed as a percentage of the maximum full length transcripts accumulated (in this case the band intensities at 1 min pre-incubation for both *pL* and RNA1).

complexes with time or inactivation of some complexes (discussed further in section 4.3.1.). It was possible that the formation of maximal levels of active open complexes before the first time point reflects a slow action of heparin rather than rapid open complex formation, however previous experiments testing the speed of heparin activity had indicated that the action of heparin is very rapid (section 4.2.1). This experiment showed that, when using conditions that attempt to be equivalent to those used when measuring *in vivo* interference, 80% of the maximal level of open complexes take less than 20 seconds to form from free RNAP and DNA. These open complexes then clear the promoter and produce full-length transcripts with a half-life of 6.5 minutes. The rate of clearance of these open complexes is thus much slower than the rate of formation, indicating that *pL* most likely accumulates open complexes *in vivo* that become sitting ducks for collisions with elongation complexes from *pR*.

4.2.4. Open complexes formed at *pL in vitro* are sensitive to elongation from *pR*.

A sitting duck mechanism of interference for *pL*, involving collisions with elongating RNAP from *pR* and open complexes formed at *pL* which are slow to clear, is only relevant if these collisions are detrimental to *pL* activity. The potential 'damage' of collisions between heparin resistant complexes formed at *pL* and elongating polymerases from *pR* was determined by using the previously described *in vitro* transcription assay system. Based on the relative rates of transcript production from *pR* and *pL* determined in section 4.2.2. (fig. 4.5) it was expected that over 80% of the complexes formed at *pR* would fire before even 10% of the complexes at *pL* had been released. Therefore, if heparin resistant complexes were formed on pBC1-*pRpL* and elongation allowed to occur, a large proportion of the open complexes formed at *pL* would be hit by elongation from downstream *pR* open complexes. To examine the outcome of these collisions, the amount of final full-length transcript produced from both *pR* and *pL* was compared with the amount of final full-length transcript produced from *pR* and *pL* in the absence of collisions. For the experiment, reactions containing the templates pBC1-*pRpL*, pBC1-*pRpL*$^{-}$ or pBC1-*pR*$^{=}$*pL* were performed by first adding RNAP at a 25 fold molar excess to DNA, and allowing open complex formation to occur for 4 minutes (enough time to form stabilised amounts of active open complex at both *pR* and *pL*, see fig 4.7). NTPs and heparin were then added and elongation allowed to continue for 60 minutes, to ensure that all active open complexes from *pL*

would produce full-length transcripts. The transcript pattern for each reaction was obtained (fig 4.8) and bands of full-length transcript for *pR*, *pL* and RNA1 were quantified for each reaction, then normalised to the uracil content of each transcript. To compare the amounts of *pL* and *pR* transcript produced from each reaction, transcription units were normalised to the amounts of RNA1 transcript produced in that reaction (which is expected to be relatively constant between templates and reaction) and finally expressed as a percentage of the amount of RNA1 transcript (table 4.1).

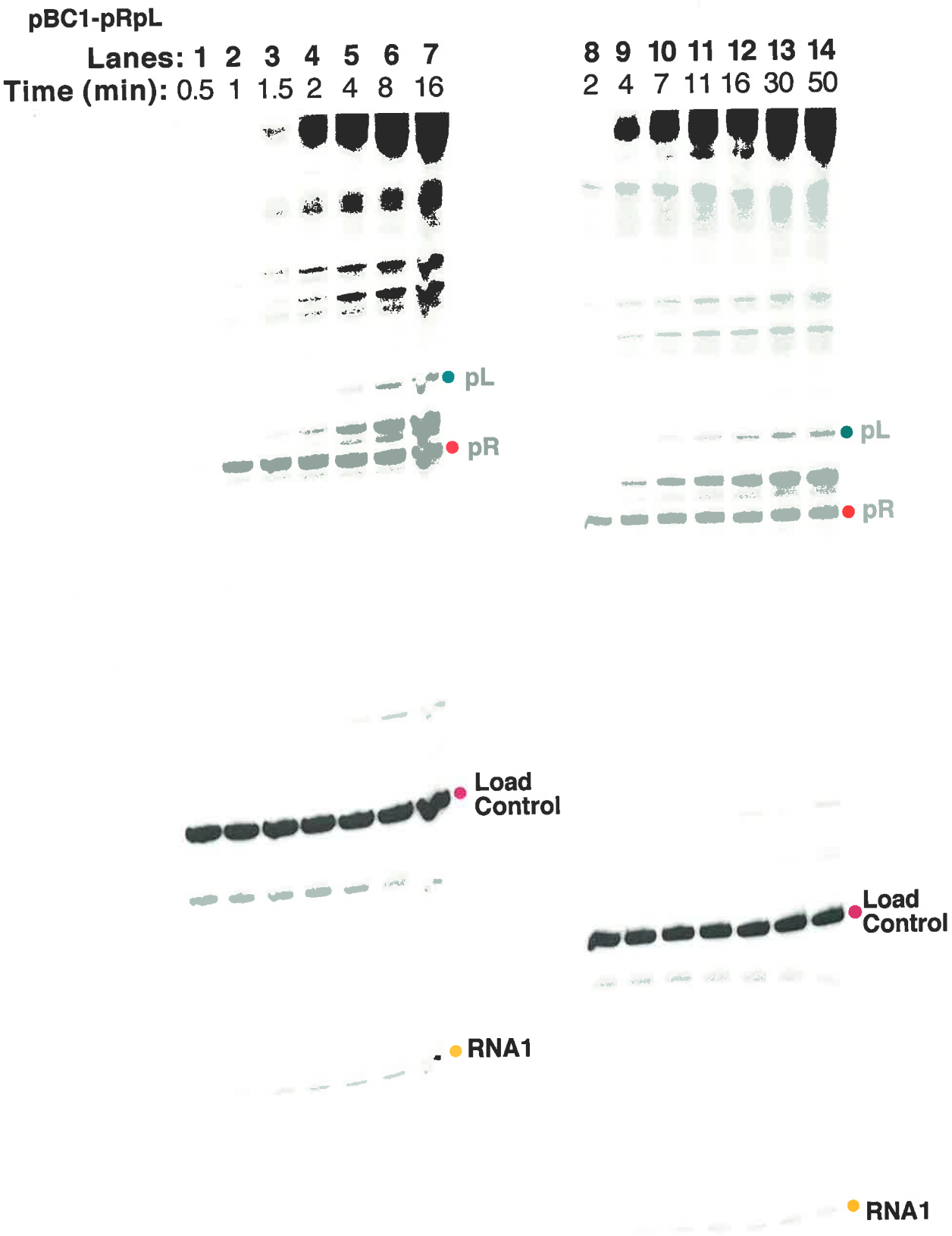
The amount of transcript produced from *pR* in the absence of *pL* activity was 169 ± 9% (of RNA1), and did not change significantly in the presence of active *pL* (165 ± 4%). This suggests that these collisions with *pL* open complexes do not affect the final progress of the elongating polymerase involved in the collision. The amount of *pL* transcript produced was reduced from 100 ± 8% to 66 ± 4% in the presence of active *pR*. This indicates that collisions with converging RNAP are detrimental to full-length transcript production from one third of the heparin resistant complexes.

If all promoters were saturated and active in the reaction then the amount of transcript produced for RNA1, intrinsic *pL* and intrinsic *pR* should be the same after normalising to the amount of U incorporated in each transcript. This appears to be the case for *pL* and RNA1 as *pL* transcripts were 100% that of RNA1 however *pR* produced over 60% more transcripts than either *pL* or *pR*RNA1. This indicates either inactivation of open complexes at *pL* and *pRNA1* or non-saturating conditions. Since the fraction of the open complexes at *pR* promoters that produce an elongating polymerase is not known, it is not possible to know what proportion of open complexes formed at *pL* are subject to a collision with an elongating polymerase from *pR*. As one third of *pL* open complexes are inactivated by collisions, a difference in the number of collisions at *pL* would indicate different sensitivities to collisions. If the fraction of *pR* open complexes that fire was as low as 33% then it is estimated that every *pL* complex exposed to a collision was inactivated. If all of the *pR* open complexes fired, then only one in three *pL* complexes were inactivated by collision. This experiment clearly shows that collisions with open complex 'sitting ducks' at *pL* are detrimental to their activity *in vitro*, causing inactivation of at least one third of all open complexes formed at *pL*.

Figure 4.8

Transcript pattern showing the effect of one round of convergent transcription on the amount of full-length transcript accumulated from *pR* and *pL* open complexes.

RNAP and DNA were incubated using standard reaction conditions for 4 minutes prior to the addition of NTPs and heparin. Elongation was allowed to continue for 60 minutes then stopped with stop/load buffer (with loading control). 10μl of the elongation reactions were loaded onto a 6% polyacrylamide gel and the transcript pattern was analysed. Reactions were repeated 4 times. Shown are the transcript patterns, for the DNA template pBC1-*pRpL*$^{-}$, pBC1-*pR*$^{=}$*pL*, pBC1-*pRpL*, lanes 1 to 4, 5 to 8 and 9 to 12 respectively, which were used for quantitation in table 4.1.


pBC1-pRpL
Lanes: 1 2 3 4 5 6 7
Time (min): 0.5 1 1.5 2 4 8 16
8 9 10 11 12 13 14
2 4 7 11 16 30 50
pL
pR
pL
pR
Load Control
Load Control
RNA1
RNA1

Promoter	Number of full length transcripts (% of RNA1)
pL(pR–)	100 ± 8
pR(pL–)	169 ± 9
pL(pR+)	66 ± 4
pR(pL+)	165 ± 4

Table 4.1
Quantitation of the number of full length transcripts produced from pR and pL after a collision.
Bands for full length transcripts for pL, pR or RNA1 from fig 4.8 were quantitated as described in materials and methods, normalised to the loading control in each lane and also the U content of each transcript, then expressed as a percentage of the number of transcription units calculated for RNA1 in each lane. The values are the averages and standard deviations from four experiments.

It was possible that collisions not only inactivate a third of *pL* open complexes, but also reduce the clearance rate of those complexes not inactivated. Additionally, pausing of elongation from *pR* during a collision would be expected to reduce the *rate* of full-length transcript production from *pR* even though the final *amount* of *pR* full-length transcripts was unaffected. To assess whether head-on collisions between open complexes and elongating RNAP cause polymerase at either *pL* or from *pR* to pause at the site of collision, (possibly until the release of open complexes from the promoter), the rates of full-length transcript accumulation for *pL* and *pR* from pBC1-*pRpL* were examined in the presence of convergent transcription. An example of the transcript pattern obtained for these clearance assays is shown in fig 4.9. Full-length transcripts for *pL* and *pR* were quantitated as before, plotted against elongation time and the data used to find the exponential relationship for the rate of clearance (fig 4.10). The graphs for transcript accumulation were nearly identical to those seen for *pR* and *pL* in the absence of collisions (shown in fig 4.10 for comparison). The calculated $t_{50\%}$ from *pR*(pL^+) was 0.9 ± 0.1 minutes which is only slightly longer than the $t_{50\%}$ measured for *pR*(pL^-) of 0.7 ± 0.1 minutes (see fig 4.10). This indicates that transcription from *pR* is mostly unaffected by collisions with open complexes at *pL*. Transcript production from the *pL* open complexes that produced full-length transcripts after collisions with elongation from *pR* occurred at a $t_{50\%}$ of 6.2 ± 0.4 minutes which is equivalent to the rate of production of *pL* transcripts without *pR* activity, ie. 6.5 ± 0.3 minutes. This implies that the *pL* transcripts escaping interference by transcription from *pR* were uninhibited in their rate of initiation. Thus, initiation intermediates at *pL* are either completely unaffected by collision or are rendered completely incompetent to produce full-length transcripts. At the most 67% of *pL* initiation intermediates are unaffected and a minimum of 33% are rendered incompetent.

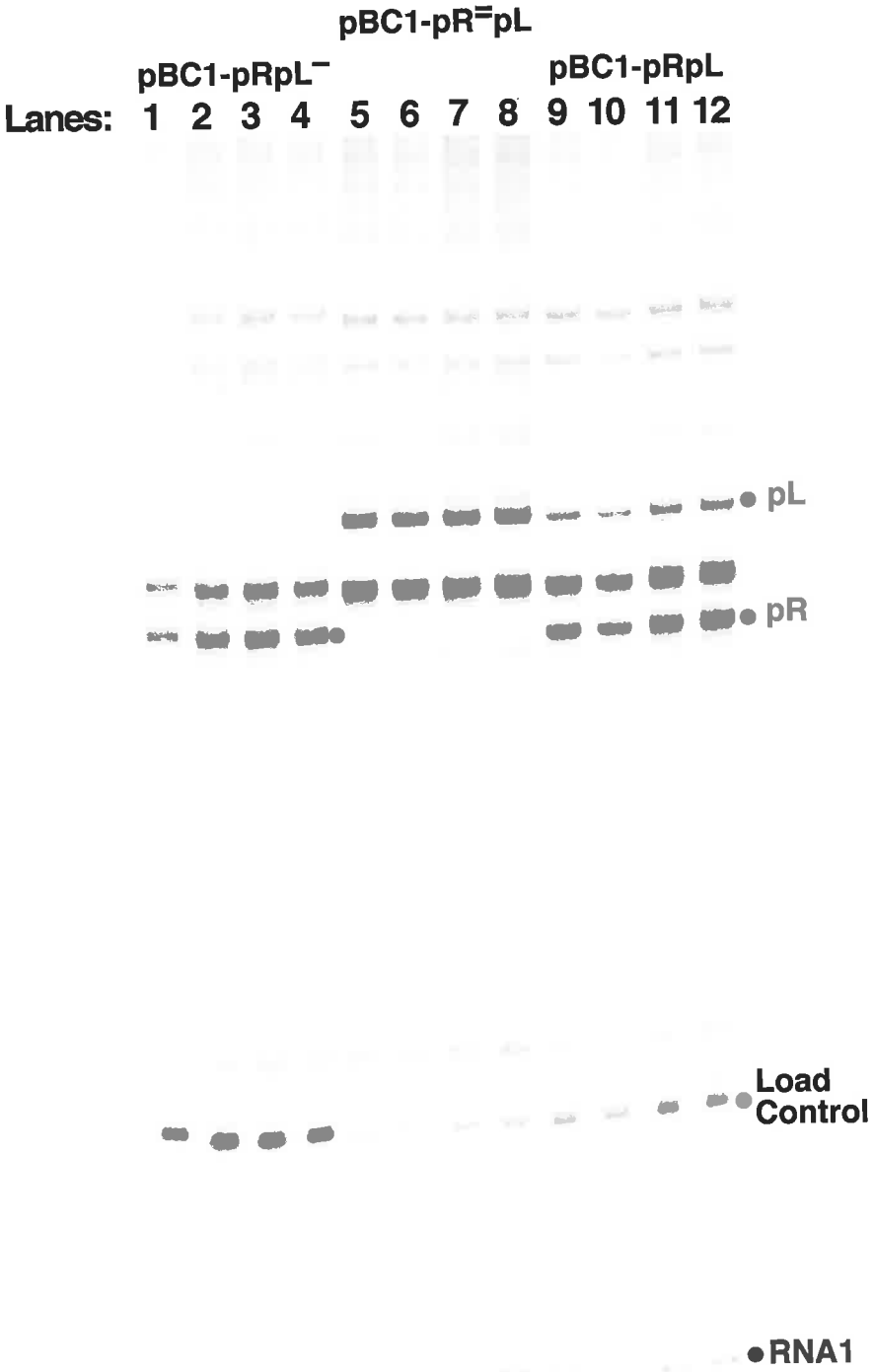
4.2.5 Attempts to establish an *in vitro* model of interference.

If interference of *pL* activity was due only to a sitting duck mechanism, then the reduction in *pL* activity caused by one round of collisions between *pL* open complexes and elongation from *pR* should be amplified in a multiple round transcription assay, such that *pL* activity is reduced 5.6 fold. Steady state transcription of *pR* and *pL* should be obtained using a multiple round assays by simply omitting heparin to allow RNAP to re-initiate cleared promoters. Transcription reactions were set up using equivalent

Figure 4.9

Transcript pattern showing the effect of convergent transcription on the rate of full-length transcript accumulation from *pR* and *pL* open complexes during single-round transcription assays.

Transcript pattern showing the rate of full-length transcript accumulation from open complexes formed at pBC1-*pRpL*. Reactions were performed as described for the escape assay in fig 4.4, except using pBC1-*pRpL* as the template. The transcript patterns are examples of the data used to obtain the results described in fig 4.10. Lanes 1 to 7 correspond to a transcription reaction terminated 0.5, 1, 1.5, 2, 4, 8 and 16 minutes after the addition of NTPs respectively and was used to quantitate the rate of full-length transcript accumulation from $pR(pL^{+})$ open complexes; lanes 8 to 14 are from another transcription reaction terminated 2, 4, 7, 11, 18, 30 and 50 minutes after the addition of NTPs and was used to quantitate the rate of full-length transcript accumulation from $pL(pR^{+})$ open complexes.


pBC1-pR=pL
pBC1-pRpL⁻
pBC1-pRpL
Lanes: 1 2 3 4 5 6 7 8 9 10 11 12
pL
pR
Load Control
RNA1

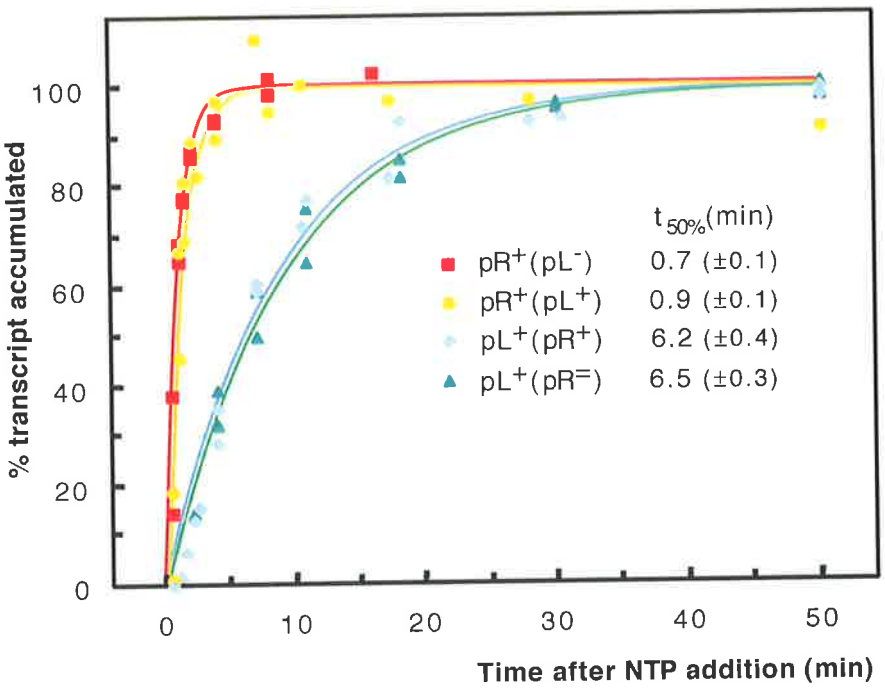


Figure 4.10 Graph showing the effect of convergent transcription on the rate of full length transcript accumulation from *pR* and *pL* open complexes during single-round transcription assays.

Bands for full length transcripts $pR(pL^{-})$ (red squares), and $pL(pR^{=})$ (green triangles) without convergent transcription and *pL pR* transcripts during convergent transcription $pR(pL^{+})$ (yellow circles) and $pL(pR^{+})$ (blue diamonds) from the transcript patterns of clearance assays described in fig 4.9 and fig 4.4 were quantitated and graphed as described in the legend of fig 4.5. Values for $t_{50\%}$ are shown.

conditions to those used previously but reactions were initiated by adding DNA template to a mixture of RNAP and rNTPs with no heparin present. To analyse whether stead state transcription was achieved, the kinetics of transcript accumulation from a multi-round assay containing pBC1-*pRpL* was analysed. The transcript pattern of the reaction is shown in fig 4.11 and the bands for full-length *pL*, *pR* and RNA1 transcripts were quantitated and transcription units calculated were plotted against reaction time (fig 4.12A). The kinetics for accumulation of *pR* and RNA1 transcripts were found to be similar to that of the single-round clearance assay. For steady state, multiple-round transcription, transcript accumulation for *pR* and RNA1 should be increasing linearly with time. Instead, a rapid rise to a maximum was seen, indicating no steady state of transcript production. The lack of multiple rounds of transcription was confirmed by measuring the kinetics and quantities of transcription from an identically prepared experiment but which contained heparin (fig 4.11 shows the transcript pattern and fig 4.12B shows transcript accumulation with time, expressed as transcription units). The kinetics of transcript production for both single and multiple round assays were nearly identical. It seemed as though multiple round transcription assays were actually capable of only one round. This was presumably due to some condition in the reaction that became limiting during transcription. Attempts to determine which conditions (rNTPs, salt, RNAP or DNA) were limiting were unsuccessful (not shown). This problem has been reported by other laboratories (personal communications, and (Sukhodolets *et al.*, 2001)) where it is suggested that RNAP molecules become nonfunctional after completing only one or two rounds of transcription under conditions similar to those used here. The inactivation has been shown to be the result of transcription, as RNAP incubated with DNA and transcription buffer without NTPs was still active after a prolonged incubation (Sukhodolets *et al.*, 2001). The problem has been shown to be circumvented by the addition of a host protein RapA, that associates with high affinity to core RNAP, and stimulates RNAP recycling during *in vitro* transcription from supercoiled templates (Sukhodolets *et al.*, 2001). It would be interesting to know whether the additional of RapA to *pRpL in vitro* transcription reactions could achieve multiple rounds of transcription and interference, however time limits prevented these experiments being performed.

Figure 4.11

Transcript pattern showing the elongation kinetics of multi-round versus single-round transcription assays using the template pBC1-*pRpL*.

Transcript patterns are from standard transcription reactions. Two identical standard transcription reactions were set up containing RNAP and DNA template which were pre-incubated at 37°C for 30 minutes to allow open complex formation to occur. At time point zero, NTPs were added to the multi-round reaction (lanes 1-7) and a NTPs/heparin mix was added to the single round reaction (lanes 8-14) to initiate elongation. To monitor the accumulation of transcripts from each reaction, aliquots of the reactions were taken and added to stop/load buffer containing load control. Lanes 1-7 and 8-14 correspond to a transcription reaction terminated 20 secs, 40 secs, 1, 1.5, 2, 3, and 10 mins after initiation. Transcripts were quantitated and results shown in fig 4.12.

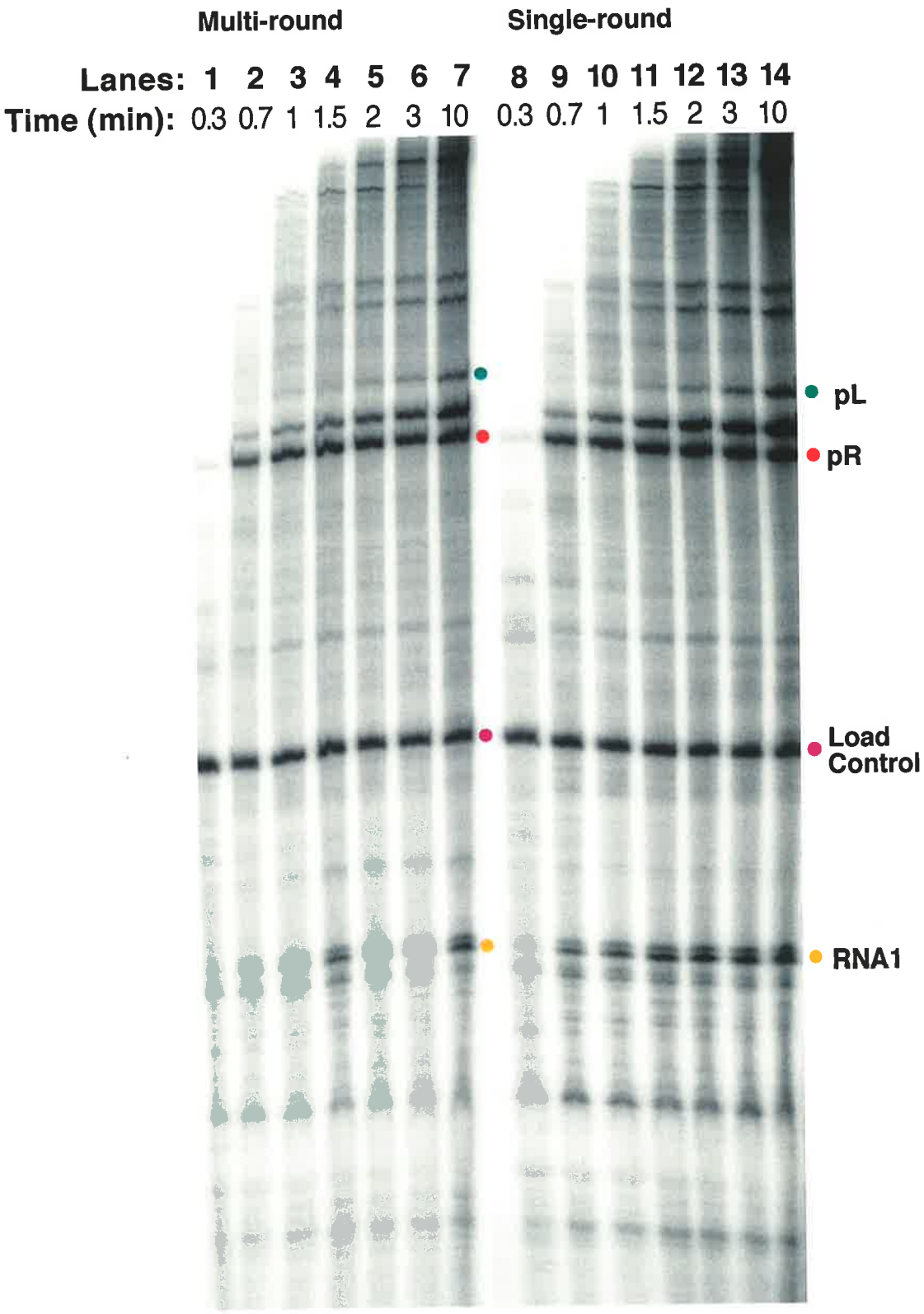
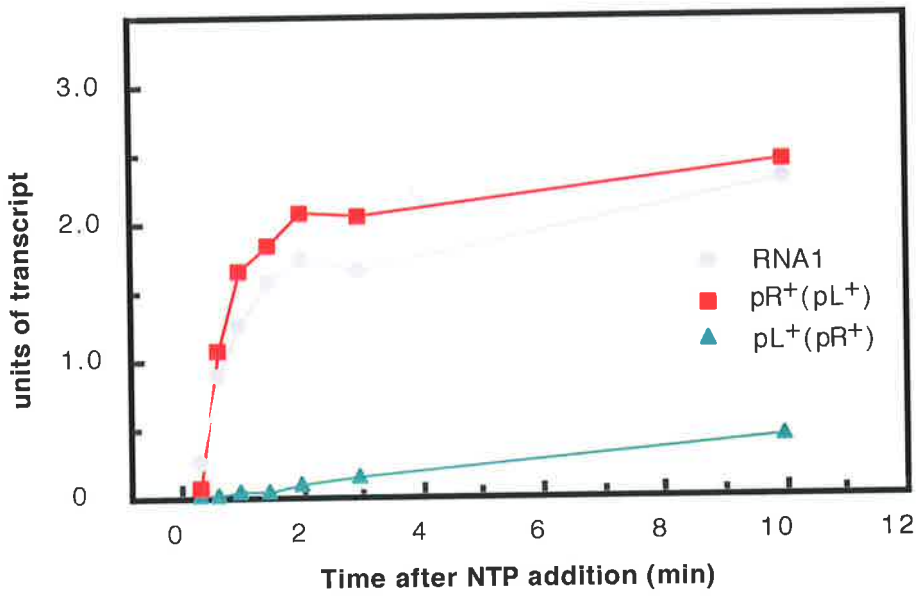

Multi-round
Single-round
Lanes: 1 2 3 4 5 6 7 8 9 10 11 12 13 14
Time (min): 0.3 0.7 1 1.5 2 3 10 0.3 0.7 1 1.5 2 3 10
pL
pR
Load Control
RNA1

Figure 4.12

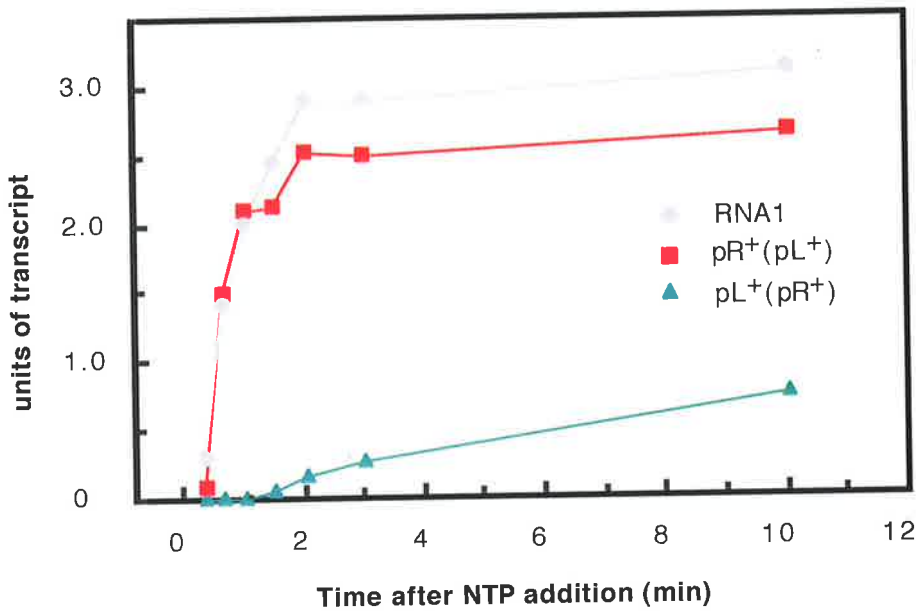
The elongation kinetics of multi-round (A) versus single-round (B) transcription assays using the template pBC1-*pRpL*.

Bands for full-length transcripts from *pR*(pL^+), *pL*(pR^+) and RNA1 from transcript patterns of fig 4.11 were quantitated and plotted against elongation time. The amount of transcript was normalised to the loading control in each lane and to the amount of U's in each transcript and are expressed as 'units of transcript'.

A. Multi-round transcription.



B. Single-round transcription.



4.3 Discussion

4.3.1. Kinetics of promoter initiation at *pL*

In this chapter the kinetics of one round of transcription from *pL* were measured under *in vitro* conditions that attempt to simulate an *in vivo* environment. The intention was to discover the nature of interference that occurs due to elongation over *pL* by examining the potential of *pL* to form 'sitting duck' RNAP initiation intermediates that are sensitive to collisions with elongation from *pR*. The kinetics of transcription were divided into two stages: (i) the rate of formation of heparin resistant complexes and (ii) the rate of formation of full-length transcripts from these heparin resistant complexes. The first stage was found to occur rapidly, with 80% of the complexes forming within 20 seconds. The second stage was much slower, with 50% of the complexes producing full-length transcripts within 6.5 minutes.

Formation of heparin resistant complexes.

The formation of heparin resistant complexes was measured by examining the amount of *pL* transcript produced from transcription reactions where RNAP and DNA template had been pre-incubated at various times. The speed at which the events of promoter recognition, closed complex formation, and isomerisation to form open complexes at *pL* occurred were rapid under these conditions. Transcript production for *pL* seemed to initially increase with pre-incubation time then after ~1 minute decreased (with an exponential rate of decay) probably to some equilibium of ~50%, but not to zero (substantial transcription from *pL* was still observed after a 60 minute pre-incubation of DNA and RNAP, see experiments in 4.2.1 and 4.2.2). This decrease in transcript production indicates either an inactivation or a dissociation of open complexes. Open complex formation was once thought to be essentially an irreversible reaction, however this view of transcription has since been reconsidered and replaced by a reversible kinetic mechanisms of initiation (Record *et al.*, 1996). If the rate of dissociation of open complexes formed at *pL* was much slower than the rate of formation, than it is possible that the decrease in active open complexes with time reflects the rate at which an equilibrium between formation and dissociation is reached. Note that the rate of dissociation may be of little significance *in vivo* during continuous steady state transcription in the presence of NTPs. Alternatively, the decrease in active open

complexes maybe a reflection of the rate of formation of inactive complexes at *pL*. In what has been termed the branched pathway, a part of the enzyme-promoter complex is arrested at the promoter forming a moribund complex which has been shown at the $\lambda P_R AL$ promoter to be slowly converted into an inactive dead-end complex (Sen *et al.*, 2000). A generality of the branched pathway mechanism among the promoters for *E. coli* σ^{70}-holoenzyme has been suggested (Susa *et al.*, 2002). If moribund complexes form at *pL* it is possible that in the absence of NTPs there is a rate at which an equilibrium between active and moribund complexes is established, and the decrease in active open complexes with pre-incubation time observed for *pL* reflects this rate. Note that the branched pathway has only been found during *in vitro* transcription and the occurrence of a similar pathway *in vivo* is yet to be determined. The reason for the decrease in open complexes with pre-incubation time has no bearing on the conclusions made about the potential for *pL* to be involved in a sitting duck mechanism, namely that the rate at which *pL* forms open complexes is rapid but the rate at which these produce transcripts is much slower.

Experiments here show that the amount of full-length transcript produced from *pL*, RNA1 and *pR* during *in vitro* transcription, differs. After a 30 minute pre-incubation of DNA and RNAP, *pR* produced the greatest amount of transcripts, and *pL* the least. This would not be expected if polymerase was at saturating concentrations and every open complex formed produced a full-length transcript. This result may reflect either non-saturating conditions at some promoters or the branched pathway of initiation at some promoters leading to the formation of inactive, moribund complexes which do not produce productive, full-length transcripts.

Rate of promoter clearance.

The rate of promoter clearance was shown to fit a first order exponential. This indicates that the process of clearance is a stochastic biochemical process, such that within a population of open complexes at the same promoters, some complexes will immediately begin producing transcripts and others will take longer, most likely spending more time involved in abortive initiation (the process of promoter escape). The rate of promoter escape reflects the 'abortive probability' of an initiation complex to abort elongation of nascent transcript at each position downstream of the initiation start site (Hsu, 1996).

The abortive probability at each position of a promoter can be measured by analysing the rate of production of short abortive transcripts versus the rate of productive initiation from a promoter. I suspect that the reason for the slow rate of productive transcript synthesis from *pL* is a reflection of a slow rate of promoter escape, which would lead to an overproduction of abortive transcripts from *pL*. It would therefore be of interest to analyse the abortive transcript pattern and its rate of accumulation for *pL* and observe whether the abortive probability of nascent *pL* transcripts could explain its slow rate of clearance. To perform this type of analysis on supercoiled templates would require the use of a circular template that contains no other promoters which could also contribute to the pattern of abortive transcripts. For this purpose the minicircle system for measuring *in vitro* transcription from a naturally supercoiled template developed by Choy and Adhya (1993) could be used for *pL*, as it was for measuring the rate of synthesis of abortive and full-length transcripts from the *gal* P2 promoter (Cashel *et al.*, 2002).

The distribution of clearance times measured was expected to be continuous about a mean. The mean is expected to be dependent on the reaction conditions and the sequence of the promoter. Sequence determinants for promoter clearance are not well defined but are expected to be equivalent to those of promoter escape, known to involve the core promoter recognition region, the initial transcribed sequence and the conformational state of the template DNA (Hsu, 2002). The use of known sequence determinants of promoter escape are discussed with reference to *pL* and another promoter *pc* (analysed in chapter 5) in the final discussion (Chapter 6). Reaction conditions known to affect promoter escape are the addition of transcript cleavage stimulatory factors GreA and GreB which can act on short RNA-bearing initiation complexes to stimulate promoter escape *in vitro* and *in vivo* (Hsu *et al.*, 1995). These factors are not expected to be present in my reactions, however it would be interesting to know if the addition of GreA and GreB reduced the clearance time of *pL*. Although kinetic experiments performed here do not determine the specific rate limiting step in promoter initiation, they do show that for *pL* the events leading up to clearance are relatively rapid but once open complexes form, due to the stochastic nature of promoter clearance, some will be able to clear rapidly enough to avoid collisions with *pR* elongation but most will be involved in collisions with a *pR* elongation complex. Due to

the fact that *pR* is about 10 fold more active than *pL in vivo*, I expect that during multiple round transcription, one open complex at *pL* on average could be hit by 10 or more rounds of elongation from *pR*.

4.3.2. Head-on collisions between an open complex and elongating RNAP.

The outcome of head-on collisions between elongating RNAP molecules and open complexes was previously investigated by Horowitz and Platt (1982). Their study used convergent *lac* and *trp* promoters *in vitro* to suggest that head-on collisions between open complexes and elongating RNAP cause the elongating polymerase to pause at the site of collision until the release of open complexes from the promoter (described further in the introduction 1.4.4). Additionally, in single round assays using the trfAp and trbAp promoters, an open complex at either promoter appeared to have the same probability of producing a transcript regardless of any open complexes formed downstream. The authors indicated that an opposing RNAP is not an insurmountable barrier to transcription *in vitro* (Jagura-Burdzy and Thomas, 1997). The ability of RNAP to transcribe through an open complex without terminally dissociating from the template DNA was also suggested by the experiments performed here. Contrary to these previous studies I showed that although the progress of the elongating polymerase is not affected, at least 30% of the open complexes involved in a collision become completely incompetent to produce full-length transcripts. Moreover, an examination of the rates of transcript accumulation for *pL* and *pR* from pBC1-*pRpL* indicated no pause in the production of either *pR* or *pL* full-length transcripts, suggesting that during a collision between elongation from *pR* and open complexes at *pL* the progress of the elongating polymerase is completely unaffected. Possible reasons for the discrepancy between these results and those of Horowitz and Platt (1982) is either the use of supercoiled DNA compared to linear restriction fragments as the template for collisions, and also the use of different promoters. The stability of open complexes formed at *pL* and at the *lac* promoter may not be equivalent and could result in different barriers to elongation and thus different outcomes during collision events.

The fact that only 30% of active open complexes formed at *pL* failed to produce transcripts after transcription from *pR* raises the possibility that not every collision between open complexes has the same outcome. If it is correct that during *in vivo*

transcription of *pL* the rate limiting step for initiation is clearance from the promoter, then convergent transcription from a promoter that is 10 fold more active (such as *pR*) would be expected to collide with an open complexes formed at *pL* at least more than once. If every collision caused the inactivation of 100% of the open complexes, then it is difficult to imagine why *pL* activity is not reduced to a much lower level of activity during interference. A heterogenous outcome of collisions could be rationalised by a number of possibilities. i) collisions with complexes at *pL* that have progressed through different stages of promoter clearance have different stabilities resulting in different outcomes. ii) There exists different populations of elongation complexes from *pR* that create different collision outcomes. Evidence for different populations of RNAPs based on recent studies of transcription from single RNAP molecules is conflicting. One group suggests that that an elongating RNAP population is composed of RNAPs in distinct states that elongate at different intrinsic rates and are more or less likely to respond to pausing signals (Davenport *et al.*, 2000). However, another group has shown that individual RNAPs exhibit homogeneous elongation dynamics, with differences among RNAPs arising from random switching between a single active elongation mode and the paused state (Adelman *et al.*, 2002). iii) Inactivation of open complexes during head-on collisions is not an all or none event but a stochastic process with a defined probability, much like the determination of the efficiency of termination by rho-independent terminators (Mooney *et al.*, 1998).

An alternative explanation to collisions causing the reduction in *pL* transcripts is that the presence of active convergent *pR* causes a decrease in the number of heparin resistant complexes formed at *pL* prior to the addition of NTPs. This possibility is not ruled out by experimental evidence presented here. However experiments were performed at a 25 molar excess polymerase to template concentration such that no promoter competition is expected. Steric hindrance from convergent open complexes formed on the same template is not expected to occur for promoters spaced more than 37 bp apart ((Jagura-Burdzy and Thomas, 1997) as discussed in section 1.4.4). Moreover, experiments performed *in vivo* in chapter 3 ruled out the possibility of interference occurring at the level of promoter binding. This theoretical possibility could be tested by further experiments analysis of the occupancy of RNAP at *pL* in the presence and absence of active *pR* (for example by footprinting analysis). To be

consistent with collision events accounting for the decrease in *pL* transcripts, *pR* should not affect the occupancy of *pL* by RNAP.

4.4 Summary

Chapter 3 demonstrated that the most likely explanation of interference is due to the RNAP from *pR* convergently transcribing over the *pL* region. In this chapter two main mechanisms of interference caused by elongation over *pL* were proposed. The first was promoter occlusion whereby the progress of an elongating RNAP over *pL* blocks σ^{70}RNAP recognition at *pL*. Based on theoretical grounds the extent of exclusion was predicted to be minimal for transcription from *pR*.

The second model was termed the 'sitting duck' mechanism where an elongating RNAP from *pR* collides most frequently with *pL* initiation intermediates formed prior to the most rate limiting step of initiation causing 'damage' to that initiation intermediate. For any given weak promoter these 'siting ducks' may be a closed complex, an open complex (and other intermediates) or an abortive complex waiting to clear. In this chapter the potential of *pL* to form sitting ducks was determined by analysing the kinetics of a single round of *in vitro* transcription. *pL* was found to quickly form open complexes but the clearance of these open complexes from the promoter was shown to be very slow, indicating that *pL* has a high potential to form sitting ducks in the form of open complexes waiting to clear. The sensitivity of these complexes to transcription from *pR* was then demonstrated by an *in vitro* transcription system, which analysed the amount of full-length transcripts produced from *pL* open complexes after collisions with convergent elongating RNAP. After these collisions the amount of transcript produced from *pL* was reduced by one third, demonstrating that a sitting duck mechanism of interference at *pL* is possible and likely. A prediction from this model of interference is that weak promoters with different properties (and thus a different potential to form sitting ducks) would give different responses to strong convergent transcription. This is the subject of the next chapter.

CHAPTER FIVE

Convergent promoter interference in bacteriophage P2 supports the sitting duck model

Chapter 5

Convergent promoter interference in bacteriophage P2 supports the sitting duck model

5.1 INTRODUCTION

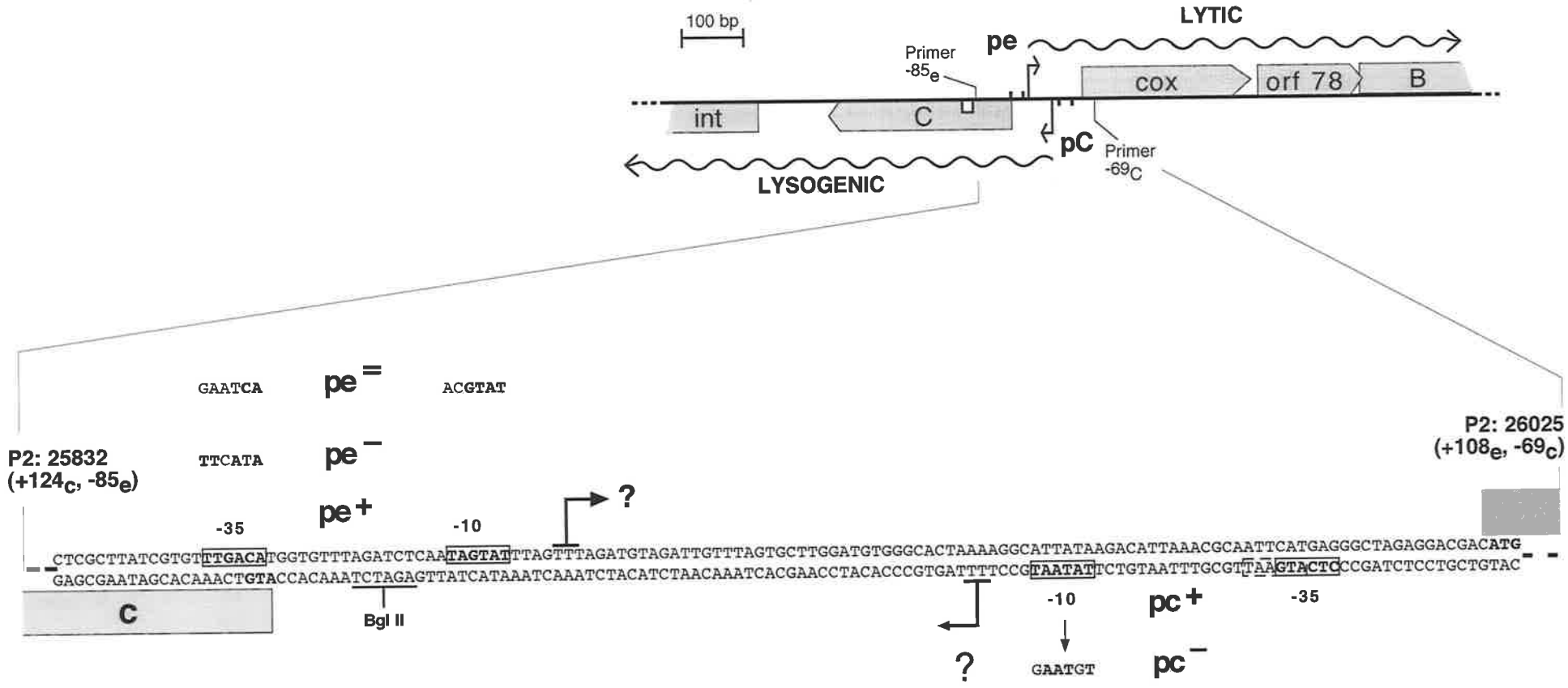
The conclusion from chapter 3 was that transcription from *pR* has to traverse *pL* to display interference. In Chapter 4, the nature of this interference was found to be most likely due to a sitting duck mechanism. This mechanism of interference would predict that weak promoters with different rate limiting steps should give different responses to strong convergent transcription. For example, a weak promoter that binds polymerase poorly but upon binding, clears rapidly, would be less sensitive than a weak promoter that binds polymerase well but is slow to clear. The switch promoters of the temperate coliphage P2 are arranged face-to-face, as with its close relative 186, but with start sites only 39 bp apart rather than 62 bp (see fig 5.1). Saha *et al* (1987) reported a 30 fold interference of the lysogenic promoter by the presence of the lytic promoter (results are summarised in fig 5.2). It was therefore expected that either the lytic promoter of P2 was substantially stronger than that of 186 (and thus more aggressive), or its lysogenic promoter has a greater 'sitting duck potential' than that of 186 (and is thus more sensitive). In an attempt to confirm the mechanism of interference suggested for 186, this chapter is concerned with testing these predictions. Section 5.2 details the measurement of interference using my lac Z reporter system, and section 5.3 and 5.4 explores the difference in interference between 186 and P2 promoters, testing the validity of predictions from the sitting duck mechanism.

Note that in this chapter labels for specific P2 clones use the following nomenclature: *pepc* refers to the 195 bp fragment used in fig 5.2, *pepc* or *pepc* refers to clones made in pBC2 to assay either *pe* in the presence of active *pc* or *pc* in the presence of active *pe*, respectively. The promoter being assayed is underlined, and the presence of mutated sequences are indicated by pe^{-} or pc^{-} or $pe^{=}$. +/–$\#_c$ and +/–$\#_e$ refer to the number of P2

Figure 5.1

Convergent promoters of the P2 developmental switch.

Diagram of the convergent promoters *pe* and *pc* from the developmental switch region of P2 (shown above) and the DNA sequence (shown below) of the fragment used in this study. The switch region is drawn to scale. Filled boxes indicate positions of –10 or –35 hexamers for promoters *pe* or *pc*, the white box indicates the IHF binding sequence. On the sequence, bent arrows indicate the predicted start sites of transcription, predicted –10 and –35 hexamers are boxed and the promoter mutations used are indicated (the broken box at *pc* indicates the –35 hexamer predicted by Saha *et al.*, 1987). $\#_e$ and $\#_C$ respectively indicate the distance from the start of *pe* and *pc* transcription (assumed to be the middle T). Also indicated is the % occurrence of consensus sequences in the list of known *E. coli* σ^{70} promoters which can be used as a rough guide for predictions of promoter strengths and the effects of promoter mutations.



consensus	T	T	G	A	C	A							T	A	T	A	A	T
(non template strand)							17 bp spacer											
% occurence (Record et al. 1997)	69	79	61	56	54	54	43						77	76	60	61	56	82

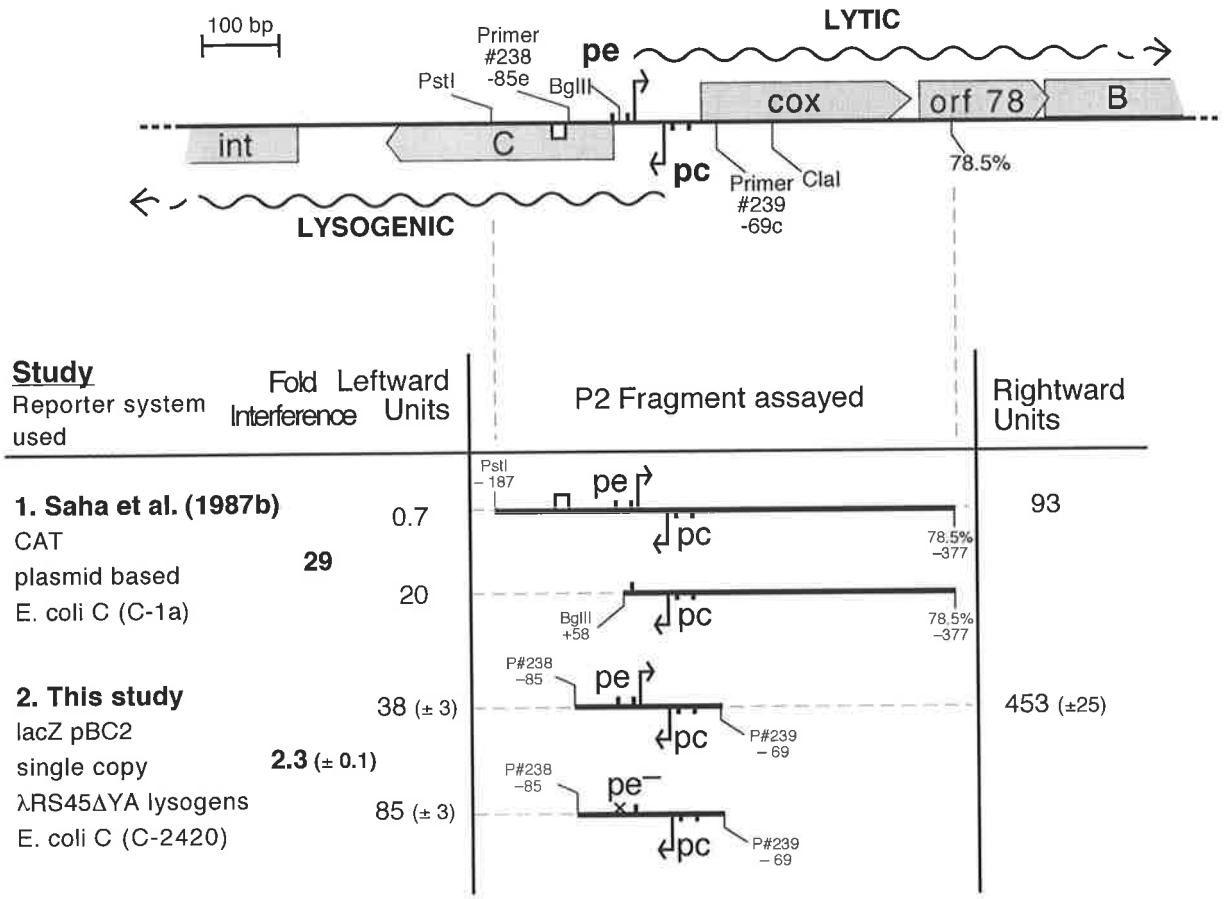


Figure 5.2 Comparison of P2 physiological promoter strengths and interference from the literature and from the present study.

The constructs used to measure promoter activity in each study are shown (drawn to scale). Filled boxes indicate positions of –10 or –35 hexamers for promoters *pe* or *pc*, crosses indicate a mutated hexamer, white boxes indicate the IHF binding sequence. The positions and numbers of the primers (P#238 and P#239) used to generate P2 promoter inserts used in this study are indicated. In the table of P2 fragments assayed, numbers above the line indicate the position from the predicted +1 site of pe and numbers below the line indicate the position from the predicted +1 of *pc*. Fold interference for study 1 is the ratio of leftward units in the absence and presence of *pe* sequences. Fold interference for this study was calculated as previously discussed for *pRpL* (Chapter 2) but using ratios of pe^- over pe^+ leftward *lacZ* units. Errors show 95% confidence limits.

base pairs upstream (–) and downstream (+) from the start site of transcription (from fig 5.1) of *pc* or *pe* respectively.

5.2 Interference between the switch promoters of P2

5.2.1. The *in vivo* level of interference between P2 promoters is 2 fold when measured using a single copy *lacZ* promoter assay.

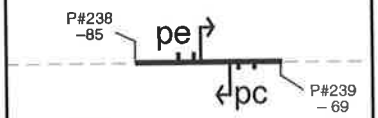
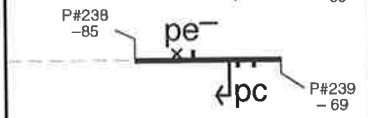
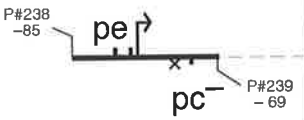
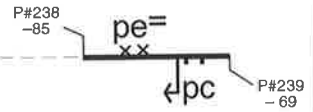
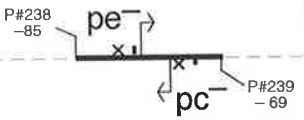
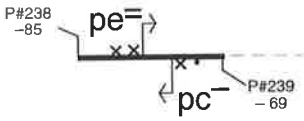
Saha *et al* (1987) used a plasmid based CAT assay system to measure promoter strengths. To confirm the high degree of transcriptional interference, the activities of the P2 switch promoters were determined using the experimental system developed in Chapter 2. A 195 bp P2 fragment from –85 of *pe* to –69 of *pc* was used for the promoter assays. Saha *et al* [1987b] measured the intrinsic activity of the lysogenic promoter, *pc*, by deleting the convergent promoter region. In this study the intrinsic promoter strength of the lysogenic promoter was measured in the presence of mutations introduced into the P2 lytic promoter, *pe*, to avoid possible context effects caused by its deletion. *pe* has a consensus –35 region (TTGACA) and mutations of this region were chosen based on the frequency of bases occurring in each position of the –35 region (Lisser and Margalit, 1993) (summarised by Record *et al.*, (1996) and shown in fig 5.1), and also studies showing the effect of base substitutions in the promoter –35 regions on promoter strength (Youderian *et al.*, 1982; Szoke *et al.*, 1987; and Kobayashi *et al.*, 1990). The consensus –35 region of *pe* was mutated from TTGACA to TTCATA as shown in fig 5.1, giving pe^{-}. The effect of the mutations on promoter strength was determined from the 195 bp P2 fragment in the presence of a mutated convergent *pc* promoter (see section 5.2.3) to avoid the potential of interference masking the correct activities (fig 5.3(e)). Using MC1061.5 as a host strain $\underline{pe^{-}}pc^{-}$ yielded 7 ± 0.3 units, a reduction in *pe* activity of over 150 fold, demonstrating that the mutations made in *pe* essentially inactivate the promoter.

P2 promoter fragments were cloned into the *XbaI* site of the *lacZ* reporter vector pBC2 to generate pBC2-*pepc*, pBC2-*pepc* and pBC2-$pe^{-}\underline{pc}$ and then promoter-*lacZ* fusions were transferred to λRS45ΔYA. The promoter assays reported in Saha *et al* (1987) were performed in an *E. coli* strain C, as opposed to the K-12 strains used so far in this study.

Figure 5.3

Measurements of P2 promoter activity and interference using different constructs.

Activities of all constructs are lacZ units measured from the constructs indicated, which had been inserted into the *XbaI* site of pBC2. The fusions were located on single copy λRS45ΔYA-pBC2-*pepC.lacZ* prophages using MC1061.5 as the host strain. Leftward indicates transcription measured by placing the constructs indicated into pBC2 *lacZ* reporter constructs such that *pc* directs transcription of the *lacZ* gene. Rightward indicates equivalent constructs placed into pBC2 *lacZ* reporter constructs such that *pe* directs transcription of the *lacZ* gene. Fold interference was calculated from ratio of pe^{-} (or $pe^{=}$) over pe^{+} leftward *lacZ* units, as discussed for 186 promoters in Chapter 2. Entry (a) was used as the pe^{+} comparison for entries (b) and (d). All lacZ assays were performed at least 12 times and errors show 95% confidence limits. Symbols for the constructs used are as described in the legend for fig 5.2.

Entry	Name	Fold Interference	Leftward Units	Construct	Rightward Units
(a)	$pe^{+}pc^{+}$		64 (± 6)		982 (±110)
(b)	$pe^{-}pc$	**2.2 (± 0.2)**	150 (± 9)		ND
(c)	$pepc^{-}$		ND		1084 (±66)
(d)	$pe^{=}pc$	**2.2 (± 0.3)**	141 (± 8)		ND
(e)	$pe^{-}pc^{-}$		8 (± 0.2)		7 (±0.2)
(f)	$pe^{=}pc^{-}$		ND		6 (±1)

So that a comparison could be made between studies, *E. coli* C was used as the host *lacZ* reporter strain. Measuring the intrinsic activity of *pc* and *pepc* activity, the lysogenic promoter was about 5 fold weaker than the lytic promoter, 85 ± 3 units versus 453 ± 25 units, respectively (fig 5.2). This was equivalent to the comparative strengths measured by Saha *et al* 1987, 20 versus 93 units. In marked contrast to the findings of Saha *et al* (1987), the presence of active convergent lytic transcription did not reduce *pc* activity 30 fold but rather 2.3 ± 0.1 fold, from 85 ± 3 to 38 ± 3 units (fig 5.2). This indicated that *pc* may not be more sensitive to interference than *pL*.

5.2.2. Comparison between P2 and 186 promoter switch promoter activities.

To enable a better comparison of the relative promoter strengths and the level of interference between the P2 and186 switch promoters, the assays performed in section 5.2.1 were repeated using the same host strain used to measure 186 interference, ie. MC1061.5. In comparison to the results using the C strain, all the promoter activities measured were nearly doubled (fig 5.3). Intrinsic *pc* activity increased from 85 ± 3 to 150 ± 9 units, *pepc* activity increased from 453 ± 25 to 982 ± 110 units, and *pepc* activity increased from 38 ± 3 to 64 ± 6 units. Interference of *pc* activity was calculated to remain unchanged at 2.2 ± 0.2 fold. When comparing these activities to those of 186 *pRpL*, the major difference is the extent of interference, 5.6 fold for 186 *pL* and 2.2 fold for P2 *pc* (compare with fig 2.3(c)). There is also a difference between the comparative strengths of the intrinsic lysogenic activities and the convergent lytic activity, which for P2 is ~6.5 fold, and for 186 this is ~9.4 fold. This is mostly due to a stronger lytic activity for 186, 1340 ± 113 for *pR* versus 982 ± 110 units for *pe*. The intrinsic lysogenic promoter activities are within 95% confidence limits of each other, 142 ± 3 for *pR*=*pL* compared to 150 ± 9 for *pe*⁻*pc*. All future promoter assays were performed in MC1061.5.

5.2.3. Interference between the P2 promoters is not reciprocal.

To measure whether the convergent activity of *pc* affects the activity of *pe*, intrinsic *pe* activity was measured from the same 195 bp P2 fragment used previously but containing mutations that inactivate *pc*. Mutations in the –10 region of *pc* (which is a consensus sequence) were chosen using a similar rationale to the mutations made in *pe* and created using site directed mutagenesis, changing TATAAT to TGTAAG (fig 5.1).

The mutation was shown to reduce *pc* activity nearly 20 fold, from 141 ± 8 to 6 ± 1 units, (fig 5.3(e)). The activity of *pe* in the presence of pc^- was measured at 1084 ± 66 units, which is within the 95% confidence limits of *pepc* (fig 5.3(c)). Thus active convergent *pc* activity did not affect lytic promoter activity, a result equivalent to that found with the effect of *pL* activity on *pR* transcription (Chapter 2).

5.2.4. Interference is not affected by different *pe* mutations.

A second inspection of the mutations made in the –35 region of *pe*, revealed that the mutations also changed the AUG start codon of the *C* gene transcribed from *pc*. This mutation is expected to disrupt translation of the initial *C* gene sequence normally occurring on the 195 bp P2 fragment used for the promoter assays. A disruption of translation may create a context effect during the determination of intrinsic *pc* activity, which is not present when measuring *pepc* activity. A reason for this context effect may be premature termination of the pe^-pc transcript caused by rho-dependent termination of the *lacZ* transcript, which contains a long untranslated region (Stanssens *et al.*, 1986). To address this concern, alternative mutations of *pe* –10 (from TAGTAT to **AC**GTAT) and –35 (from TTGACA to **GAAT**CA) regions were made as shown in fig 5.1, and the mutated *pe* was labelled $pe^=$. The mutations were shown to reduce *pe* activity 180 fold, from 1084 ± 66 units to 6 ± 1 units (fig 5.3(f)). Using $pe^=$, the intrinsic *pc* activity was re-determined to be 141 ± 8 units (fig 5.3(d)) which is within error of the activity of pe^-pc, (150 ± 9 units). Interference was calculated to remain at 2.2 ± 0.3 fold. These results indicate that for the 195 bp P2 fragment used, translation of part of the *C* gene from *pc-lacZ* transcripts had no effect on the reporting of intrinsic *pc lacZ* activity.

5.3 Can the sitting duck model of interference explain the difference in interference between 186 and P2 promoters?

The purpose of this study was to find evidence that confirms the proposed sitting duck mechanisms of interference between *pR* and *pL* by investigating another example of convergent promoters that demonstrate a different level of interference. The *pepc* promoter system was shown here to produce a 2.2 fold level of interference which is much lower that the 5.6 fold effect of *pR* on *pL*. This is despite the activities of the sensitive (lysogenic) promoters being essentially the same (of approximately 150 lacZ

units). Although the difference in interference was the reverse of what was anticipated, the original purpose of the study remains, namely, to test whether the difference in interference is a reflection of differences in promoter parameters relating to the sitting duck potential of that promoter. Without assuming the mechanism of interference, three possible explanations exist for the difference in interference between 186 *pRpL* and P2 *pepc*. Either (i) the lytic promoter of 186 is more 'aggressive' than that of P2, (ii) the lysogenic promoter of 186 is more sensitive than that of P2, (which may or may not be reflected in the sitting duck potential of each promoter), or (iii) the closer spacing of P2 convergent promoters (with 39 bp between +1 sites) compared to the 186 convergent promoters (with 62 bp between +1 sites) is able to reduce the extent of interference. The influence of spacing was investigated first.

5.3.1. Increasing the spacing between convergent P2 promoters has no effect on interference.

The influence of promoter spacing was investigated by determining whether the lower level of interference was retained when the spacing between P2 promoters was increased to that of the 186 promoters. The available literature on the P2 switch promoters has not experimentally determined the start sites of *pc* and *pe* transcription, the distances quoted between +1 sites are thus hypothetical. Multiple start sites are possible for some promoters, therefore when deciding the spacing between P2 promoters it was more relevant to ensure that the –10 regions between 186 and P2 convergent promoters were equivalent. –10 regions of a promoter are less variant than +1 and are also well determined for each promoter being studied. The –10 regions of 186 promoters are separated by 76 bp whereas those of P2 are separated by 49 bp. To obtain a separation of 76 bp the *pe* fragment from -85_e to 35bp downstream of the *pe* –10 region, was cloned convergent to a *pc* fragment from -69_c to 35 bp downstream of the *pc* –10 region, with an additional *NheI* site in the center of the promoters (see fig 5.4A). The resulting clones, *pe*(*NheI*)*pc* and *pe*$^{-}$(*NheI*)*pc*, maximise the wildtype downstream promoter sequence possible for such clones. Clones were transferred to single copy using MC1061.5 as the host strain and *lacZ* activity assayed (fig 5.4B(a)). Although measurements of the intrinsic activity of *pc* were reduced from 150 ± 9 to 109 ± 5, interference was still retained at 2.1 ± 0.2 fold. The activity of *pe*(*NheI*)*pc* was essentially unchanged at 922 ± 46 units.

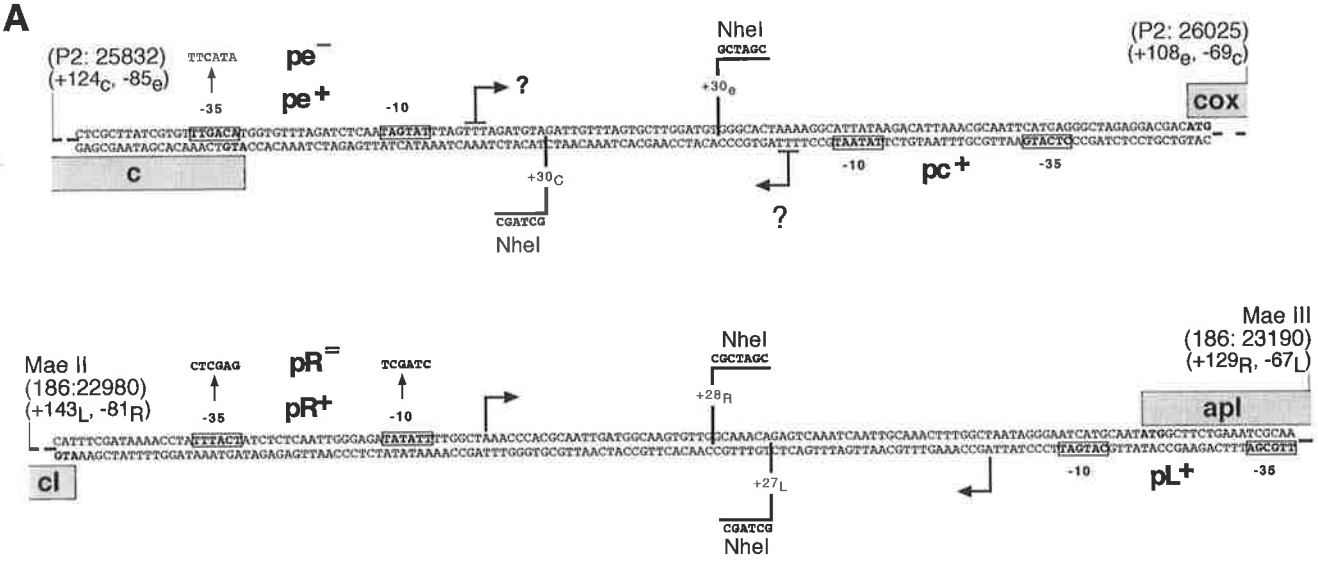
Figure 5.4

Measurements of promoter activity and interference for combinations of convergent 186 and P2 promoters placed at the same distance as wildtype *pR* and *pL*.

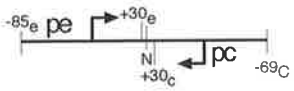
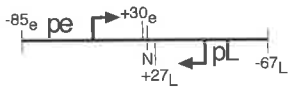
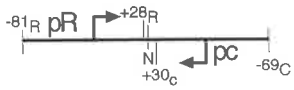
A. Sequence of the promoter fragments (or modules) used to clone different combinations. The *NheI* site was located on primers used to generate each promoter fragment. The *NheI* site was used as the linker between convergent promoter fragments and the position of the *NheI* site for each promoter fragment was chosen such that final promoter combinations have the same distance between –10 hexamers as that of wildtype *pR* and *pL*. The sequences of promoter mutations used in this experiment are shown.

B. Promoter activities and interference. Leftward indicates transcription from the indicated constructs that have been inserted into the *XbaI* site of pBC2 *lacZ* reporter constructs such that *pL* or *pc* directs transcription of the *lacZ* gene. Rightward indicates the activity from equivalent pBC2 constructs placed in the other orientation such that active *pR* or *pe* directs transcription of the *lacZ* gene. The fusions were located on λRS45ΔYA-pBC2-*p#p#.lacZ* prophages, using MC1061.5 as the host strain. pR^+ or pe^+ values indicate the amount of lacZ units when the converging promoter of that construct is active, $pR^=$ or pe^- values are the units obtained for the equivalent construct with a mutated converging promoter. Fold interferences were calculated from ratios of $pR^=$ or pe^- over pR^+ or pe^+ leftward lacZ units from equivalent clones as described for *pRpL* (Chapter2). Errors show 95% confidence limits. In the table of P2 fragments assayed, numbers above the line indicate the position from the predicted +1 site of *pe* or *pR* and numbers below the line indicate the position from the predicted +1 of *pc* or *pL*.

A



B

	Name	Fold Interference	Leftward $pR^{=}$ or pe^{-} pR^{+} or pe^{+}		Construct	Rightward
(a)	**pe(NheI)pc**	**2.1** (± 0.2)	109 (± 5)	51 (± 2)		922 (± 46)
(b)	**pe(NheI)pL**	**3.0** (± 0.1)	211 (± 10)	70 (± 3)		
(c)	**pR(NheI)pc**	**3.5** (± 0.2)	88 (± 4)	25 (± 2)		
(d)	**pR(NheI)pL**	**6.1** (± 0.4)	199 (± 10)	32 (± 2)		1736 (± 107)

To more accurately compare the difference in interference between 186 and P2 promoters, a *NheI* restriction site was also inserted centrally between the 186 switch promoters, but their natural spacing was retained. The intrinsic activity of the constructs, *pR*(*NheI*)*pL* and *pR*$^{=}$(*NheI*)*pL*, was then determined and found to be higher for all promoters compared to those previously measured (compare fig 5.4B(d) with fig 2.3(c)). Interference of *pL* activity was 6.1 ± 0.4 fold, similar to the wild type 186 situation. The difference between interference of equally spaced convergent promoters clearly remained and is thus not a property of the different distances between promoters.

5.3.2. The lytic promoter of P2 is weaker less aggressive than that of 186.

To test if the lytic promoter of P2 was less aggressive than that of 186, and therefore responsible for the lower level of interference, the ability of *pe* to interfere with *pL* activity was examined. *NheI* promoter fragments of *pe* and *pe*$^{-}$ were placed face-to-face with the *NheI* promoter fragment carrying *pL* to produce *pe*(*NheI*)*pL* and *pe*$^{-}$(*NheI*)*pL* (see fig 5.4A). The 186 spacing between –10 regions of the convergent promoters was maintained. Intrinsic *pL* activity measured in the presence of *pe*$^{-}$ was equivalent to that measured in the presence of *pR*$^{=}$, 211 ± 10 versus 199 ± 10 units respectively, but in the presence of active convergent transcription *pe*(*NheI*)*pL* was 70 ± 3 units compared to *pR*(*NheI*)*pL* 32 ± 2 units (fig 5.4(b) versus (d)). The interference of *pL* activity caused by *pe* was 3 ± 0.1 fold, which is half of that observed when *pL* is convergent to *pR* and indicating *pe* to be half as aggressive as *pR*. Aggressiveness is known to be proportional to promoter strength [Elledge and Davis, 1989] and comparisons of rightward transcription by *pe*(*NheI*)*pc* and *pR*(*NheI*)*pL* suggest *pe* to be approximately half as strong as *pR*, 922 ± 46 versus 1736 ± 107 units respectively.

5.3.3. The lysogenic promoter of P2 is less sensitive than that of 186.

The comparative sensitivities of P2 and 186 lysogenic promoters were tested by comparing interference of the two lysogenic promoter in the face of either P2 lytic transcription (see above) or 186 lytic transcription. To measure interference of *pc* in the face of *pR*, *NheI* promoter fragments of *pR* or *pR*$^{=}$ were placed face-to-face with the *NheI* promoter fragment carrying *pc*, to produce convergent promoters *pR*(*NheI*)*pc* and

pR$^{=}$(*NheI*)*pc*. These constructs maintain the 186 spacing between promoters. LacZ assays were performed using MC1061.5 as the host (fig 5.4B(c)). Intrinsic *pc* activity measured in the presence of convergent *pR*$^{=}$ was 88 ± 4 units, slightly lower than that measured previously in the presence of *pe*$^{-}$ (fig 5.4B(c) versus (a)). *pc* activity decreased to 25 ± 2 units in the presence of active *pR* and interference was 3.5 ± 0.2 fold which was an increase from the 2.1 fold due to *pe*. Based on the previous observation that *pR* was a more aggressive promoter than *pe*, some increase in interference was expected, confirming the conclusions about the comparative impact of the two lytic promoters. However, the extent of interference of *pc* by *pR* was still less than that of *pL* interference by *pR* (3.5 fold versus 6 fold). Furthermore section 5.3.2. showed that in the face of *pe* transcription the activity of *pc* was again less interfered with than *pL* (2.2 fold versus 3 fold). Thus, despite the fact that the intrinsic activity of *pL* was nearly double that of *pc*, *pc* is consistently less sensitive to interference than *pL* when either *pR* or *pe* are used as the interfering promoter.

5.4 Is the sensitivity of the lysogenic promoter a reflection of its sitting duck potential?

The weak interference of the P2 promoters was shown to be a consequence of the combination of both a less aggressive lytic promoter and a less sensitive lysogenic promoter compared to those of 186. The potential for a sitting duck mechanism of interference to explain the difference in lysogenic promoter sensitivity to interference is explored here.

5.4.1 Establishing an *in vitro* transcription system for *pc*.

The prediction from the sitting duck model of interference is that the differences of a promoters sensitivity to converging transcription can be explained by its potential to form 'sitting ducks', which is determined by its kinetic properties. This model predicts that the more sensitive 186 lysogenic promoter and the less sensitive P2 lysogenic promoter are weak promoters for different reasons. *pc* should have a poor ability to bind RNAP, but upon binding polymerase the process of escape and clearance should occur rapidly, thus giving little opportunity for collisions with converging, elongating RNAP molecules. In contrast, as shown in Chapter 4, the more sensitive 186 lysogenic

promoter can bind RNAP and form open complexes quickly, but the rate of clearance of bound polymerase is slow, thus increasing the chance for detrimental collisions with converging RNAP. *In vitro* transcription assays performed in Chapter 4 were repeated to discover the comparative sitting duck potential of *pc*.

The 195 bp P2 fragment carrying *pe⁻pc* (as used in fig 5.3(b)) was cloned into the *in vitro* transcription vector pBC1 in an orientation shown in fig 5.5. To confirm the presence of *pc* initiated transcripts in an *in vitro* transcription reaction, *pe⁻pc⁻* DNA was also cloned into pBC1 in the same orientation. The transcript pattern for pBC1-*pcpe⁻* and pBC1-*pc⁻pe⁻* was obtained (fig 5.5) and the unique band for full-length *pc* transcript of the expected size (239 nt) was identified. Three additional major bands were also observed that were not seen using pBC1 alone, these have been labelled X (~190 nt), Y (~153 nt) and Z (~220 nt) (see fig 5.5). The sizes were calculated based on the logarithmic migration of the band compared to that of other bands of known transcript lengths. The origins and potential influence of these transcripts are discussed in section 5.5.2.

5.4.2. Rate of clearance of *pc*.

The rate of promoter clearance was measured for *pc*, as previously performed for *pL* and *pR* (section 4.2.2.). An example of the transcript pattern used for these experiments is shown in fig. 5.6. Results were quantitated and analysed as for *pL* and *pR* (section 4.2.2.). The percentage of the maximum full-length transcript accumulation from *pc* heparin resistant open complexes was plotted against reaction time (fig 5.7), and results for *pL* and *pR* from chapter 4 are included for comparison. The rate of accumulation of full-length transcript from heparin resistant complexes formed at *pc* was dramatically faster than the rate measured for *pL* and more similar to that of *pR*. This is reflected in the calculated time for 50% transcript accumulation which for *pc* was 1 ± 0.1 mins, over 6-fold faster than the $t_{50\%}$ for *pL*, 6.5 ± 0.3 mins, and only slightly slower than the $t_{50\%}$ for *pR*, 0.7 ± 0.1min. The much slower rate of clearance for *pL* compared to *pc* is consistent with the sitting duck model of interference, and may explain the difference in sensitivity of *pc* and *pL* to interference, as sitting ducks accumulate at *pL* in the form of

Figure 5.5

Templates for *pc in vitro* transcription and their transcript patterns.

The templates used to analyse the *in vitro* transcription of *pc* are shown with the orientations of promoter inserts and expected *pc* transcript sizes. pBC1, pBC1-*pRpL*$^{-}$, and pBC1-*pR*$^{=}$*pL* transcript patterns are included for reference. The transcript pattern produced from these templates (pBC1-*pRpL*$^{-}$, pBC1-*pR*$^{=}$*pL*, pBC1, pBC1-*pcpe*$^{-}$ and pBC1-*pc*$^{-}$*pe*$^{-}$) in a standard single round *in vitro* transcription assay are also shown (lanes 1 to 5 respectively). RNA1 and major transcripts unique to the promoter insert are indicated.

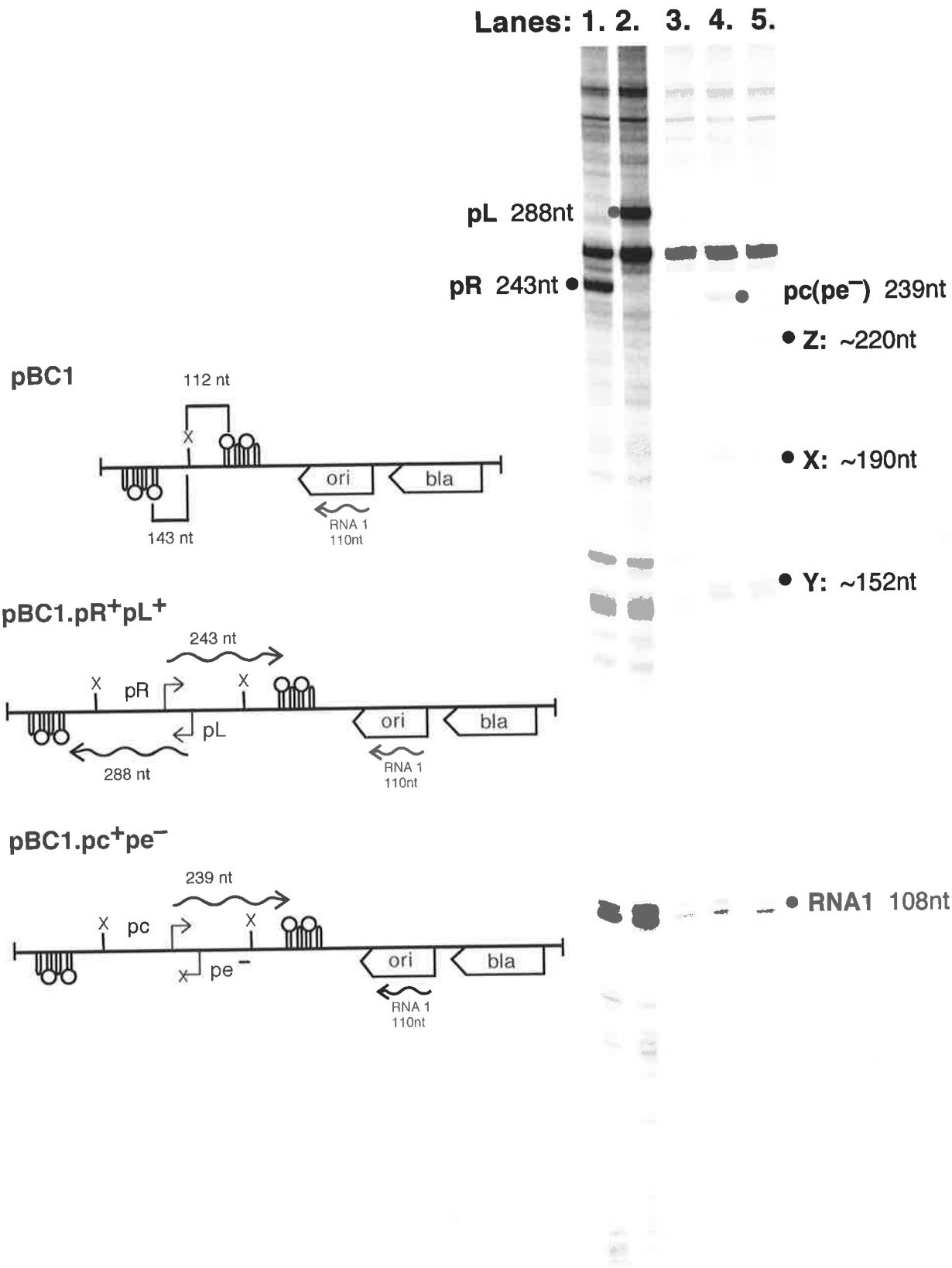
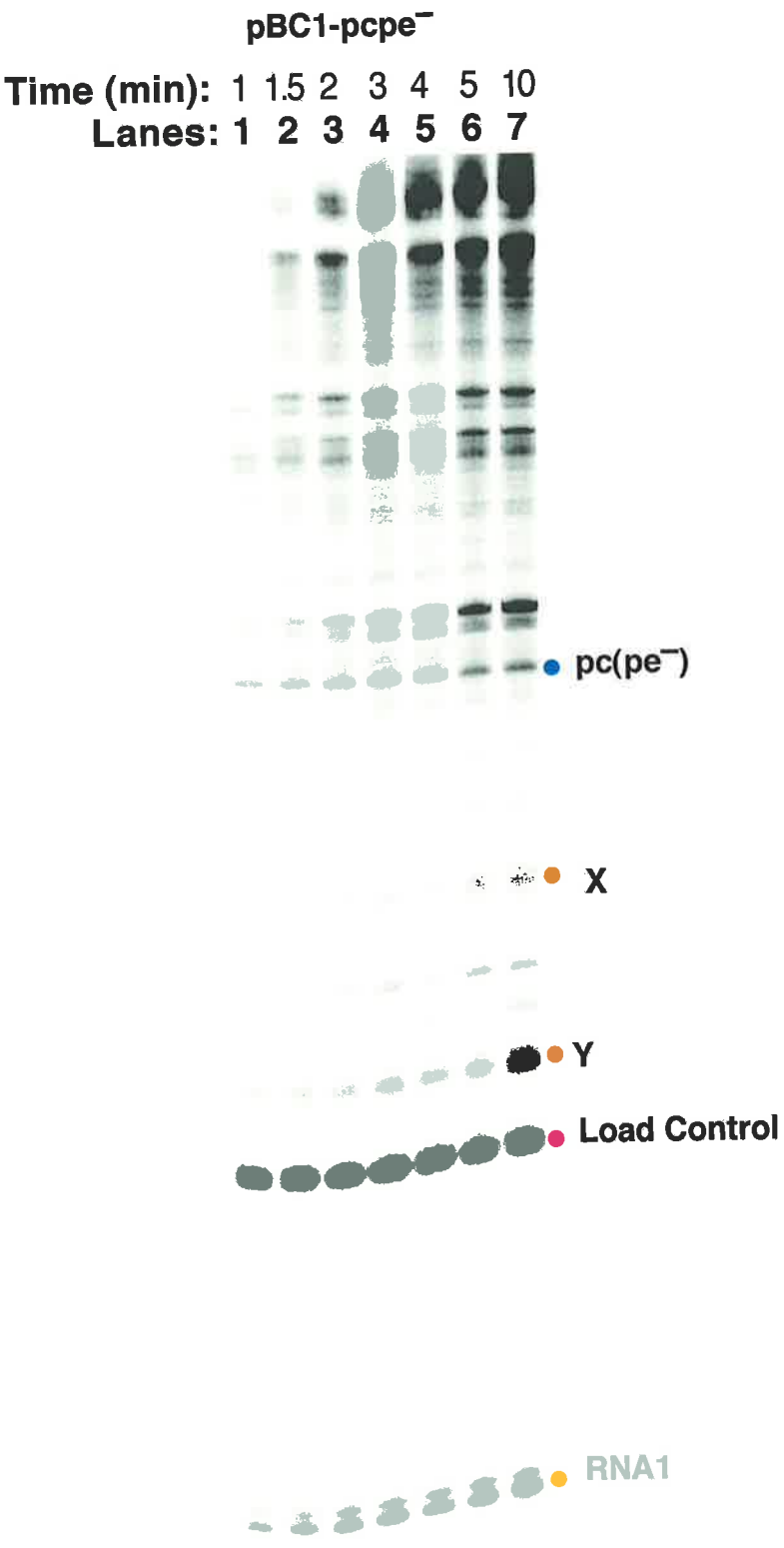

Lanes: 1. 2. 3. 4. 5.
pL 288nt
pR 243nt
pc(pe⁻) 239nt
Z: ~220nt
X: ~190nt
Y: ~152nt
RNA1 108nt
pBC1
112 nt
X
ori
bla
143 nt
RNA 1
110nt
pBC1.pR+pL+
243 nt
X
pR
X
pL
ori
bla
288 nt
RNA 1
110nt
pBC1.pc+pe−
239 nt
X
pc
X
pe −
ori
bla
RNA 1
110nt

Figure 5.6

Transcript pattern showing the rate of full-length transcript accumulation from open complexes formed at pBC1.*pcpe*$^{-}$.

An example of the transcript pattern used to quantitate the rate of accumulation of the transcripts (lanes 1 to 7) from heparin resistant complexes formed at the template pBC1-*pcpe*$^{-}$. *In vitro* transcription clearance assays were performed as described in fig 4.4 where DNA and RNAP were pre-incubated for 30 minutes to ensure complete/stable open complex formation at all relevant promoters, then elongation was initiated by the addition of NTP/heparin mix at time zero. At various times after the addition of NTPs, 7 μl aliquots of the reaction were added to 7 μl of stop/load buffer (which also contained the loading control). 10 μl of each time point sample were then loaded onto a 6% polyacrylamide gel. Lanes 1 to 7 correspond to transcription reactions terminated 1, 1.5, 2, 3, 4, 5 and 10 minutes after the addition of NTPs. This is an example of the pattern from one of two experiments used for quantitation in fig 5.7.


pBC1-pcpe⁻
Time (min): 1 1.5 2 3 4 5 10
Lanes: 1 2 3 4 5 6 7
pc(pe⁻)
X
Y
Load Control
RNA1

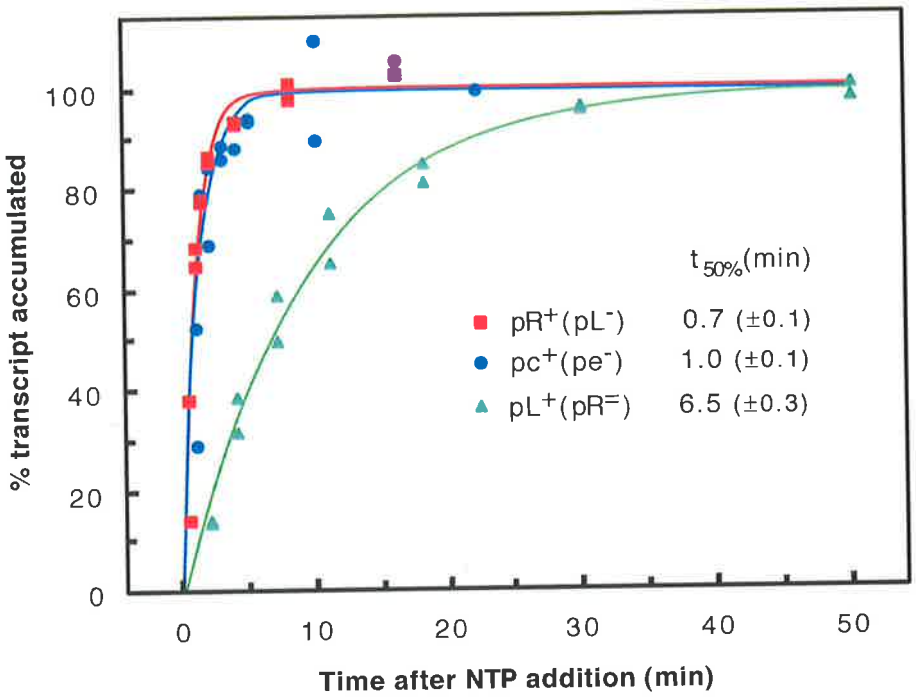


Figure 5.7 Graph comparing the intrinsic rate of full length transcript accumulation from *pc*, *pL* and *pR* open complexes *in vitro*.

Bands for full length transcripts *pc*(pe^{-}) (blue circles), from the transcript patterns of the escape assay described in fig 5.6 were quantitated and plotted against elongation time. Results from fig 4.5 showing the accumulation of full length transcripts for *pR*(pL^{-}) (red squares), and *pL*($pR^{=}$) (green triangles) are included for comparison. Quantitation was as described in fig 4.5 and equation 4.1 was used to calculate the half times ($t_{50\%}$) for promoter clearance shown. Results are from two independent experiment.

open complexes waiting to fire, but open complexes that form at *pc* do not accumulate due to their rapid rate of clearance.

It was possible that the difference in rates of transcript accumulation reflect differences in elongation rates. Full-length *pc* transcript is smaller than that from *pL* (239 nt compared to 288 nt). However, based on an expected *in vitro* elongation rate of about 40-50 nt/sec, (Richardson and Greenblatt, 1996), the difference in the time of transcript elongation of an extra 49 nt was expected to be about a second (assuming no pausing) and could not account for observed differences in the time scale of minutes.

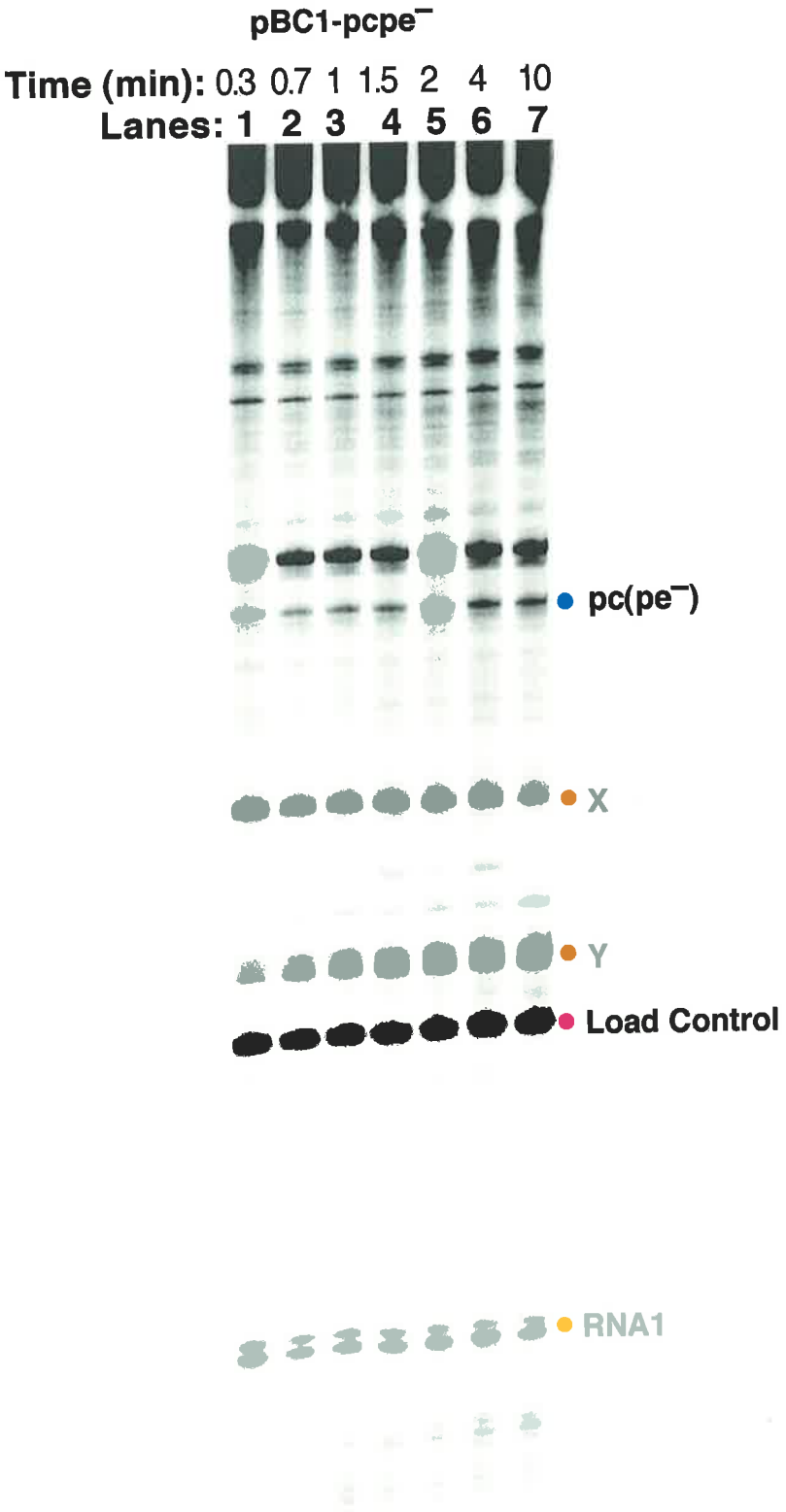
5.4.3. Rate of open complex formation of *pc*.

Although the different rates of clearance are consistent with a sitting duck model to explain interference, the model would also argue that differences in the rate of open complex formation could also explain differences in interference. During convergent transcription a promoter that forms closed complexes which are slow to isomerise will lead to collisions between elongating polymerase and either closed complex or isomerisation intermediate 'sitting ducks'. However, under the conditions used in these *in vitro* experiments maximum levels of transcriptionally competent complexes formed at *pL*, *pR* and *pRNA1* within 20 seconds (section 4.2.3.). This suggested that at excess polymerase concentrations, open complex formation at *pL* is not rate limiting and unlikely to create 'sitting ducks' at this level. To assess the ability of *pc* to form sitting ducks prior to the formation of open complexes, the rate of formation of heparin resistant complexes at *pc* was assayed as described for *pL* (section 4.2.3), using the template pBC1-*pcpe*$^{-}$. The transcript pattern obtained from the reaction is shown in fig 5.8. Bands for full-length *pc* transcript and RNA1 were quantitiated, as described previously for *pL*, then plotted as % of maximum transcript produced against pre-incubation time with RNAP and template (fig 5.9). The amount of RNA1 transcript produced showed a similar pattern of production as RNA1 from pBC1-*pR*$^{=}$*pL*, with maximum levels obtained at the first time point (20 secs) and then the amount of active open complexes seeming to slowly degrade. The accumulation of *pc* transcript however seemed to follow a first order exponential, reaching maximum levels at around 4 minutes, with 50% of competent complexes estimated to form in ~12 seconds. The inactivation of open complexes seen for *pL* and RNA1 was not evident for *pc*. The rate

Figure 5.8

Transcript pattern showing the rate of formation of heparin resistant complexes (or open complex formation) that are able to produce transcripts from pBC1-*pcpe*$^{-}$.

The reaction is as described for fig 4.6 except that the template used in the reaction was pBC1-*pcpe*$^{-}$. Incubation times of DNA and RNAP prior to the addition of heparin and NTPs (pre-incubation time) were 20 secs, 40 secs, 1, 1.5, 2, 4 and 10 minutes (lanes 1 to 7, respectively).


pBC1-pcpe⁻
Time (min): 0.3 0.7 1 1.5 2 4 10
Lanes: 1 2 3 4 5 6 7
pc(pe⁻)
X
Y
Load Control
RNA1

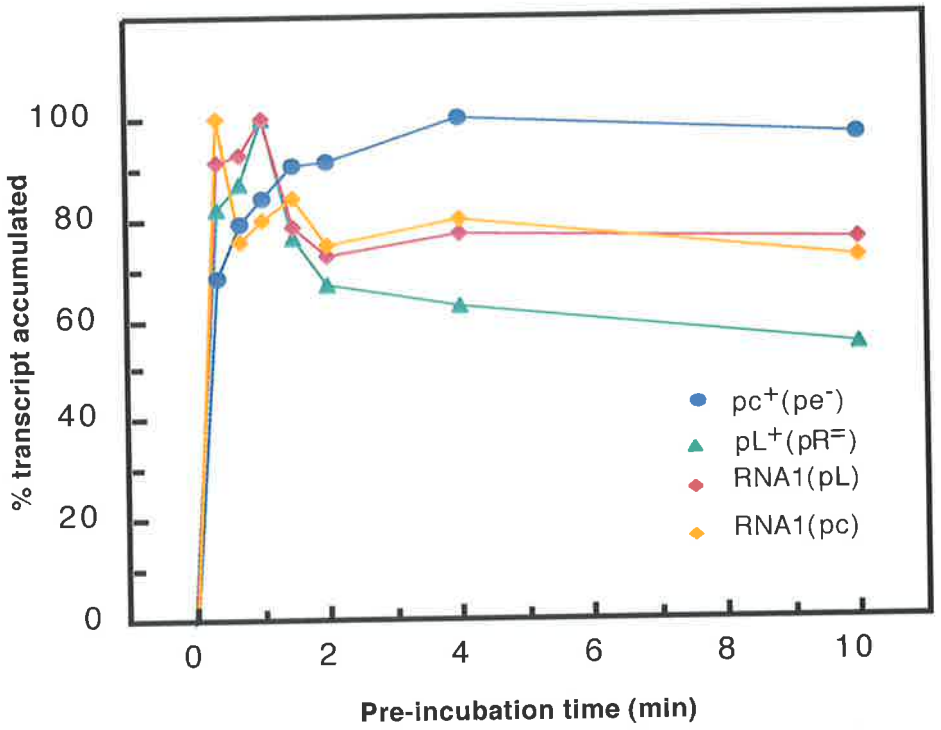


Figure 5.9 Comparison of the rate of formation of heparin resistant complexes (or open complex formation) that are able to produce transcripts from pBC1-*pcpe*$^{-}$ and pBC1-*pR*$^{=}$*pL*.

Bands for full length transcripts for *pc*(*pe*$^{-}$) and RNA1(*pc*) (ie. RNA1 produced from the template pBC1-*pcpe*$^{-}$) from the transcription pattern shown in fig 5.8 were quantitated and plotted against pre-incubation time. Result for transcripts *pL*(*pR*$^{=}$) and RNA1(*pL*) (ie. RNA1 produced from the template pBC1-*pR*$^{=}$*pL*) from fig 4.7 are also shown for comparison. The appropriate bands were quantitated, normalised to the amount of loading control in each lane, and expressed as a percentage of the maximum amount of specific transcript produced for that reaction (for *pc* this was the 4 min time point, for RNA1(*pc*) this was the 20 second time point).

of open complex formation in these conditions therefore appears to be slower for *pc* than for *pL* which may help to explain the weak activity of *pc in vivo* (discussed further in section 5.5). In comparison to the 6.5 minutes taken for promoter clearance from *pL*, the difference in the rate of open complex formation is minimal. The ability to form sitting ducks in the form of closed complexes waiting to isomerise may be greater for *pc* than *pL*, and possibly explains some of the interference observed at *pc*. However this is considered minor compared to the differences in ability of *pL* versus *pc* to form sitting ducks in the form of open complexes waiting to clear.

5.5 Discussion

5.5.1. Why is P2 interference much less than that expected from the literature?

Using the reporter system developed in chapter 2 to measure P2 convergent promoter activities, an unexpectedly low level of interference was measured compared to that reported by Saha *et al* (1987). This low level of interference was not influenced by different host strains. Comparing these two studies, two major differences in the reporter systems used could explain these disparate results: either the type of P2 promoter fragments assayed or the type of promoter assay performed.

It was possible the longer P2 fragment length used by Saha *et al.* (1987) (–187 bp from the *pe* start site to –377 bp from the *pc* start site, see fig 5.2) to determine interference was responsible for conferring additional interference. The additional sequence includes an IHF site upstream of *pe*, which could affect promoter activities and thus interference. This possibility was tested by assaying promoter activities in the presence of equivalent additional DNA upstream of *pe* (data not shown) and no major change in promoter activities was observed. Interference increased only slightly to about 3 fold, which was far from the 30 fold previously reported. The influence of additional P2 DNA upstream of *pc* was not examined, however using the model of interference obtained for 186, it is difficult to imagine how increasing the size of the fragment assayed could increase interference of *pc* by a mechanism that involves the passing of polymerase over *pc*. Increasing the size of the 186 fragment assayed did not alter the extent of interference (see chapter 2). In terms of comparing the mechanism and extent of interference

between 186 and P2 convergent promoters it was more relevant to use constructs containing the minimal promoter regions.

It was possible that the high level of interference observed by Saha *et al* (1987) was due to an overestimate of intrinsic *pc* strength as a consequence of context effects caused by deletion of downstream *pc* sequences, from the BglII site in between the –35 and –10 region of *pe*. *pc* activity was therefore compared between clones producing *pc*-CAT transcripts that differ in their 5' untranslated region by 168 nt. This region includes the translation of a truncated C gene. This translation could easily affect the stability of the transcript, transcription and downstream translation of the CAT gene. The measurement of intrinsic *pc* activity using an equivalent sized *pc* fragment carrying mutations which inactivate *pe* is a more ideal experiment.

Saha *et al* (1987) used plasmid based CAT assays to measure promoter strength, whereas this study used chromosomal lacZ assays. The advantages of single copy chromosomal assays over plasmid based assays have previously been reported where at least three problems with the use of plasmid based assays were identified (Linn and Pierre, 1990; and Simons *et al.*, 1987). (i) Plasmid copy number can vary with the size of the DNA insert and the strength of the cloned promoter (Adams and Hatfield, 1984). The presence of a strong promoter such as *pe* could easily interfere with plasmid replication, reducing the copy number. Thus a potentially lower than normal copy number caused by the presence of active *pe* would skew measurements of *pc* activity, and increase the apparent level of interference. This effect could partially be accounted for by normalising reporter units of a strain to its plasmid DNA content, however the study by Saha *et al* (1987) did not include this control. (ii) Certain complete or truncated gene products can be detrimental to the cell when expressed at high levels from multicopy plasmids. This provides selective pressure for reduced expression of the gene or reduction of in the plasmid copy number. The plasmid constructs used by Saha *et al* (1987) contained over half the *C* gene expressed from *pc* and *pe* expressed the complete *cox* gene whose translation was truncated by introducing a frame shift mutation in the *ClaI* site of the gene (see fig 5.2) (the *ClaI* site was digested and endfilled), and also part of orf 78. (iii) Multiple copies of some genes or over expression of their products could potentially titrate regulatory components present in low copy

numbers in the cell and thus lead to abnormal expression. The presence of multiple copies of the strong promoter *pe* could potentially titrate limiting transcription components, leading to increased interference with *pc* activity. All these problems could result in alterations in CAT expression and measurements of interference, but are not expected to be relevant when using single copy chromosomal lacZ assays, and could account for the discrepancies in interference.

5.5.2. The origin and influence of transcripts X, Y and Z.

Transcripts X (~190 nt) and Y (~152 nt) appear only when the templates pBC1-*pcpe*$^{-}$ and pBC1-*pc*$^{-}$*pe*$^{-}$ are used. Therefore the origins of these bands must be a consequence of the inserted DNA. Transcript Z (~220 nt) is specific to pBC1-*pc*$^{-}$*pe*$^{-}$. Extraneous transcripts must arise from either cryptic promoters not previously predicted, or premature termination (or pausing) of transcripts from known promoters.

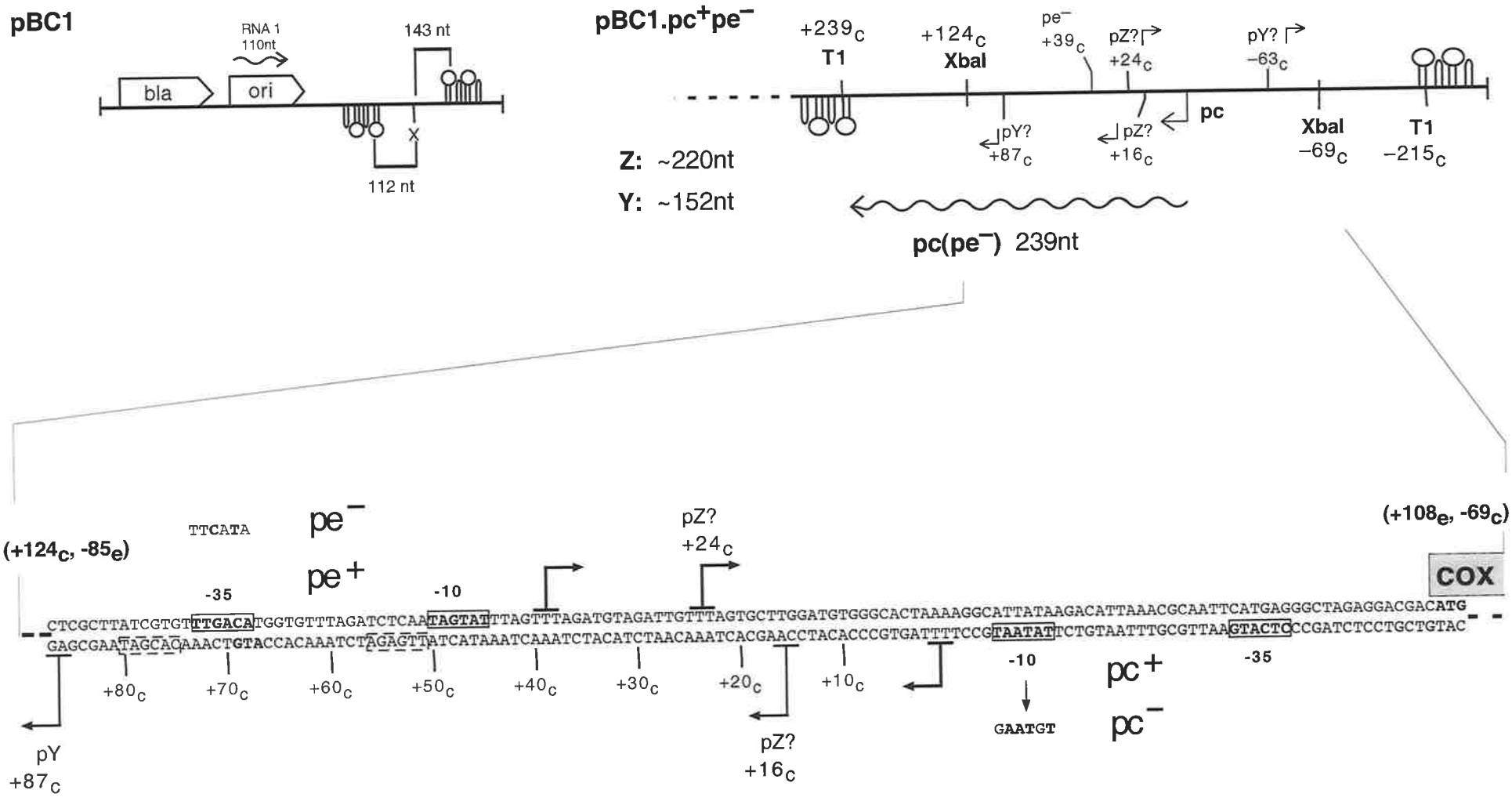
In both the transcript patterns in fig 5.5 and that of the promoter clearance assays in fig 5.6, transcript X is faint. The band for transcript X only becomes significant in the transcript pattern assaying the rate of open complex formation in fig 5.8. Here a strong band is observed at the first timepoint (20 seconds) but this band rapidly loses intensity with increased incubation time of template and RNAP. This indicates that the origin of X is probably a cryptic promoter, pX, on the *pepc* insert which quickly forms open complexes, however these open complexes are unstable and either dissociate or form inactive complexes. At incubation times greater than 15 minutes, transcription from this promoter is very much reduced. The rate of clearance from *pc* was measured from templates that were pre-incubated for 30 mins with RNAP, which explains why transcription from pX is minimal. If most of the open complexes formed at pX have dissociated, then the minimal activity from this promoter in the promoter clearance assay is not expected to significantly alter the clearance rate of *pc*. However if inactive complexes have formed at pX, this may pose some barrier to transcription from *pc*. At the very least this barrier would be expected to hinder the time taken for full-length transcript accumulation from *pc*, thus the conclusion from these experiments that promoter clearance from *pc* is quicker than from *pL* is still valid.

Analysis of transcript Y is complicated by the presence of a faint transcript of an equivalent size present in the transcript patterns from pBC1 alone (fig 5.5). Quantitation of transcript Y in fig 5.5 from the transcript pattern for pBC1-*pcpe*$^{-}$ and the other templates shows that Y is over 3 fold more intense than its background band. If transcript Y was the result of premature termination of elongation from *pc*, then one would expect the accumulation of Y to occur prior to the accumulation of full-length *pc* transcript. Fig 5.6 shows the opposite. Accounting for the background transcript of an equivalent size, the intensity of transcript Y after 15 minutes of elongation is less than half that observed in fig 5.5 and fig 5.8 when elongation continued for 60 minutes. This indicates that the accumulation of transcript Y is much slower than that of full-length *pc*. Thus it is not caused by premature termination from *pc* and must be the product of another cryptic promoter, pY, that is very slow to produce full-length transcripts. Assuming that the *pcpe* insert does not contain any cryptic termination signals it is presumed that transcript Y must originate from a cryptic promoter within the *pcpe* fragment and terminates at one of the T1 terminators either side of the insert. Based on the size of this transcript (~152 nt) the origin of the pY must be either ~87 bp downstream of *pc*+1 (+87$_c$), directing transcription in the same orientation as *pc* (tandem pY), or ~63 bp upstream of *pc*+1 (–63$_c$), directing transcription away from *pc* (divergent pY) (see fig 5.10). The presence of a divergent promoter is unlikely based on the *in vitro* transcript pattern from a template pBC1-*pR*$^{=}$(*NheI*)*pc*, in which *pc* directs transcription in an opposite orientation to pBC1-*pcpe*$^{-}$ with respect to the vector (data not shown). From this template a divergent pY would be expected to produce a transcript 31nt shorter (due to the assymetry of the placement of the terminators from the *XbaI* site used for all cloning), but no such transcript was seen. A tandem pY would not be present on the *pR*$^{=}$(*NheI*)*pc* insert (as the *pe*$^{-}$ region where pY would be located, is absent), and no additional transcripts corresponding to pY were seen. Analysis of the P2 sequence at the location of tandem pY revealed a possible weak promoter sequence with –35 region **TTGAGA**, an 18bp spacer and –10 region CACGAT, (consensus matches are in bold). The *in vitro* properties of this promoter are that it is slow to produce full-length transcripts and is slower than *pc* at forming active open complexes (see fig 5.8), reaching half maximum at 30-40 seconds. These properties indicate it would be a weak promoter *in vivo*. This is confirmed by assay of the leftward *in vivo*

Figure 5.10

Diagram showing the potential location of pY and pZ in pBC1-*pcpe*$^{-}$.

Distances downstream (+) and upstream (–) from the start of pc transcription are given as $\#_c$. Bent arrows indicate likely positions of promoters, and broken boxes indicate potential –10 and –35 hexamers defining pY. For further details see section 5.5.2.


pBC1
RNA 1
110nt
143 nt
bla
ori
X
112 nt
pBC1.pc+pe−
+239c
T1
+124c
XbaI
pe−
+39c
pZ?
+24c
pY?
−63c
pY?
+87c
pZ?
+16c
pc
XbaI
−69c
T1
−215c
Z: ~220nt
Y: ~152nt
pc(pe−) 239nt
(+124c, -85e)
TTCATA
pe−
pe+
pZ?
+24c
(+108e, -69c)
COX
-35
-10
CTCGCTTATCGTGTTTGACATGGTGTTTAGATCTCAATAGTATTTAGTTTAGATGTAGATTGTTTAGTGCTTGGATGTGGGCACTAAAAGGCATTATAAGACATTAAACGCAATTCATGAGGGCTAGAGGACGACATG
GAGCGAATAGCACAAACTGTACCACAAATCTAGAGTTATCATAAATCAAATCTACATCTAACAAATCACGAACCTACACCCGTGATTTTCCGTAATATTCTGTAATTTGCGTTAAGTACTCCCGATCTCCTGCTGTAC
+80c
+70c
+60c
+50c
+40c
+30c
+20c
+10c
-10
-35
GAATGT
pc+
pc−
pY
+87c
pZ?
+16c

lacZ activity from *pc*$^{-}$*pe*$^{-}$ at 8 units (fig 5.3(e)). This low *in vivo* activity suggests that pY is not significant *in vivo* and is not expected to influence *in vivo* measurement of *pc* interference. The presence of an active *in vitro* promoter downstream of *pc* may influence the rate of elongation from *pc*, but as with pX, if any effect occurs this would most likely be negative and would not influence the conclusion that full-length transcript production from *pc* is much faster than that from *pL*.

Transcript Z only appears in the absence of transcription from *pc*. This is consistent with a weak cryptic promoter in the vicinity of *pc* such that RNAP binding at *pc* will block the access and binding of another RNAP molecule to that promoter. The size of the transcript (223 nt) is consistent with the location of a cryptic promoter (pZ) close to *pc*, ie either +16c in the same direction as *pc* or +24c in the opposite direction of *pc* (see fig 5.10). The location of a weak promoter in either direction would be expected to be blocked by polymerase binding at *pc*. The fact that the transcript disappears when *pc* is active suggests that pZ is very weak and RNAP activity at *pc* easily outcompetes the activity of pZ. pZ is not expected to be significant *in vivo* based on the low *lacZ* activity of *pe*$^{-}$*pc*$^{-}$ constructs.

5.5.2. Why does *pc* produce a low level of *in vitro* transcripts?

An inspection of transcript pattern for pBC1-*pcpe*$^{-}$ compared to pBC1-*pR*$^{=}$*pL* (eg fig 4.3 and 5.6) suggests that the maximum amount of transcript accumulated during a single round *in vitro* transcription assay for *pc* is less than that produced from *pL*. Quantitating the maximal values for full-length *pc* and *pL* accumulated in the promoter clearance experiments and then normalising to the activity of RNA1, I found that nearly twice as much transcript was produced from *pL* than from *pc*. These values reflect the amounts of heparin resistant complexes formed at a given promoter which are able to produce full-length transcripts. As explained in section 4.3.1, a low value could reflect either a reduced number of open complexes due to a reduced ability to bind polymerase at this polymerase concentration (25-fold molar excess to DNA template concentration), or a reduction in the fraction of open complexes producing full-length transcripts.

Reduced open complex formation is consistent with a weak promoter that binds RNAP poorly, having a low equilibrium binding constant K_B. If this were the case then the concentration of polymerase used is not saturating for *pc*, and the maximum amount of *pc* transcript accumulated should increase with higher polymerase concentrations (except that experimentally, specific polymerase activity decreased with high excesses of polymerase for all promoters examined). A crude assay observing the rate of open complex formation, which includes the rate of polymerase binding, was performed and indicated a slower rate of open complex formation at *pc* compared to *pL*. Additionally open complex formation at *pc* showed quite a disparate type of accumulation with time compared to *pL*, increasing as a first order exponential to a maximum in about 4 minutes but not decreasing with extended pre-incubation incubation. In contrast both RNA1 and *pL* accumulated a maximum level of transcripts very quickly but then a decrease in active open complexes was observed with time. The absence of a decrease in *pc* transcript accumulation possibly indicates an inability to form inactive, or 'moribund', complexes. This would be consistent with an explanation of poor binding rather than increased moribund formation for the low level of overall *pc* activity *in vitro* (see above). Alternatively this difference in open complex formation could indicate that inactivation of open complexes is occurring over the same time frame in which maximal levels of open complexes are being formed, such that the rates of formation and inactivation superimpose. If this were true then the rate of open complex formation at *pc* may be even slower than indicated.

Reduced transcript production from an equivalent amount of open complexes is characteristic of a higher number of inactive or 'moribund' open complexes forming at that promoter (Susa *et al.*, 2002). The presence of increased moribund complexes could be confirmed by performing footprints on the promoter to determine the occupancy of the promoters after elongation has occurred or analysing the abortive transcription pattern from *pc*. It is known that transcript cleavage stimulatory factors GreA and GreB can facilitate the conversion of the moribund complex into an active clearance-competent complex (Hsu *et al.*, 1995), therefore if an increase in moribund formation is shown, addition of GreA and GreB to the reaction should increase the amount of *pc* transcript produced. For *pc* to fit the criteria of having a reduced potential to form sitting ducks compared to *pL* but also be of an equivalent strength or weaker than *pL*,

then *pc* must have a reduced rate of initiation steps that don't increase its sitting duck potential. The only rate limiting step which is expected to be unaffected by a sitting duck mechanism of interference is that of promoter recognition. Poor binding at *pc* rather than increased moribund complex formation is thus the favoured explanation for its weak *in vitro* activity, as this is more consistent with a weak promoter that escapes interference. However, future experiments are required to distinguish between these possibilities.

Is it possible to gauge the characteristic kinetics of *pc* from its sequence? The sequence of *pc* has a very poor –35 region (CTCATG which contains only two matches to the consensus) and a consensus –10 region (TATAAT). It was once hypothesized that the –35 sequence affected primarily the initial binding of polymerase, and the –10 sequence affected primarily opening of the DNA (Shih and Gussin, 1983). This would suggest that *pc* binds polymerase poorly but rapidly allows opening of the DNA, which would be consistent with escape from interference by a sitting duck model. However, testing of this proposal has found that mutations in the –35 and –10 regions affect both parameters (Szoke *et al.*, 1987) making it difficult to predict the type of promoter *pc* is based on its sequence. The only prediction that can be made with confidence is that *pc* is weak due to its weak –35 region. More kinetic experiments are required better determine the reason for the weak *in vivo* and *in vitro* activity of *pc*.

5.3.4. Conclusions.

The purpose of this study was to test the proposed sitting duck mechanisms of interference between *pRpL* by investigating another example of convergent promoters that demonstrates a different level of interference. The *pepc* promoter system were shown here to produce a 2.2 fold level of interference which is much lower that the 5.6 fold effect of *pR* on *pL*. This is despite the activity of the sensitive (lysogenic) promoter of P2 being the same or lower than that of 186. The differences in interference were shown not to involve spacing differences. An explanation involving either the aggressiveness of the lytic promoters or the sensitivities of the lysogenic promoters was then examined by measuring the interference caused by different 186 and P2 switch promoter combinations while maintaining the same spacing between the –10 regions. The weak interference of the P2 promoters was shown to be a consequence of the

combination of both a less aggressive lytic promoter and a less sensitive lysogenic promoter. The difference in aggressiveness can probably be explained by a difference in the strengths of the lytic promoters. *pR* was nearly twice as strong as *pe* and conferred over one third more interference than *pe*. In the face of *pR* or *pe* transcription, *pc* was shown to be about half as sensitive to interference than *pL*. Using the system developed in chapter 4, it was shown that the rate of promoter clearance from *pc* was 6.5 fold faster from *pc* than at *pL*, suggesting that *pL* has a greater potential than *pc* to form 'sitting ducks' for collisions with convergent RNAP, in the form of open complexes waiting to fire. These results are consistent with a sitting duck model of interference between the closely spaced promoters *pRpL* and *pepc*.

CHAPTER SIX

Final Discussion

Chapter 6

Final Discussion

6.1. The mechanism of transcriptional interference between the convergent switch promoters of 186.

The activities of the convergent promoters in the developmental switch of bacteriophage 186 were measured using an *in vivo* assay system developed in Chapter 2 featuring chomosomal single copy *lacZ*-promoter operon fusions. The lytic promoter, *pR*, was found to be about ten times stronger than the lysogenic promoter, *pL*, found 62 bp downstream in a convergent orientation. Using mutations that inactivate the convergent promoter, it was shown that the stronger promoter reduced transcriptional activity of the weaker promoter by 5.6 fold, and that this interference was not reciprocal. Based on experiments performed in subsequent results chapters, it is proposed that the major cause of this interference is by a 'sitting duck' mechanism, in which *pR* convergent transcription over *pL* negatively affects rate-limited promoter initiation intermediates that form at *pL*. A discussion of how this proposal was conceived follows.

6.1.1. *pR* bound RNAP does not inhibit *pL* activity

The interference of *pL* activity could theoretically be caused by either RNAP activity at *pR* or the result of elongation from *pR*. Experiments in Chapter 3 demonstrated that *pR* bound polymerase does not inhibit *pL* activity by either steric hindrance, competitive inhibition or a roadblock mechanism.

Steric hindrance

Increasing the spacing between the promoters an extra 100 bp brought no loss in interference but rather a slight increase which is inconsistent with a steric hindrance model of interference. Given that the RNAP footprint extends to only 20 bp downstream from the start site of a promoter (Record *et al.*, 1996), and that the spacing between *pR* and *pL* is 62 bp, a lack of steric hindrance is not surprising. This spacing should easily allow for co-binding of RNAP to both promoters. Although co-binding of *pR* and *pL* has not been directly demonstrated it is assumed to occur at least *in vitro*,

on linear templates, based on the studies by Jagura-Burdzy and Thomas, (1997). In their study, mobility shift assays of DNA fragments containing differentially spaced convergent promoters and RNAP, were used to show that simultaneous open complexes could form at convergent promoters spaced 49 and 37 bp apart. The direct testing of co-binding of *pR* and *pL in vivo* remains a formality.

Competitive inhibition

The loss of interference observed when the promoters were rearranged to be divergent, indicated no role for a local polymerase sink mechanism of interference acting at *pR*. This type of interference is expected to be dependent on a limiting free polymerase concentration in the cell, which suggests that this is not the case for the strain used in these assays, grown under the conditions of these assays (ie. *E. coli* strain MC1061.5 grown at log phase in LB medium). This type of mechanism does however remain a theoretical possibility during conditions of limiting concentrations of holoenzyme. These conditions may be experienced *in vivo* using different strains grown at specific conditions. For example a 10 fold difference in free RNAP concentration has been estimated in bacterial strains exhibiting different growth rates as a result of growth in different media (Liang *et al.*, 1999). When an alternative strain was used to measure *lacZ* activities (*E. coli* strain NK7049) absolute values of promoter activities decreasing two-fold, which may reflect differences in holoenzyme concentration between MC1061.5 and NK7049; however no increase in interference was observed when using this strain (chapter 2), again suggesting that holoenzyme concentrations are not limiting enough to allow competitive inhibition. The opportunity for this type of interference to occur *in vitro*, under manipulated limiting concentration of polymerase has previously been suggested by a number of studies, including, the apparent *in vitro* interference observed between the *dnaQ-rnh* convergent promoters when low polymerase:DNA template ratios were used (Nomura *et al.*, 1985b) and between divergent λ promoter P_R and P_{RM} during *in vitro* studies at low enzyme to DNA ratios (Owens and Gussin, 1983). For the *in vitro* experiments performed here, care was taken to ensure an excess concentration of RNAP.

Roadblock mechanism

It is unlikely that the roadblock mechanism is a major cause of interference since there is a loss of interference when transcription from *pR* is terminated between *pR* and *pL* (Chapter 3). This indicates that either RNAP is not bound to *pR* long enough during steady state transcription to become a sufficient roadblock, or that a bound polymerase does not inhibit elongation from a converging polymerase. Consistent with the latter, the single round *in vitro* transcription assays performed in chapter 4, which analysed the outcome of a collision between an elongating RNAP

complex and an open complex, demonstrated no inhibition of the elongating complex. Moreover, if an open complex did affect convergent elongation, then based on the kinetic properties of *pR* and *pL* (Chapter 4), it would be expected that 'open complex roadblocks' would exist longer at *pL* than at *pR*. However, as *in vivo pR* activity was shown to be unaffected by *pL* (Chapter 2), it is likely that promoter bound polymerase poses no obstacle to elongation.

6.1.2. Mechanisms of interference involving elongation from *pR*.

Interference involving elongation could act by a mechanism involving head-on collisions between elongating polymerase molecules, antisense transcription, occlusion of RNAP binding to *pL* and a sitting duck mechanism. For interference of *pL* by *pR* activity, a sitting duck mechanism is favoured.

Head-on collisions

Predictions of the contribution of head-on collisions between elongating polymerases can be made based on calculations of the probability of these types of collisions occurring. With a wildtype spacing of 62bp between promoters, any *pR* initiated polymerase has just over 1 second to collide with RNAP from *pL* before it passes over *pL* +1, assuming polymerase travels at 40-50 nt/s (Vogel and Jensen, 1994). Even if *pR* was as strong as λP_L, which has been measured as firing at an average rate of once every 4.6 seconds (Liang *et al.*, 1999), then the intervening DNA would be free of converging RNAPs 80% of the time. In other words, at most only about 1 in 5 of the polymerases fired from *pL* will be involved in a head-on collision.

These calculations could be flawed if the sequence between *pR* and *pL* contains information that leads to a major decrease in the elongation rate between promoters, such as pausing sites. If these signals exist and interference by collisions was the only mechanism of interference, then doubling of this region should double the amount of interference. The *pR*(+100)*pL* construct contains a direct repeat of the 62 bp co-transcribed region (plus a further 38 bp) (Chapter 3). Although interference from this construct increased from 5.6 to 7.2 fold, its extent indicates that if elongating polymerase collisions do contribute to the increase in interference, their role during the interference of naturally spaced *pRpL* is only minor. In addition, a terminator located between *pR* and *pL* (now spaced 206 bp apart) resulted in a 3.3 fold level of interference. In this construct the tA terminator is placed at least 70 bp downstream from *pR*, this sequence retains the native 62bp co-transcribed region and is expected to be co-transcribed at the same frequency as wildtype *pRpL*. If the 5.6 fold interference observed for wildtype *pRpL* was solely due to head-on collisions, then interference of *pR*(tA^+)*pL* should have remained at 5.6 fold rather than the measured 3.3 fold. Although this may suggest that

collisions could contribute at least 3.3 fold interference, much of this residual interference can be explained by incomplete termination of *pR* transcription. While collisions are not the major mechanism operating for *pR* and *pL*, the fact that interference did increase minimally for promoters with increased spacing suggests some potential role for collisions in the interference seen. The influence of head-on collisions for more widely spaced promoters is explored later in discussions of a general mechanism of interference.

Antisense transcription.

A closely spaced divergent *pR* producing antisense *pL* transcripts was used in chapter 3 to show that production of the antisense transcript alone does not confer interference. The potential of transcriptional interference caused by the production of antisense transcripts from different convergent promoters was also investigated by a number of other groups, using an *in trans* supply of antisense transcript of plasmid origin (see introduction 1.4.4.). My findings are consistent with those of Elledge and Davis (1989) which showed that interference in their promoter system only occurred when the promoters were located on the same plasmid (*in cis*), but not when located on different plasmids (*in trans*). However, for the convergent promoters *pc* and *pe* of bacteriophage P2, a three fold inhibition of weak *pc* activity was reported when the *pc* antisense transcript was supplied *in trans* from a plasmid bearing strong and active *pe* (Saha *et al.*, 1987b). This study used a multi-copy, CAT assay to analyse the activities of P2 promoters. In Chapter 5, a more reliable single copy *lacZ* reporter assay was used to re-analyse P2 promoter activities, and those studies showed a marked disagreement with the measurements of these promoters. This fact casts some doubt on the validity of the *in trans* experiments.

It is possible that interference by antisense RNA requires that the complementary sequences in fact directly overlap (ie. the antisense transcripts are transcribed from the same DNA location). However in the *pR*(tA$^+$)*pL* construct (described below), all the *pR* transcripts, including those terminated at tA, show at least 70 bases of antisense and overlapping sequence with the *pL* transcripts that have extended past tA. If overlapping antisense transcript was required for interference then interference in the *pR*(tA$^+$)*pL* construct (3.3 fold) should have been as high as the native *pRpL* construct (5.6 fold), which has 62 bp of overlapping and antisense sequence.

A mechanism of interference involving an overlapping transcript and an elongating polymerase is a 'tethering' mechanism. In the event of simultaneous elongation from *pR* and *pL*, the close proximity of the nascent 3'antisense transcripts could lead to the tethering of one RNAP to the other via the formation of a 62 bp double stranded RNA which could then lead to the restricted progress of each

polymerase. The influence of a tethering mechanism was tested by observing interference from constructs that contain a RNaseIII site downstream of *pL* but before the *lacZ* gene. Clipping of the transcript by RnaseIII at this position should release the *pL* transcript attached to RNAP from any tether formed by overlapping *pR* transcription. With the inclusion of an RnaseIII cleavage site interference was measured at 4.8 fold. This small reduction in interference compared to the 5.6 fold interference of *pL* activity in the absence of a RNaseIII cleavage site, indicated that a mechanism involving overlapping and tethered RNA makes at best only a minor contribution to the overall interference of *pL* activity.

The lack of interference by antisense transcripts suggests that the formation of double stranded RNA (dsRNA) at the start of the *pL* transcripts either do not affect transcription or do not form. The likelihood of dsRNA forming is supported by primer extension analysis of pE transcripts *in vivo* attempting to locate the start of the pE transcript (I. Dodd, personal communication). These experiments resulted in extension of the primer only to the region where *pR* transcription begins. Two possible interpretations of this result exist, both requiring the formation of dsRNA: (i) dsRNA forms at the 5' ends of *pL* transcripts and prevents primer extension beyond the point of complementation, (ii) dsRNA forms but is quickly degraded by the dsRNA ribonuclease activity of the cell. Long stretches of double stranded RNA can be substrates for rapid degradation by the ribonuclease RNase III (Nicholson, 1999). Therefore, if *pL* transcripts bind to complementary *pR* transcripts, then the double stranded RNA could potentially be quickly degraded by RNase III. It is possible that mechanisms of interference involving the antisense transcript will be more significant in strains that lack the function of these RNA degrading enzymes. However, the fact remains that the natural interference seen is not due to antisense hybridisation.

Elongation over pL

The placement of a unidirectional intrinsic terminator between *pR* and *pL* significantly reduced interference from 9.2 fold to 3.3 fold. This is consistent with the cause of interference being elongation from *pR* passing over *pL*. Transcription over *pL* would be expected to interfere with initiation events at *pL* and could involve either promoter occlusion or a sitting duck mechanism of interference.

6.1.3. Interference by elongation over *pL*.

The transient, multiple round nature of this mechanism.

Interference is a mechanism of repression. However, the major difference between repression of transcription by the binding to DNA of conventional repressor molecules and transcriptional interference by elongation over *pL*, is that the negative signal of conventional repression involves the static binding of a repressor complex to DNA at a defined equilibrium, whereas the negative signal of interference is transient and only works whilst RNAP is passing over *pL*. The transient nature of this negative signal means that the extent of interference will not only depend on how much 'damage' to *pL* activity each wave of this negative signal does, but also the rate and duration of the signal and the rate of recovery of *pL* from this signal. For a negative signal to have a lasting impact such that interference becomes significant then it must do enough damage to *pL* initiation such that *pL* does not have time to fully recover before being hit by the next damaging round of transcription from *pR*. The balance between the rate and extent of damage by *pR* elongation versus the rate of recovery by *pL* is crucial in determining the amount of interference that will occur.

Promoter occlusion.

Promoter occlusion (see fig 4.1A), originally defined by Adhya and Gottesman (1982) to explain the interference by tandem promoters, is the process where RNAP molecules initiating upstream block the access of other RNAP molecules to downstream promoters. This can result from direct steric hindrance or from distortion of DNA structure. A theoretical analysis of occlusion is given (summarised in table 6.1).

The *rate of damage* will depend on the rate of *pR* transcription. The rate of *pR* transcription has not been directly examined but can be estimated from comparisons of lacZ activities of *pR* with promoters of known firing rates. This can be predicted from the speed of transcription over *pL* and the frequency of *pR* transcription. The rate of firing of λ*pL in vivo* is estimated to be once every 4.5 seconds (Liang *et al.*, 1999). Based on comparative *lacZ* reporter studies, 186*pR* is about 4-fold less active than λ*pL* in reporter assays (I. Dodd, personal communication), and therefore will be expected to fire once every 18 seconds. Note that this comparison is not ideal and a more complete study of interference will require direct measurements of the transcription rate of *pR*.

The *duration of damage* will depend on the time that binding of RNAP to *pL* is blocked by one round of elongation from *pR*. This will depend on the nature of occlusion and the rate of transcription from *pR* over *pL*. In the case of polymerase binding to *pL* being blocked by only

Mechanism of Interference	Occlusion	Sitting duck mechanism
Rate of damage	rate of pR transcription: ~every 18 seconds	rate of pR transcription: ~every 18 seconds
Duration of damage (maximum)	Steric hindrance (105bp): 2-2.6 seconds Topological changes (240bp): 4.8-6 seconds	not applicable
Extent of damage	100%	minimum of 33%
Recovery rate	rate of diffusion (effectively immediate)	the rate of pL initiation (maximum): ~10 times that of pR initiation (180 seconds)
Maximum expected interference (measured is 5.6 fold or 82%)	Steric hindrance: (2.6/18) 1.2 fold or 14% Topological changes: (6/18) 1.5 fold or 33%	$1/(1-0.33)^{10}$ = 55 fold or 98.2%

Table 6.1
Predictions of the extent of interference by different mechanisms.
For details see the text.

physical hindrance from elongating RNAP, this will be determined by the overlap of the DNaseI footprints for elongating polymerase (about 30 bp (Record *et al.*, 1996)) and a closed complex (about 75 bp (Record *et al.*, 1996)). From the initial point of overlap to the last, this is a maximum elongation distance of 105 bp (the length of an elongating RNAP plus the length of a closed complex). If binding is blocked by transcriptionally induced changes in *pL* DNA topology in addition to steric hindrance, then the length of transcribed DNA that could cause damage will be larger. This distance is difficult to predict, however the fact that interference decreased by the placement of terminator ~70 bp downstream of *pL* indicates that the front edge of possible topological changes must be less than this. Additionally, the fact that divergent transcription occurring 171 bp upstream (from the start sites of transcription) did not inhibit *pL* activity suggests that the back edge of topological changes must be less than this. This means that if topological changes are responsible for occlusion then the maximum elongation distance this could occur in will be ~240bp. For transcription of mRNA in *E. coli*, the speed of elongation has been measured at 40-50 nucleotides per second *in vivo* (Vogel and Jensen, 1994). Assuming no significant pause sites exist for convergent transcription over the *pL* region, the duration of damage by occlusion will be a maximum of 2-2.6 seconds for steric hindrance or 4.8-6 seconds for topological changes.

The *extent of damage* is assumed to be 100% for the time of occlusion, ie. whilst the promoter is being blocked, no polymerase can bind.

The *recovery rate* is expected to be very rapid because once RNAP has passed over *pL*, the promoter region is immediately available for binding by free σ^{70}RNAP, the rate of promoter recognition will depend on diffusion rates and free polymerase concentrations but is generally thought to be very rapid (Record *et al.*, 1996).

The *extent of interference* by occlusion is therefore dependent only on the rate and duration of the damaging signal. Combining these parameters leads to a maximum possible occlusion of 2.6 seconds every 18 seconds by steric hindrance or 14%, and 6 seconds every 18 seconds by possible topological changes or 33%. Therefore at the most, occlusion from *pR* will contribute a level of interference of only 1.5 fold. For a 5.6 fold level of interference (or 82% reduction of *pL* activity) a much faster transcription rate from *pR* will be required. Occlusion may become more significant for promoters with very high rates of transcription such as the maximal activity of the *rrn* promoters, shown to be more than one transcript per second at high growth rates (Liang *et al.*, 1999).

The possibility of occlusion at *pL* would be more likely if *pL* contains sequences which reduce the speed of convergent elongation over the promoter (eg. pause sites). The presence of pause sites is defined by particular sequence elements (such as hairpin loops) but a slow rate of elongation can also be induced by extrinsic factors such as ppGpp (Mooney *et al.*, 1998). Pausing can be experimentally determined by following the production of transcripts from a promoter *in vitro* with time. A pause site is indicated by the transient appearance of shorter than full-length transcripts prior to the production of full-length transcript. Although this is equivalent to the clearance assays performed in Chapter 4, determination of pausing from these experiments is not ideal. An assay for pausing should examine shorter elongation time points and be able to observe shorter transcripts (possible with electrophoresis using a higher percentage polyacrylamide gel) and preferably from only a single promoter (possible by using DNA minicircles (Choy and Adhya, 1993)). The possibility of pausing of *pR* elongation over *pL* affecting the extent of occlusion remains to be tested.

The significance of occlusion has been discussed in terms of theoretical predictions; there is also some experimental evidence against occlusion. The extent of occlusion as discussed here will be independent of the properties of the occluded promoter (with some notable exceptions, see later). Therefore if the majority of interference that occurs at *pL* is due to occlusion by *pR* transcription over *pL*, then the extent of interference of another promoter, eg *pc*, placed convergent to *pR* should be equivalent. Additionally, interference of *pL* and *pc* by convergent transcription from *pe* should also be the same. In Chapter 5 it was shown that in the face of either *pR* or *pe*, interference of *pL* activity was consistently greater than that of *pc*. This supports the idea that occlusion at *pL* is not a significant cause of interference. The exception to this analysis is that occlusion by topological changes may depend on a promoter's sensitivity to these changes at the level of polymerase binding. Additionally, if the elongation rates over *pL* and *pc* differ dramatically due to pause sites (for example) then occlusion will differ.

A sitting duck mechanism.

If occlusion was the only mechanism of interference, then any polymerase molecules that manage to bind *pL* between waves of *pR* transcription must not be inhibited by the next converging polymerase, and go on to produce full length *pL* transcripts. This will happen either if the speed of *pL* initiation is faster than the time between waves of blocking polymerase or if initiation steps of *pL*, subsequent to promoter binding, are unaffected by *pR* activity. However, if RNAP initiating at *pL* is sensitive to *pR* elongation, and the rate of *pL* initiation is slow, then the sitting duck mechanism becomes possible. If the strength of *pR* is 10 fold greater than *pL*, then presumably the

rate of initiation from *pR* is 10 fold faster than that of *pL*. Thus it is likely that before one round of initiation events are completed at *pL* there will be on average 10 polymerase molecules (initiated from *pR*) passing over the *pL* region. This elongation from *pR* is most likely going to collide with the first initiation intermediate formed at *pL* that is slow to progress to the next step. A priori it seems likely that this collision would affect the initiating complex.

Simplistically, the rate limiting step of a weak promoter can define two types of sitting ducks. A promoter that is slow to form open complexes from closed complexes will produce sitting ducks in the form of isomerisation intermediates. Whereas a promoter that is slow to clear open complexes that form quickly will produce sitting ducks in the form of open complexes waiting to fire (presumably in the process of abortive initiation). A promoter that exhibits poor binding of RNAP (ie has a poor ability to form stable closed complexes) but once stable closed complexes form they progress rapidly through the process of isomerisation and clearance, will not be expected to accumulate initiation intermediates that are sensitive to collisions with elongating polymerase. This third type of promoter would be expected to escape a majority of the interference by a sitting mechanism.

To examine the potential of *pL* initiation intermediates to form sitting ducks, the rates of formation of open complexes and promoter clearance were examined *in vitro* using conditions similar to those found in the cell. *pL* was shown to efficiently form heparin resistant complexes with RNAP but these complexes were slow to clear the promoter and produce full-length transcripts (Chapter 4). In fact there was time for at least 3 complete rounds of *pR* transcription before even 50% of the complexes at *pL* produced transcripts. It was concluded that *pL* has the ability to form sitting ducks (in the form of open complexes waiting to clear) which are susceptible to collisions with elongation complexes from *pR*. To demonstrate that the sitting duck mechanism operates at *pL*, it was necessary to show that collisions between open complexes formed at *pL* and elongation complexes from *pR* were detrimental to *pL* activity. Using the *in vitro* transcription assay developed for measuring promoter properties, each collision between elongating RNAP from *pR* and heparin resistant sitting duck polymerases at *pL in vitro* was shown to reduce the activity of the *pL* sitting duck by at least one-third. During these collisions the activity of the elongating polymerase from *pR* was uninhibited, consistent with the non-reciprocal nature of interference.

To estimate the extent of interference by a sitting duck mechanism, based on these observations, the same analysis used for occlusion was applied (summarised again in table 6.1.).

The *rate of damage* will occur at the rate of *pR* transcription, previously predicted to be every 18 seconds. Damage is caused by the single moment of polymerase collision, thus the duration of the damaging signal will be irrelevant. The *extent of damage* from one round of *pR* transcription, as measured in Chapter 4 using *in vitro* transcription, is questionable depending on the occupancy of *pR* in the reactions, however at the least, collisions inactivated about one third of the open complexes at *pL*. The *recovery rate* from inactivating collisions was not measured, however assuming that these collisions completely dissociate RNAP from the open complex, leaving *pL* available for binding and re-initiation, it will take as long as the normal rate of initiation events at *pL*. Based on the experiments performed here, open complex formation at *pL* may quickly occur before the next round of *pR* transcription, however these open complexes will still need to progress through the slow events of promoter clearance before being able to fully recover from the initial damage. As *pR* is ~10 fold more active than *pL*, assuming open complexes at *pL* form rapidly *in vivo*, each open complex will be hit with on average 10 rounds of elongation from *pR*. Even if each round damages only 5% of the open complexes at *pL*, because of the expected slow rate of recovery from this damage, successive rounds of *pR* transcription will amplify this damage such that the extent interference quickly escalates. Calculations of the *extent of interference* caused by this compounding damage is given by the equation:

$$I = \frac{1}{(1-d)^n} \qquad \textbf{equation 6.1}$$

Where I is the fold interference, d is the extent of damage caused each collision event (expressed as a fraction of the activity of *pL*) and n is the number of damaging rounds of elongation that pass over *pL* before one round of *pL* transcription will be completed. For a 5.6 fold interference and an average of 10 rounds of transcription from *pR* for every round from intrinsic *pL*, the amount of damage required per round will be 16 %. The damage caused to open complexes at *pL* by *pR* transcription *in vitro* is at least twice this. The amount of interference expected for 10 round of 33% damage is 55 fold. Thus 33% damage from each round could easily account for the observed 5.6 fold interference. In fact these calculations suggest that interference of *pL* caused by a sitting duck mechanism should be far greater than 5.6 fold. There are two possible explanations for this apparent lack of interference: i) there exists a large variation in the firing intervals of *pR* and *pL*, or ii) some *pL* intermediates can resist and recover quickly from collisions. The frequency of sitting duck collisions described by equation 6.1 does not account for variation in the distribution of *pR* and *pL* firing times. Promoter initiation is expected to be a stochastic process such that although the average time of initiation events at *pL* will occur 10 fold slower than those at *pR*, it is possible that

for example one sixth of the initiations at *pL* may be fast enough to fit into the gap between *pR* initiations. Thus, initiation from *pL* would occasionally be able to occur with minimal damage. A large distribution of initiation rates would also lead to the recovery rate of some damaged *pL* intermediates being more rapid than that used for the calculation of interference by equation 6.1. Based on this analysis of a sitting duck mechanism, if every collision of elongating RNAP and open complexes caused 100% of the open complexes to dissociate then far more interference would be expected. As only 5.6 fold interference was observed, this leads to the surprising conclusion that a proportion of collisions are not detrimental and open complexes have the ability *in vivo* to withstand a challenge from a converging RNAP molecule passing over the DNA it has bound. This result was also suggested *in vitro*. It would be interesting to further investigate the consequences of these collisions, to examine how polymerase can accommodate this polymerase traffic with only minor structural or topological consequences.

Interference in P2: Support for the sitting duck mechanism.

Support for the sitting duck mechanism of interference at *pL* was obtained by investigating another example of convergent transcription, the lytic/lysogenic switch of bacteriophage P2 (chapter 5). Using the reporter assay developed in Chapter 2, the intrinsic strength of each promoter was approximately equivalent to those of *pR* and *pL*, yet the interference between the P2 promoters was measured as being lower than that of 186, ie. 2.2 fold rather than 5.6 fold. The reason for this lack of interference was shown to be due to both a less aggressive lytic promoter and a more sensitive lysogenic promoter, rather than a difference in promoter spacing. The sitting duck model of interference would predict that a promoter's reduced sensitivity to interference is a consequence of a reduced ability to form sitting ducks for collisions. The ability of a weak promoter to escape from interference yet still retain its weak activity could be achieved by having a very weak ability to bind polymerase, but once bound, the subsequent steps of promoter initiation occur rapidly. The ability of *pc* to form sitting ducks was measured using the same *in vitro* transcription assays utilised for *pL*. Promoter clearance at *pc* was shown to occur far more rapidly than at *pL*, which was consistent with a sitting duck mechanism of interference to explain the differences in sensitivity. Open complex formation at *pc* was shown to be not as rapid as that at *pL* suggesting some possibility of *pc* to form sitting ducks at the level of isomerisation intermediates. This may explain some of the low level of interference still occurring at *pc*. However the experiment examining the rate of open complex formation does not discern between promoters that are slow to form closed complexes and those that have slow rates of isomerisation. The full potential (or lack thereof) of sitting duck interference at *pc* as well as the ability of *pc* to bind RNAP (compared to *pL*) remains to be investigated.

Future experiments.

For a more thorough proof of the sitting duck mechanisms it would be interesting to see if changing the properties of *pL* and *pc* could change their sensitivity to interference in a manner predicted by the mechanism. Sequence determinants for reduced promoter clearance have been shown to involve the core promoter recognition region, the initial transcribed sequence (ITS) and the conformational state of the template DNA (Hsu, 2002). It would be interesting to know whether swapping the ITS of *pc* and *pL* could also swap the clearance rates of either promoter and their sensitivity to interference.

Similarly, it would be of interest to examine the extent of interference for a range of different promoters with a variety of known promoter properties when placed face-to-face with *pR*, at a similar distance to that of *pL*. The kinetic properties of a large number of different promoters have been measured. Ideally it would be best to use a set of promoters which are derivatives of the same promoter, which all exhibit a similar weak activity but which have a range of different rate limiting steps.

The siting duck mechanism of interference assumes that *pR* transcription over *pL* interferes with initiation events of *pL* rather than elongation. Although this is implied by the terminator experiment, a reduction in *pL* initiation during interference has not been directly tested. Direct examination of *pL* initiation could be examined by either probing the amount of initiation occurring at *pL* during interference. Experiments in chapter 4 demonstrated a reduced ability of *pL* to clear the promoter. Promoter that are slow to clear can produce extensive potassium permanganate footprints during steady state *in vivo* expression, owing to the long duration of open complexes formed at these promoters (Ellinger *et al.*, 1994). (Potassium permanganate probes single stranded stranded DNA such as that formed in the active site of an open complex (Sasse-Dwight and Gralla, 1989)). It is therefore expected that intrinsic *pL* will also produce an extensive $KMnO_4$ footprint *in vivo*. A decrease in intensity of this footprint would be expected during interference, if there is indeed a reduced amount of open complex formed at *pL* and if interference by elongation from *pR* causes a reduction in *pL* initiation. Alternatively, elongation from *pL* could be probed using a similar strategy to that used by Prescott and Proudfoot (2002) to show that transcriptional interference in yeast is caused by collisions. The relative quantities of *pL* transcripts could be assayed at different positions downstream of *pL* possibly by primer extension analysis using different primers. If interference reduces initiation at *pL*, then the quantity of all *pL* transcript lengths will remain relatively constant in the presence and absence of *pR* transcription, but will be reduced overall when *pR* is active (at a level that reflects the *pL lacZ* activites). If *pL* elongation is reduced but initiation is

not, then during interference the quantity of transcripts probed at the immediate 5' end would be equivalent to that produced for intrinsic *pL* activity, however probes further downstream will progressively detect lower levels of transcript compared to the intrinsic activity as elongation is terminated due to collisions.

6.1.4. Alternative mechanisms of interference involving elongation from *pR*.

Interference by elongation over *pL* could cause interference by either a direct interaction with elongating polymerase from *pR* and initiation events occurring at *pL* (eg. occlusion by steric hindrance or as a result of collisions with sitting ducks), or as an indirect consequences of *pR* elongation. Two possible indirect mechanisms involve transcriptionally induced changes in other host proteins or in DNA toplogy. The possible influence of changes in DNA topology has been previously raised with respect to occlusion, but is more formally addressed here.

Transcriptionally induced changes in host proteins.

The most likely host proteins which could affect *pL* activity in a *pR*-dependent manner are DNA-binding proteins such as the architectural proteins (eg. H-NS, IHF and HU) which are known to affect promoter initiation in a variety of ways both positive and negative (Atlung and Ingmer, 1997; Goosen and van de Putte, 1995). It is possible that transcription from *pR* could either increase or decrease the binding of these proteins to negatively regulate *pL* initiation. A weak IHF binding site has been predicted upstream of *pR* within the *CI* gene. However I do not expect this site to be involved in interference of *pL* as comparisons of interference between the short 186 used in these experiments which do not carry this IHF site and longer clones containing the site do not indicate an increase in interference (see Chapter 2). However, a possible influence of IHF is suggested in the interference of P2 promoters, as the introduction of additional DNA upstream of *pe* was shown to increase interference from 2.2 fold to 3 fold. This additional DNA carried a known IHF binding site. To demonstrate whether architectural proteins or other host proteins are required to mediate interference experiments could be performed to attempt to reproduce transcriptional interference in an *in vitro* environment. Attempts to demonstrate interference of *pL in vitro* were complicated by the inability to achieve mutiple round transcription. However, the demonstration of damage to *pL* caused during a single round *in vitro* transcription from *pR* does show that no additional proteins are required for interference caused by a sitting duck mechanism.

DNA topology.

It has been suggested by a number of authors that transcriptional interference could be caused by changes in DNA topology that occur as a result of convergent transcription (Brantl and Wagner,

1997; Elledge and Davis, 1989; and Eszterhas *et al.*, 2002). Active transcription of topologically constrained DNA, such as chromosomal DNA, is known to cause changes in the local DNA supercoiling, with waves of positive supercoils induced ahead of transcribing polymerase and negative supercoils found behind (Liu and Wang, 1987). As changes in DNA topology have been known to inhibit or promote the expression of a number of promoters (Opel *et al.*, 2001), it is conceivable that the changes in topology caused by transcription of *pR* over *pL* could be responsible for interference. To simplify discussion of topological changes, the potential effects on *pL* activity are categorised in terms of interference mechanisms that act through the DNA over a long-range, mid-range or short-range. The DNA-range of action is defined by the physical distance between an open complex formed at *pL* and an elongating RNAP from *pR* which is either before *pL* or after.

Possible long-range effects. The ability of elongating RNAP to act on *pL* initiation at a long-distance by changing DNA topology has been shown to be minimal by two experiments. A long-range supercoiling mechanism of interference would be predicted to remain even if *pR* transcription was stopped before it reached *pL*. However when interference using the *pR*(tA$^+$)*pL* construct was examined, a decrease in interference from 9.8 to 3.3 fold was observed (fig 3.2). Although a supercoiling mechanism may be responsible for part of the residual 3.3 fold interference, it is not considered to make a major contribution. Therefore, any potential long-range topological affect induced by an elongating polymerase in front of *pL*, which has at least 70 nt between its site of polymerisation (the point of tA termination) and the start site of *pL* transcription, is negligible. If *pL* activity is negatively affected by long-range changes in DNA topology caused by elongation from *pR* that has passed over *pL*, then this interference should have remained when *pR* was placed divergent to *pL* (ie. *pR*(div)*pL* or *pR*(divMM)*pL*. No effect on *pL* activity, either positive or negative, was observed for active divergent transcription. Therefore, any potential long-range topological effect induced by an elongating polymerase after *pL* which has at least 171 nt between its site of polymerisation (the start site of divergent *pR*) and the start site of *pL* transcription, is also negligible.

Mid-range effects. I have defined mid-range effects as those which act at a lesser distance to those tested by the clones *pR*(tA$^+$)*pL* and *pR*(div)*pL* (ie. +70 bp to –171 bp) but greater than the distance at which elongating RNAP from *pR* physically collides with an open complex formed at *pL* (ie. +40 to –75, see following). The point at which physical collisions will occur can be roughly estimated from the point at which DNaseI footprints for either complex would begin to overlap. An elongating RNAP protects 40bp of DNA from cleavage, with the polymerisation site being roughly in the middle of the complex. An open complex protects DNA from +20 to –55 from the start of

transcription (Record *et al.*, 1996). Thus physical collisions between molecules would be expected to occur when the site of polymerisation in the elongating polymerase is ~40 nt downstream and ~75 nt upstream of *pL* +1. Note these values may change depending on the helical phase occupied by each polymerase. Mid-range topological changes could negatively influence *pL* activity in a number of ways. Changes in topology caused by DNA bending and twisting at the site of elongation could act to reduce promoter recognition, as σ^{70}-holoenzyme, trying to bind *pL*, will be attempting to bend and twist the same DNA in an opposite orientation. During promoter occlusion, this topological affect will have the overall affect of increasing the 'sphere-of-influence' of an elongating RNAP beyond its physical boundaries. This will act to increase the duration of the damaging signal. As previously discussed (Section 6.1.3.), even if this is increased a few seconds, the extent of interference by occlusion will still not account for 5.6 fold interference.

Short range affects. Short range effects are those that occur at the same distance in which physical clashes between polymerases are expected to begin to occur. For the purposes of this study, any possible short-range mechanisms are expected to be equivalent to the mechanisms of occlusion and the sitting duck mechanism but with an added description of the exact molecular details of the interaction(s) involved. I suspect that topological stresses caused by colliding polymerases would become the main factor causing inactivation and possibly dissociation, regardless of wether steric clashes can occur or not. During interference by head-on collisions between elongating polymerases, short range topological stresses as a result of these collisions may also be the cause of inactivating elongation.

In summary although changes in DNA topology caused by *pR* transcription may be the cause of disruptions of *pL* activity, these changes are expected to act very locally and only when associated with elongation polymerases passing over *pL*. This type of interference by polymerase is only an alternative explanation of interference in that it describes a mechanism of *pL* repression that occurs immediately prior to direct physical clashes with elongating RNAP.

6.2 General conclusions about convergent promoters and interference.

Chapter 1 presented evidence for an abundance of convergent transcription in biology. The potential for wide scale gene regulation by transcriptional interference of these promoters is only realised once the general nature of interference is established. From this and previous studies a number of general conclusions about the potential for and extent of interference between other convergent promoters is possible. The mechanism of interference and hence the extent of

interference will depend upon four properties of the convergent promoters- the activity of the aggressive promoter, the kinetic properties of the sensitive promoter, the speed of transcription and the spacing of the promoters.

Properties of the convergent promoters.

The extent of interference has been shown to be proportional to the strength of the aggressive promoter (Elledge and Davis, 1989). Increasing the frequency of initiation at the aggressive promoter is expected to contribute to an increase in interference by either an occlusion, sitting duck or a head-on collisions mechanism. The exact relationship of this dependence remains to be determined. Generally the properties of the sensitive (usually weak) promoter are not expected to significantly alter the extent of occlusion or head-on collisions but are crucially important for a sitting duck mechanism. Sensitive promoters which bind RNAP poorly but clear rapidly are expected to be minimally affected by a sitting duck mechanism of interference. Those promoters that have slow rates for initiation steps after promoter recognition will accumulate 'sitting ducks' that are targets for detrimental collisions with convergent transcription. This study has demonstrated the potential for weak promoters that are rate limited during the transition from open complex to productive transcription, to be sensitive to converging transcription. However, wether weak promoters that form rate limited isomerisation intermediates are also sensitive to collisions is yet to be investigated.

The speed of elongation.

The extent of interference caused by head-on collisions between elongating RNAP molecules is expected to be highly sensitive to the speed of elongation between the convergent promoters. Elongation speed and thus interference mechanisms can be influenced by pause sites and other factors such as ribosomal anti-termination sequences (which increase elongation rates to 80-90 nt/sec) or high ppGpp concentrations (which decrease elongation rates to 20 nt/sec) (Vogel and Jensen, 1994). Placement of a significant pause site is expected to dramatically increase the chance of these collisions. Occlusion will be greatly enhanced by slow elongation over a promoter. Interference by a sitting duck model is expected to be unaffected by elongation rate.

Promoter spacing.

The mechanism of interference will dependent on the distance between convergent promoters.

Very closely spaced promoters, 0-35 bp apart, are those that do not allow co-binding of RNAP *in vitro* and may mediate a mechanism of interference that acts at the level of RNAP binding (Jagura-

Burdzy and Thomas, 1997). However it is difficult to determine whether steric hindrance occurs during *in vivo* steady state transcription. Due to the constant activity of the interfering promoter *in vivo*, promoter bound RNAP complexes may rapidly clear and rarely remain bound long enough to effect steric hindrance. This problem was experienced when examining the ability of RNAP bound at a promoter to act as a repressor during transcription of the divergent promoters from bacteriophage λ, P_R and P_{RM} which have a single base-pair deletion in P_{RM} such that start sites are separated by 81 bp. Using combinations of *in vitro* techniques it has been shown that RNAP activity at P_R interferes with open complex formation at P_{RM} in the absence of NTPs (Woody *et al.*, 1993). Although these interactions are relevant *in vitro*, *in vivo* assays indicate that RNAP bound at P_R does not inhibit P_{RM} activity in the cell (Woody *et al.*, 1993). Apparently, both P_R and P_{RM} promoters clear rapidly enough such that neither is occupied for a significant fraction of time to allow interference. However in bacteriophage 434, where the interpromoter distance between P_R and P_{RM} is even shorter (65 bp), such that the –35 regions of each promoter nearly coincide, a threefold inhibition of P_{RM} activity caused by steric occlusion from P_R bound RNAP was shown to occur both *in vitro* (Xu and Koudelka, 2000) and *in vivo* (Bushman and Ptashne, 1986). This suggests that steric occlusion by promoter bound polymerase can occur *in vivo* but may be dependent on either the precise orientation of promoters or the particular properties of the promoters involved.

Medium spaced promoters, at least 37 bp or greater, would be expected to mediate interference as shown in this study ie. mostly by a sitting duck mechanism of interference. The extent of interference caused by this mechanism is presumed to be independent of promoter spacing. However observations in chapter 3 demonstrated that increasing the spacing between promoters gradually increases the interference between promoters. For *pRpL* promoters spaced 62 bp apart (*pRpL*), 162bp apart (*pR*(+100)*pL*) and 206bp apart (*pR*(tA=)*pL*), the amount of interference was calculated to be 5.6 fold, 7.2 fold and 9.8 fold respectively (Chapter 3). This implies that as promoter spacing is increased alternative mechanisms of interference become significant. For more distantly spaced promoters, the influence of these other mechanisms presumably becomes dominant over that of sitting duck interference. Possible explanations for this spacing dependent mechanism include head-on collisions between elongating RNAP, antisense effects or topological effects.

For distantly spaced promoters, the influence of head-on collisions between elongating RNAP molecules becomes significant. With a wildtype spacing of 62bp between promoters any *pR* initiated polymerase has just over 1 second to collide with RNAP from *pL* before it passes over *pL* +1, assuming polymerase travels at 40-50 nt/s (Vogel and Jensen, 1994). Even if *pR* was as strong

as λP_L, which has been measured as firing at an average rate of once every 4.6 seconds (Liang *et al.*, 1999), then at most only about 1 in 5 of the polymerases fired from *pL* is expected to be involved in a head-on collision. If the travelling time between promoters is increased either by increasing the distance between promoters or the presence of inter-promoter pause sites, then the frequency of these type of collisions will contribute significantly to the mechanism of interference. For example if 186 *pR* and *pL* were placed at a inter-promoter distance of 1kb, polymerase will be expected to take on average 20-25 seconds to elongate over this distance. With *pR* expected to transcribe at an average rate of once every 18 seconds, the chance of elongating complexes from *pL* colliding with at least one head-on polymerase from *pR* will be 100%. This mechanism of interference depends on the outcome of any head-on collisions. This is currently unknown, but is suggested to be the explanation for interference occurring between convergent promoters placed over 3 kb apart (Prescott and Proudfoot, 2002).

Although the antisense transcript was demonstrated here to not be involved in interference, increasing the distance between promoters also increases the region of overlapping complementary transcript which may then be sufficient enough to cause interference. This mechanism is not expected based on the study of interference between convergent promoters placed 1.5 kb apart by Elledge and Davis, (1989). They showed that an *in trans* supply of antisense transcript does not cause interference.

Increasing the spacing between promoters could potentially increase the topological constraint of elongating complexes prior to transcription over the convergent promoter. For closely spaced promoters such as *pRpL*, transcription from *pR* of the 62 bp prior to *pL* may not be a large enough distance to establish a topologically constrained domain of transcription, such that *pR* transcriptionally induced changes in supercoiling and other topological stresses will be minimal prior to *pL*. That is when elongation from *pR* reaches *pL* it may still be elongating around the DNA rather than feeding the DNA through it active site. It is possible that increasing the distance between promoters would increase the chance of more topologically constrained transcription from *pR* occurring prior to transcription over *pL*. Additionally, if the transcript from the aggressive promoter was translated prior to transcription over the sensitive promoter, the loading of ribosomes onto the transcript would also increase topological constrained domain downstream of *pL*.

Future experiments.

The outcome of head-on polymerase collisions could possibly be analysed *in vitro* in real time using surface plasmon resonance (SPR). SPR reports small local changes in refractive index linked

directly to alterations in concentration at a surface. The technique has been used for measuring interactions between DNA immobilised on a surface and RNAP in solution (Muskhelishvili *et al.*, 1997). If the DNA contained sequences for strong face-to-face promoters (such *pRpR*) then by manipulating the inter-promoter sequence the outcome of collisions could be examined as follows. Double stranded DNA templates would be immobilised on the surface of the biosensor chip. RNAP would first be bound to both promoters in the presence of 3 NTPs such that all the templates contained two face-to-face open complexes that had been converted into paused, opposing elongation complexes and which should be resistant to dissociation after washing with buffer. Such binding of RNAP would result in a large increase in response units. The inter-promoter sequence would be designed such that upon addition of the fourth NTP and the omission of a different NTP, the elongation complexes would be able to elongate through each other and pause at some position downstream such that if both polymerases remained, they would now be back-to-back. If both polymerases remained after the collision then no change in response units would be observed (except that created by the growing RNA chains). However, if one polymerase dissociated then the response units would be halved, alternatively should both polymerases dissociate the response units would fall back to that of the template only. To see if any remaining polymerase molecules were still able to transcribe, all four NTPs would be added; transcription should run off the ends of the template, leading to a measure of response units found for template only. The control for these reactions would be to test the response from templates containing deletions in either of the promoters.

6.3 Transcriptional interference in eukaryotic systems

Interference by convergent eukaryotic promoters is already well documented (see Chapter 1). Although transcription is complicated by the presence of histones and a much larger initiation complex, I suggest that the mechanisms of interference involved are principally the same as that discussed here. Studies of convergent promoters in plants and yeast have shown that interference can be relieved by the addition of termination signals between the promoters (Ingelbrecht *et al.*, 1991; Padidam and Cao, 2001; and Prescott and Proudfoot, 2002). This parallels the findings of the present study in *E. coli* and suggests that the mechanism of interference by elongation from the dominant promoter is generic. Based on the conserved nature of RNA polymerisation and the high degree of structual similarity between the prokaryotic and eukaryotic RNAPs (revealed by crystal structures of *T. aquaticus* RNAP core (Zhang *et al.*, 1999) and the yeast *S. cerevisisiae* RNAP II (Cramer *et al.*, 2001; Gnatt *et al.*, 2001)), the outcomes of either prokaryotic or eukaryotic polymerase collisions are expected to be similar. As the distance between promoters in eukaryotic systems are often in the kb range, and the speed of transcription is slow (20-25 nt/s (Ucker and

Yamamoto, 1984)) the prediction is that in these cases interference is mostly due to head-on collisions between elongating polymerases. This is indeed what has been found for convergent genes in budding yeast placed ~3kb apart (Prescott and Proudfoot, 2002). In this example neither a sitting duck model nor occlusion would be possible as most of the convergent elongation was shown to terminate prior to transcription over the opposing promoter. It would be interesting to know whether promoters spaced more closely such that collisions between elongating polymerases become insignificant, are still able to interfere. If so it is suspected that similar rules governing occlusion and sitting duck interference will exist. Sitting duck interference caused by collisions with initiation intermediates waiting to escape and clear the promoter are possible, as it known that activity of eukaryotic promoters can be resticted at the level of promoter clearance, and additionally there is support for the idea that complexes involved in promoter escape are inherently unstable until they synthesize RNA 10 or more nucleotides in length (Dvir, 2002).

6.4 Reinterpretation of developmental switch in light of new interference data

Interference by convergent transcription is a flexible mechanism for the control of gene expression. For the convergent lysogenic and lytic promoters of bacteriophage 186 and P2, evolutionary adjustments in the strength of the lytic promoter, the interpromoter distance and the properties of the lysogenic promoter have created pairs of promoters which reduce the activity of the lysogenic promoter of 186 5.5 fold, but in P2 only 2.2 fold. The need for these differences in interference may be explained by the different strategies used to establish lysogeny. To establish lysogeny, 186 requires sufficient concentrations of an activator, CII, translated from the early lytic transcript, to activate an alternative leftward promoter and give a transitory burst of CI repressor transcription. The repressor in turn shuts off *pR* and hence positively autoregulates *pL* transcription. In a phage infection *pL* activity alone is insufficient for establishing *pL*, presumably because of the high level of transcriptional interference. Establishment of lysogeny in P2 is not dependent on a CII function and therefore must rely on transcription from *pc* alone to establish high enough levels of the lysogenic repressor, C, to prevent lytic development and enter lysogeny. This task would be very difficult if interference with *pc* activity was high. To ensure that lytic development occurs in the absence of significant lysogenic transcription, P2 can not rely on the inhibition of *pc* expression by transcriptional interference and instead must depend on repression by another protein Cox, transcribed from *pe*.

6.5 Transcriptional interference by tandem promoters

Conclusions from this study will also bear on the mechanism of potential transcriptional interference displayed with tandemly arranged promoters. The negative impact of a promoter placed upstream of another promoter and directing transcription in the same orientation has been reported in a number of situations, (Adhya and Gottesman, 1982; and Gafny *et al.*, 1994). The mechanism proposed to be responsible for this interference is promoter occlusion. The calculations of the predicted extent of promoter occlusion caused by convergent transcription should be transferable to tandem transcription. This would predict that only very strong promoters or the presence of pause sites at the interfered promoter would be able to direct interference by an occlusion mechanism. Studies where occlusion has been suggested did involve a very strong interfering promoter, ie λP_L (Adhya and Gottesman, 1982) and ribosomal RNA promoter P1 (Gafny *et al.*, 1994). The sitting duck mechanism of interference would only be feasible if it were shown that the collision of an elongating complex with the rear of an open complex is detrimental to open complex activity. The establishment promoter of 186, pE, is located 278 bp upstream and in tandem to *pL*, and the potential for interference of *pL* activity caused by the activation of pE by CII has been investigated (Neufing *et al.*, 2001). Although activated pE was shown to be less than two fold weaker than *pR*, in the presence of active *pL* leftward transcription (from pE and *pL*) was shown to be additive, indicating a lack of interference. This result would suggest that collisions with the rear of sitting ducks at *pL* are not detrimental.

6.6 Final conclusions.

Transcriptional interference caused by convergent transcription has been reported by many groups in a number of different organisms. The importance of this interference as a method of gene regulation has also been reported, and there exists a great potential for this type of regulation in biology. This study has led to the proposal that for at least one example of convergent transcription, the mechanism of interference is mostly due to a sitting duck mechanism. Although a number of questions about the mechanisms of interference still remain, such as the ability of isomerisation intermediates to act as sitting ducks and the mechanism of interference between widely spaced promoters, this study has substantially progressed the current understanding of interference by convergent transcription. The consequences of this study for the significance of interference in other examples of convergent transcription are that the degree and type of transcriptional interference occurring is likely to dependent on the properties of the promoters involved and the spacing between promoters. Consequently regulation by interference can not simply be predicted based on the occurrence of convergent transcription, but will require measurements of the activities

of each promoter. I anticipate that many future studies examining the regulation of gene expression will observe regulation as a consequence of transcriptional interference, and that the study presented here will assist in these investigations.

CHAPTER SEVEN

MATERIALS AND METHODS

CHAPTER 7

Materials and methods

7.A. Materials.

7.A.1. BACTERIAL STRAINS.

C-2420: *E. coli* strain C-1a, F⁻ made *zai-736*::Tn*10* Δ(*argF-lac*)U169 Tet^S (Julien and Calendar, 1995) used as a C-strain host for *lacZ* reporter assays.

DH5α: F⁻ *end*A1 *hsd*R17 ($r_K^- m_K^+$) *sup*E44 *thi*-1 *rec*A1 *gyr*A (Nal^R) *rel*A1 Δ(*lacZ*YA-*arg*F) U169 *deo*R (ϕ80*dlac*Δ(*lacZ*) M15) (Bethesda Research Laboratories) used for routine cloning.

NK7049: Δ*lac*X74 *gal* OP308 *rps*L (Simons *et al.*, 1987) Lab strain E4300. *E. coli* *lac*⁻ strain and host strain for bacteriophage 186, used as a host for *lacZ* reporter assays.

MC1061.5: F⁻ *ara*D139 Δ(*ara-leu*)7696 Δ(*lac*)χ74 *gal*E15 *gal*K *hsd*R2 ($r_K^- m_K^+$) *mcr*B1 *rps*L (Str^R) (Koop *et al.*, 1987) *recA*$^+$ strain used for routine cloning and as a host for *lacZ* reporter plasmids and phage.

7.A.2. Bacteriophage.

λRS45 : Lambda phage derivative used to create single copy *lacZ* fusions. (Simons *et al.*, 1987)(see fig 2.2 and fig 7.1)

λRS45ΔYA : λRS45 with *lacY* and *LacA* sequences deleted described in (Dodd *et al.*, 2001). Lambda phage derivative used to create single copy *lacZ* fusions (see fig 2.2 and fig 7.1).

7.A.3. Primers

All primers used in this study are detailed in table 7.1.

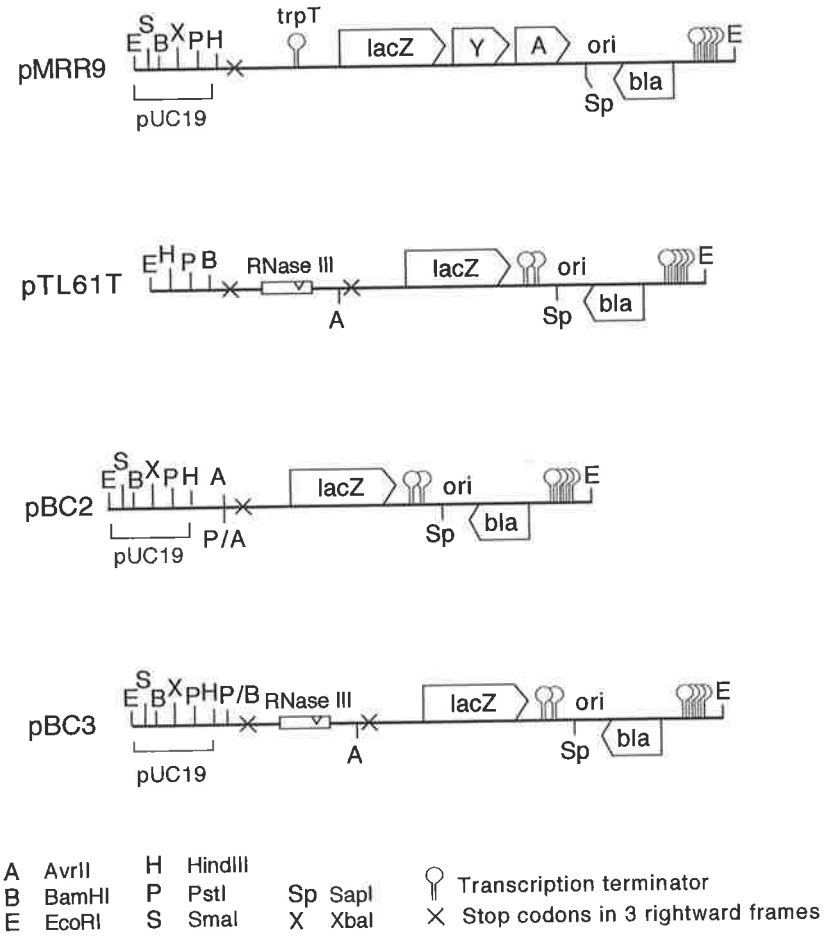


Fig 7.1. Diagram of lacZ reporter vectors.

Adapted from Dodd and Egan, 2002.

Table 7.1

Primer #	Sequence	Use
	Chapter 2 (and general use)	
181 r	186: 22980 (pR -81) GGG GTA CCT CTA GAC GTT GCT CCA TCC TAA AGA Kpn I Xba I Mae II	PCR of p*R* inserts containing MaeII region of 186 for construction of pBC2 and pBC1 plasmids.
182 l	186: 23190 (pL -67) GGG GTA CCT CTA GAG TAA CGA TAG GTG CAG GCA C Kpn I Xba I Mae III	PCR of *pL* inserts containing MaeIII region of 186 for construction of pBC2 and pBC1 plasmids.
236 r	186: 23055 (pL +68) GGG GTA CCT CTA GAT TGG CTA AAC CCA CGC AAT T Kpn I Xba I	PCR of 186:*pL* inserts for construction of pBC2 plasmids.
237 l	186: 23129 (pR +69) GGG GTA CCT CTA GAC CCT ATT AGC CAA AGT TTG C Kpn I Xba I	PCR of 186:*pR* inserts for construction of pBC2 plasmids.
256 r **257 l**	pR^{-} −10 5' CTC AAT TGG GAG **ATC GAT GTT** GGC TAA ACC C 3' 3' GAG TTA ACC CTC **TAG CTA CAA** CCG ATT TGG G 5' Cla I	Site directed mutagenesis of 186:*pR* −10 region. Changes are indicated in bold.
57 r	AGT TCC CAA GCT TGC ATG CC	Sequencing of pBC2, pMRR9 plasmids, anneals upstream of *lacZ* and MCS.
USP	AGT TCC CAA GCT TGC ATG CC	Sequencing of pBC2, pMRR9 clones and pBSSK^{+} subclones, anneals within *lacZ*. Also used to make the loading control.
154 λATTP	TTT AAT ATA TTG ATA TTT ATA TCA TTT TAC GTT TCT CGT TC	PCR detection of single or multiple λ lysogens anneals left of λ *attP*.
155 λATTB	GAG GTA CCA GCG CGG TTT GAT C	PCR detection of single or multiple λ lysogens anneals left of *E.coli attB*.
156 λINT	ACT CGT CGC GAA CCG CTT TC	PCR detection of single or multiple λ lysogens anneals left of within λ *int* gene
SK	CGC TCT AGA ACT AGT GGA TC	Sequencing of pBSSK^{+} subclones, and also used to make the loading control.

Primer #	Sequence	Use
	Chapter 3	
308 l	186: 23088 (pR +28) CTA GCT AGC GCA ACA CTT GCC ATC AAT TGC Nhe I	PCR of 186:*pR* inserts for construction of pBC2 plasmids. Also used in chapter 5.
309 r	186: 23096 (pL +27) TGG GCT AGC GAG TCA AAT CAA TTG CAA AC Nhe I	PCR of 186:*pL* inserts for construction of pBC2 plasmids. Also used in chapter 5.
236 r	186: 23055 (pL +68) GGG GTA CCT CTA GAT TGG CTA AAC CCA CGC AAT T Kpn I Xba I	PCR of 186:*pL* inserts for construction of pBC2 plasmids.
237 l	186: 23129 (pR +69) GGG GTA CCT CTA GAC CCT ATT AGC CAA AGT TTG C Kpn I Xba I	PCR of 186:*pR* inserts for construction of pBC2 plasmids.
318 r	GAT CCA GCA ATC AGA TAC CCA... Bam HI ...GCC CGC CTA ATG AGC GGG CTT TTT TTT CCC Sma I Stem-loop U-tract	Double stranded oligo used to insert tA^+ sequence.
320 l	Complementary to 318	
319 r	GAT CCA GCA ATC AGA TAC CCA... Bam HI ...GCC CGC CTA ATG AGC GGG CTC GCG TGT CCC Sma I tA^- mutation	Double stranded oligo used to insert tA^- sequence.
321 l	Complementary to 319	
401 r	Bam HI C CGG GAT CCA GCA ATC AGA TAC CCA GCC CGC CTA ATG... $tA^=$ mutation anti-# 57 ..ACT CGA GTC GCG TGT CCC GGG ATC CCC AAT TCC TGG CA Xho I Sma I	Single stranded oligo used to insert $tA^=$ sequence.
	Chapter 4	
244 l	pTL61T: 4905 (after $(T1T2)_2$) CCA CAT GTC ATA TGG ACC CAA CGC TGC CCG ACT Nde I	PCR of $(T1T2)_2$ inserts for construction of pBC1.
245 r	pTL61T: 4025 (50bp from start of T1) GGA ATT CTG AAA CGC CGT AGC GCC GAT Eco RI	PCR of $(T1T2)_2$ inserts for construction of pBC1.

Primer #	Sequence	Use
	pTL61T: 4041 (30 bp from start of T1)	
247 r	CGC AAG CTT GAT GGT AGT GTG GGG TCT C Hind III	PCR of $(T1T2)_2$ inserts for construction of pBC1.
	pTL61T: 4924 (after $(T1T2)_2$)	
248 l	TCC GAT ATC ATG CGC ACC CGT GGC C Eco RV	PCR of $(T1T2)_2$ inserts for construction of pBC1.
250	AGT TCC CAA GCT TGC ATG CC	Sequencing of pBC1 plasmids anneals near HindIII site.
	Chapter 5	
	P2: 25832 (pe -85)	
238 r	GGG GTA CCT CTA GAC TGC TCA AAT ACT CTG ATT TTC Kpn I Xba I	PCR of P2:*pe* inserts for construction of pBC2 and pBC1 plasmids.
	P2: 26025 (pc -69)	
239 l	GGG GTA CCT CTA GAG TGT TAC TTG CTT GCT CA Kpn I Xba I	PCR of P2:*pc* inserts for construction of pBC2 and pBC1 plasmids.
	P2: 25870 pe⁻ –35 P2: 25904	
251 r	5'CTC GCT TAT CGT GTT TCA TAT GGT GTT TAG ATC TC 3'	Site directed mutagenesis of P2:*pe* –35 region. Changes are indicated in bold.
252 l	3'GAG CGA ATA GCA CAA AGT ATA CCA CAA ATC TAG AG 5' Nde I Bgl II	
	P2: 25980 pc⁻ –10 P2: 25946	
253 l	5'GCG TTT AAT GTC TTG TAA GGC CTT TTA GTG CCC AC 3'	Site directed mutagenesis of P2:*pc* –10 region. Changes are indicated in bold.
254 r	3'CGC AAA TTA CAG AAC ATT CCG GAA AAT CAC GGG TG 5' Hae III	
	P2: 25729 (pe –187)	
316 r	GCT CTA GAC TGC AGG ATG TTC ATC ATG Xba I Pst I	PCR of P2:*pe* inserts for construction of pBC2 plasmids.
	P2: 25894 pe⁻ –10	
317 r	GTT TAG ATC TCA AAC GTA TTT AGT TTA GAT GTA G Bgl II	Mutagenesis of P2:*pe* –10 region by PCR. Changes are indicated in bold.
	P2: 25906	
336 l	TTG AGA TCT AAA CAC CAT GAT TCA CAC GAT AAG CGA G Bgl II pe⁻(new) –35	Mutagenesis of P2:*pe* –35 region by PCR. Changes are indicated in bold.
	P2: 25927 (pc +35 from -10)	
306 l	TGG GCT AGC GAT TGT TTA GTG CTT GGA TG Nhe I	PCR of P2:*pc* inserts for construction of pBC2 plasmids.
	P2: 25948 (pe +36 from -10)	
307 r	TGG GCT AGC ACA TCC AAG CAC TAA ACA ATC Nhe I	PCR of P2:*pe* inserts for construction of pBC2 plasmids.

r and **l** refers to the strand of 186, P2 or plasmid sequence the primer belongs to, based on the coordinates given in the GeneBank database. r is the rightward strand and l is the leftward strand. All primers were constructed by GeneWorks (Australia).

7.A.4. Plasmids.

Table 7.2 lists the plasmids used in this study in order of their appearance in each chapter. The plasmids not constructed in this study are referenced in the description, those not referenced were constructed in this work. Pictorial details of some of the clones can be gathered by referring to the relevant figures in each chapter used to describe the results and experiment. The personal (BC) glycerol stock numbers of the plasmids constructed here are given. Plasmids used to assay promoter activity were transferred to a single chromosomal copy by recombination with λ and lysogenisation of a particular host strain, and the resulting lysogenic strain assayed. The glycerol stock numbers of the lysogens used for assaying the relevant clones are given in the form BC###-Xλ or BC###-XλΔ where: 'BC###' indicates the glycerol stock number of the lysogen made from the plasmid described in that row of the table, 'X' indicates the *E. coli* strain used as a host which is either 'C' (C-2420), 'N' (NK7049) or 'M' (MC1061.5), 'λ' indicates that λRS45 was used for recombination and 'λΔ' indicates that λRS45ΔYA was used.

All plasmids in this study that were constructed by inserting DNA generated by PCR have been sequenced over the insert. Where the construction of plasmids allows for multiple orientations, correct orientations have been checked by colony PCR. The junctions of plasmids constructed by ligation of restriction fragments have been checked by either sequencing or restriction digest.

7.A.5. REAGENTS.

7.A.5.1. Enzymes.

E. coli DNA polymerase I (Klenow fragment): GeneWorks. Adelaide, Australia.

E. coli RNAP holoenzyme (sigma saturated): Epicentre technologies.

Lysozyme: Sigma Chemical Co.

Restriction Endonucleases: New England Biolabs, Boehringer Man*NheI*m or Pharmacia.

RNase A: Sigma Chemical Co. 10mg/ml stock solution, heated at 95° C for 20min to inactivate DNases.

RNase inhibitor (SUPERase-In): Ambion

Table 7.2.

Plasmid	Description	Glycerol/ Lysogen Number
	CHAPTER TWO	
	Measuring transcriptional interference between 186 *pR* and *pL* in vivo.	
pBS SK+	pBleuscriptSK+ used for general cloning and subcloning of PCR fragments. (Stratagene)	
pUC19	General cloning vector. (Yanisch-Perron et al., 1985)	
pACYC(186)CI	pACYC184 expressing the 186 CI gene. The large 819bp HaeIII to HaeIII fragment of pET3a-CI (Shearwin and Egan, 1996) was inserted into the NruI site of pACYC184. CI expression is sufficient to give immunity to 186. Strains carrying this plasmid were used as hosts strains when constructing plasmids that contain active *pR* sequences. (K. Shearwin, unpublished)	
pMRR3	Contains the *Xho I-Bgl II* (629-4244) fragment from 186 inserted into the *Sal* I-*Bam* HI sites of pUC19. Used as the PCR template to obtain *pRpL* fragments (Dodd et al., 1993).	
pMRR7	PR^- derivative of pMRR3 used as a PCR template to obtain pR^-pL fragment. (Reed, 1994)	
pMRR9	*LacZ* transcriptional fusion vector (see fig 7.1). (Reed, 1994)	
pPN467	Contains the SnaBI to BspMI region of 186 with pL^- and *apl* Δ11 changes, cloned into the *Eco* RI and *Sma* I sites of pMRR9. Used as PCR template for $pRpL^-$ fragment. (Neufing, 1997)	
pTL61T	*LacZ* transcriptional fusion vector (see fig 7.1). (Linn and Pierre, 1990)	
pBS-HS-pL^+pR^-	Plasmid containing the HincII-SnaBI (22258-23532) pL^+pR^- fragment from pMRR7 inserted into the EcoRI-PstI site of pBluescriptSK+. Used as a template for quickchange mutagenesis of the *pR* –10 region. (I.B. Dodd, unpublished)	
pBS-HS-$pL^+pR^=$	pBS-HS-pL^+pR^- derivative containing the mutation in *pR* –10 caused by quickchange mutagenesis using primers #256 and #257.	250

pBC2	*LacZ* transcriptional fusion vector (fig 7.1) made by (i) removing the RNase III cleavage site in pTL61T by cutting with *PstI* and *AvrII*, blunt ending by treatment with T4 DNA polymerase and rejoining (recreates *AvrII*) site; and (ii) replacing the *SapI-HindIII* fragment containing the *bla* gene and MCS with the equivalent fragment from pMRR9.	270 BC321-Nλ, BC322-Nλ, BC340-MλΔ, BC341-MλΔ, BC323-Cλ, BC324-Cλ
pMRR9-*pRpL*	Plasmid used to assay promoter activity. DNA fragments from the 186 *Mae*II.22980 to *Mae*III.23183 sites (MM), containing the *pL* –65 to +143 region, were prepared by PCR from the $pR^{+}pL^{+}$ template pMRR3 using primers #181 and #182 each bearing flanking *Xba*I sites. The DNA was cut with *Xba*I and inserted into the *Xba*I site of pMRR9 in an orientation such that *pL* directs transcription of the *lacZ* gene.	195 BC223-Nλ, BC224-Nλ
pMRR9-*pRpL*	As for pMRR9-*pRpL* except that the insert is orientated such that *pR* directs transcription of the *lacZ* gene.	191 BC225-Nλ, BC226-Nλ
pMRR9-$pR^{-}pL$	As for pMRR9-*pRpL* but using a $pR^{-}pL$ PCR insert prepared from the template pMRR7. The insert is orientated such that *pR* directs transcription of the *lacZ* gene.	197 BC221-Nλ BC222-Nλ
pBC2-*pRpL*	As for pMRR9-*pRpL* except that the PCR fragment was inserted into the *Xba*I site of pBC2 in an orientation such that *pL* directs transcription of the *lacZ* gene.	283 BC313-Nλ, BC314-Nλ,
pBC2-*pRpL*	As for pBC2-*pRpL* except that the insert is oriented such that *pR* directs transcription of the *lacZ* gene.	296 BC327-Nλ, BC328-Nλ, BC361-MλΔ, BC362-MλΔ
pBC2-$pR^{=}pL$	As for pBC2-*pRpL* but using a $pR^{=}pL$ PCR insert prepared from the template pBS-HS-$pL^{+}pR^{=}$. The insert is orientated such that *pL* directs transcription of the *lacZ* gene.	273 BC319-Nλ, BC317-MλΔ, BC318-MλΔ
pBC2-*pRpL*$^{-}$	As for pBC2-*pRpL* but using a *pRpL*$^{-}$ PCR insert prepared from the template pPN467. The insert is orientated such that *pR* directs transcription of the *lacZ* gene.	284 BC325-MλΔ, BC326-MλΔ
pBC2-$pR^{=}pL$	As for pBC2-$pR^{=}pL$ but with the insert orientated such that *pR* directs transcription of the *lacZ* gene. Used in the construction of pBC2-$pR^{=}pL^{-}$.	271

pBC2-*pR*$^{=}$*pL*$^{-}$	Plasmid used to assay *pR*$^{=}$ activity. The insert was constructed by a two factor PCR ligation. The *pR*$^{=}$ fragment was prepared by PCR from the template pBC2-*pR*$^{=}$*pL* using primers #57 and #237, then digested with KpnI and PleI (which cuts the DNA once between *pR* and *pL*). The *pL*$^{-}$ fragment was prepared by PCR from the template pBC2-*pRpL*$^{-}$ using primers USP and #236, then digested with PleI and HindIII. The two promoter fragments were ligated together and then inserted into the KpnI/HindIII site of pBC2.	275 BC337-MλΔ, BC338-MλΔ
pBC2-*pR*$^{=}$*pL*$^{-}$	Plasmid used to assay *pL*$^{-}$ activity. The *pR*$^{=}$*pL*$^{-}$ fragment was prepared by PCR from the template pBC2-*pR*$^{=}$*pL* using primers #181 and #182, then digested with XbaI and inserted into the XbaI site of pBC2 such that *pL*$^{-}$ directs transcription of *lacZ*.	277 BC339-MλΔ

CHAPTER THREE

What action of *pR* transcription causes interference?

pBC3	*LacZ* transcriptional fusion vector similar to pBC2 but containing the RNase III cleavage site (fig 7.1). Made by (i) cutting pTL61T with BamHI and PstI, endfilled with Klenow fragment and religated to remove some restriction sites and (ii) replacing the *SapI-HindIII* fragment containing the *bla* gene and the MCS from the modified pTL61T with the equivalent fragment from pMRR9.	658
pBC3-*pRpL*	As for pBC2-*pRpL* but the *pRpL* fragment was instead inserted into the XbaI site of pBC3 such that *pL* directs transcription of the *lacZ* gene.	669 BC688-MλΔ, BC689-MλΔ
pBC3-*pR*$^{=}$*pL*	As for pBC2-*pR*$^{=}$*pL* but the *pR*$^{=}$*pL* PCR insert was inserted into the XbaI site of pBC3, such that *pL* directs transcription of the *lacZ* gene.	671 BC690-MλΔ, BC691-MλΔ
pBC3-*pRpL*	As for pBC3-*pRpL* except that the insert is oriented such that *pR* directs transcription of the *lacZ* gene.	673 BC682-MλΔ, BC683-MλΔ
pBC2-*pR*(+100)*pL*	Used to assay 186 promoter activity with increased spacing. Constructed in two parts, (i) PCR fragments of *pR*+ were prepared from the template pMRR3 using primers #181 and #237, then digested with XbaI and inserted into the XbaI site of pBC2 such that *pR* transcribes away from *lacZ*. (ii) PCR fragments of *pL*+ were prepared from the template pMRR3 using primers #182 and #236, then digested with KpnI and inserted into the KpnI site of the plasmid made in step (i), such that *pL* transcribes *lacZ*.	377 BC398-MλΔ, BC475-MλΔ
pBC2-*pR*$^{=}$(+100)*pL*	As for pBC2-*pR*(+100)*pL* except in step (i) PCR fragments of *pR*$^{=}$ DNA was prepared from the template pBC2-*pR*$^{=}$*pL*	366 BC373-MλΔ, BC374-MλΔ

pBC2-*pR*(divMM)*pL*	Plasmid to assay divergent promoter activity. Fragments of MaeII to MaeIII (MM) *pRpL* promoter DNA were prepared by PCR from pBC2-*pRpL*$^{-}$ template using primers #181 and #182 bearing Kpn I sites. The DNA was cut with Kpn I and inserted into the Kpn I site of pBC2-*pR*$^{-}$*pL* such that *pR* directs transcription away from *lacZ* and is divergent to active *pL*, to produce pBC2-*pL*$^{-}$*pR*$^{+}$(divMM)*pR*$^{=}$*pL*$^{+}$.	393 BC434-MλΔ, BC435-MλΔ
pBC2-*pR*$^{-}$(divMM)*pL*	As for pBC2-*pL*$^{-}$*pR*$^{+}$(divMM)*pR*$^{=}$*pL*$^{+}$ except that the PCR fragment was prepared from pBC2-*pR*$^{=}$*pL*$^{-}$.	369 BC375-MλΔ, BC376-MλΔ
pBC2-*pR(div+35)pL*	DNA fragments of divergent *pR* and *pL* which include only *pR*+28 and *pL*+27 were prepared by PCR from pBC2-*pR*(divMM)*pL* and pBC2-*pR*$^{-}$(divMM)*pL* templates using primers #308 and #309 bearing *Nhe I* sites. The DNA was cut with *Nhe I* and inserted into the *Xba I* site of pBC2 such that *pL* transcribes *lacZ*	446 BC481-MλΔ, BC482-MλΔ
pBC2-*pR*(div+35)*pL*	As for pBC2-*pR(div+35)pL*, except that the insert is orientated such that *pR* transcribes *lacZ*.	448 BC483-MλΔ, BC484-MλΔ
pBC2-*pR*$^{-}$(div+35)*pL*	As for pBC2-*pR(div+35)pL*, except that the PCR insert was prepared from pBC2-*pR*$^{-}$(divMM)*pL* templates.	449 BC485-MλΔ, BC486-MλΔ
pBC2-*pR*(tA^{+})*pL*	Used to assay the effect of termination on interference. Clones containing active tA were constructed by using double stranded oligos (oligo #318 with its complementary oligo #320), containing the tA^{+} sequence, a *Bam HI* sticky end upstream of the terminator and a *Sma I* blunt end downstream. tA^{+} oligos were inserted into the *Sma I* and *Bam HI* sites of pBC2-*pR*(+100)*pL*, such that termination of transcription occurs from the direction of *pR*,	444 BC509-MλΔ, BC510-MλΔ
pBC2-*pR*$^{=}$(tA^{+})*pL*	As for pBC2-*pR*(tA^{+})*pL* except that the oligos were inserted into pBC2-*pR*$^{=}$(+100)*pL*.	466 BC513-MλΔ, BC514-MλΔ,
pBC2-*pR*(tA^{-})*pL*	As for pBC2-*pR*(tA^{+})*pL* except that the double stranded oligos used were oligo #319 with its complementary oligo #321 which contain the tA^{+} sequence with mutation in the U-tract, a *Bam HI* sticky end upstream of the terminator and a *Sma I* blunt end downstream. tA^{-} oligos were inserted into the *Sma I* and *Bam HI* sites of pBC2-*pR*$^{+}$(+100)*pL*$^{+}$, in a similar orientation as pBC2-*pR*(tA^{+})*pL*.	472 BC511-MλΔ, BC512-MλΔ
pBC2-*pR*$^{=}$(tA^{-})*pL*	As for pBC2-*pR*(tA^{-})*pL* except that the oligos were inserted into pBC2-*pR*$^{=}$(+100)*pL*.	469.1 BC545-MλΔ, BC546-MλΔ

pBC2-$pR^{+}(tA^{=})\underline{pL}^{+}$	Clones containing completely inactive tA were constructed from a single oligo (#401) containing $tA^{=}$ sequence (with a double mutation of the terminator stem loop and the U-tract to ensure that no termination or pausing occurred) and sites for Bam HI and SmaI cleavage. The oligo was made double stranded by primer extension of a short primer (#57) that was complementary to the 3' end, then digested with BamHI/SmaI. tA+ and inserted into the *Sma I* and *Bam HI* sites of pBC2-*pR*(+100)*pL*, such that termination of transcription occurs from the direction of *pR*,	694 BC699-MλΔ, BC700-MλΔ
pBC2-$pR^{=}(tA^{=})\underline{pL}^{+}$	As for pBC2-*pR*($tA^{=}$)*pL* except that the double stranded oligo was inserted into pBC2-$pR^{=}$(+100)*pL*.	696 BC701-MλΔ, BC702-MλΔ
pBC3-*pR*(tA^{+})*pL*	Clones which assay *pR* activity were made by subcloning into the EcoRV site of pBluescript KS^{+} the blunt-ended PCR product generated by using primers #181 and #182 and templates pBC2-*pR*(tA^{+})*pL*. The EcoRI//HindIII promoter fragment of the resulting subclone was then inserted into the *Eco RI* and *Hind III* sites of pBC3 to produce pBC2-*pR*(tA^{+})*pL*.	676 BC685-MλΔ, BC698-MλΔ
pBC3-*pR*(tA^{-})*pL*.	As for pBC3-*pR*(tA^{+})*pL* except that the template for the PCR product subcloning was pBC2-*pR*(tA^{-})*pL*.	677 BC686-MλΔ, BC687-MλΔ
pBC3-*pR*($tA^{=}$)*pL*.	As for pBC3-*pR*(tA^{+})*pL* except that the template for the PCR product subcloning was pBC2-*pR*(tA^{-})*pL*.	679 BC703-MλΔ, BC705-MλΔ

CHAPTER FOUR

Nature of the interference caused by the passage of RNAP across *pL*.

pBC1	In vitro transcription vector constructed in four steps see fig 7.2.	144
pBC1-*pRpL*	Template used in in vitro transcription assays. DNA fragments from the 186 *Mae*II.22980 to *Mae*III.23183 sites (MM), containing the *pL* –65 to +143 region, used for previous pBC2 clones, were prepared by PCR from the $pR^{+}pL^{+}$ template pMRR3 using primers #181 and #182 bearing flanking *Xba*I sites. The DNA was cut with *Xba*I and inserted into the *Xba*I site of pBC1 in an orientation such that *pR* transcribes over the HindIII restriction site of pBC1.	116
pBC1-$pR^{-}pL$	As for pBC1-*pRpL* but using a PCR insert prepared from the $pR^{-}pL^{+}$ template pMRR7.	161
pBC1-$pRpL^{-}$	As for pBC1-*pRpL* but using a PCR insert prepared from the $pR^{+}pL^{-}$ template pPN467.	162
pBC1-$pR^{=}pL$	As for pBC1-*pRpL* but using a PCR insert prepared from the $pR^{=}pL^{+}$ template pBS-HS-$pL^{+}pR^{=}$.	638

CHAPTER FIVE

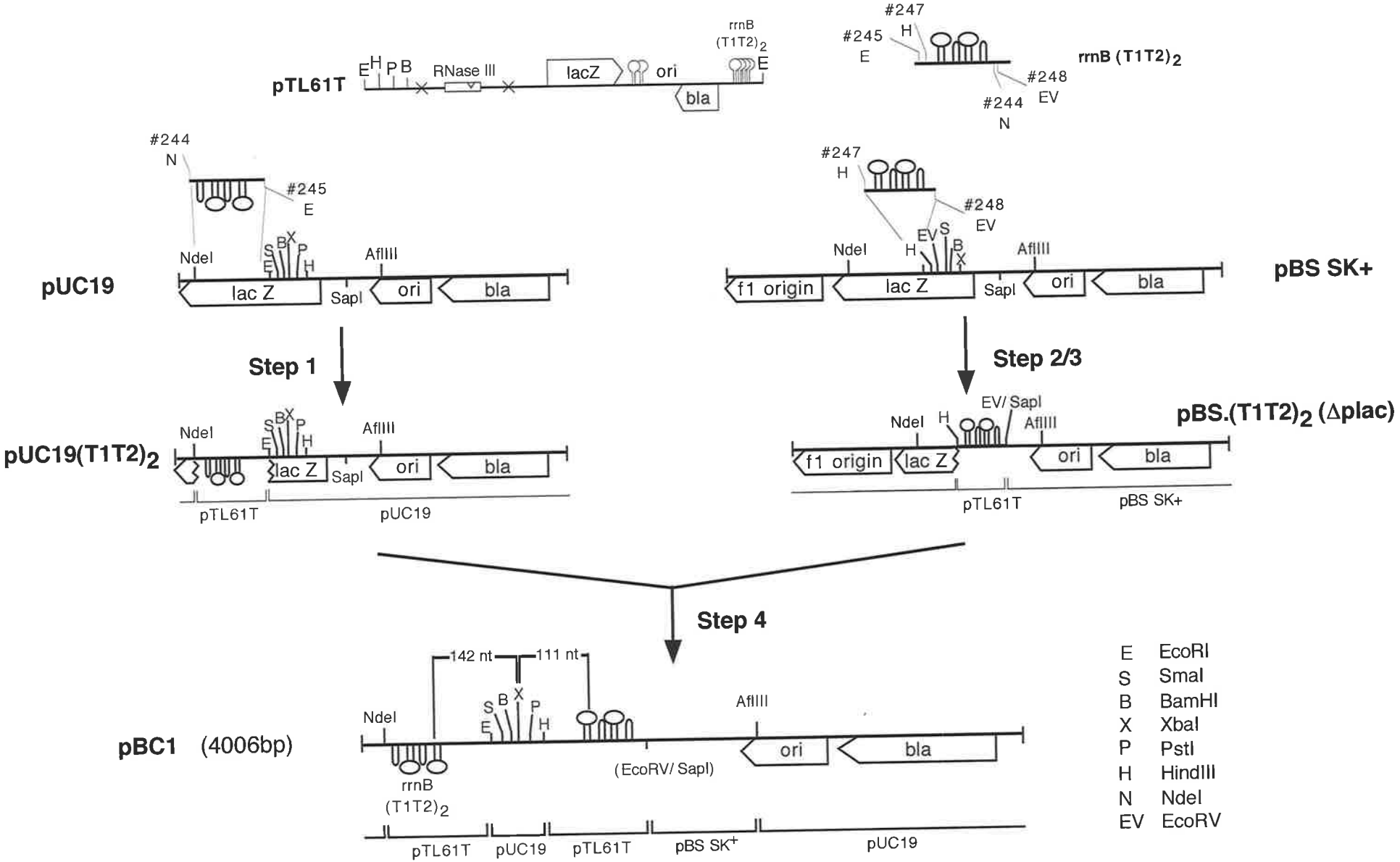
Convergent promoter interference in bacteriophage P2 supports the sitting duck model.

pDBR14	Contains the *SnaBI-DsaI* (24351-26362) fragment from P2+ inserted into the *SmaI* site of pBS SK+. Used as the PCR template for P2 fragments and as the template for mutagenesis of *pc* and *pe*. (D. Reynolds, unpublished)	
pDBR14-*pe*$^{-}$	pDBR14 derivative containing the mutation in *pe* –35 caused by quickchange mutagenesis using primers #251 and #252.	126
pDBR14-*pc*$^{-}$	pDBR14 derivative containing the mutation in *pc* –10 caused by quickchange mutagenesis using primers #253 and #254.	129
pBC2-*pepc*	Plasmid used to assay P2 promoter activity. The *pepc* fragment was prepared by PCR from the template pDBR14 using primers #238 and #239 (bearing XbaI sites), then digested with XbaI and inserted into the XbaI site of pBC2 such that *pc* directs transcription of *lacZ*.	300 BC344-CλΔ, BC345-CλΔ, BC342-MλΔ, BC343-MλΔ
pBC2-*pepc*	As for pBC2-*pepc* but with the insert oriented such that *pe* directs transcription of *lacZ*.	302 BC331-CλΔ, BC332-CλΔ, BC329-MλΔ, BC330-MλΔ
pBC2-*pe*$^{-}$*pc*	As for pBC2-*pepc* but using a *pe*$^{-}$*pc* fragment prepared from the template pDBR14-*pe*$^{-}$.	285 BC306-MλΔ, BC307-MλΔ
pBC2-*pepc*$^{-}$	As for pBC2-*pepc* but using a *pe*$^{-}$*pc* fragment prepared from the template pDBR14-*pc*$^{-}$.	304 BC333-MλΔ
pBC2-(P)*pepc*	As for pBC2-*pepc* but using a larger *pepc* fragment prepared from the template pDBR14 using primers #316 (which anneals next to the PstI site in P2) and #239.	456 BC501-MλΔ, BC502-MλΔ
pBC2-(P)*pepc*	As for pBC2-(P)*pepc* but with the insert oriented such that *pe* directs transcription of *lacZ*.	458 BC503-MλΔ, BC504-MλΔ

Figure 7.2

Construction of *in vitro* transcription vector pBC1.

In vitro transcription vector pBC1 was constructed in four steps. Step 1: the tandem repeat of *rrnB T1* and *T2* terminator region $(T1T2)_2$ present in pTL61T (bp 4025 to 4904 on pTL61T) was amplified by PCR using *Pfu* polymerase and primers #244 (bearing a *NdeI* site) and #245 (bearing an *Eco RI* site) and then inserted into the *NdeI* and *Eco RI* sites of the cloning vector pUC19 to generate pUC19$(T1T2)_2$. Step 2: the $(T1T2)_2$ present in pTL61T (bp 4041 to 4924 on pTL61T) was amplified using primers #247 (bearing a *Hind III* site) and #248 (bearing an *Eco RV* site) and then inserted into the *Eco RV* and *Hind III* sites pBluescriptSK+ to generate pBS.$(T1T2)_2$. Step 3: this vector was then digested with Eco RV and Sap I, endfilled with Klenow fragment and religated to remove the *lacZ* promoter. Step 4: pBC1 was finally constructed by ligating the small *Hind III/ Afl III* terminator-containing fragment from the modified pBS.$(T1T2)_2$ vector, with the large terminator-containing *Hind III/ Afl III* fragment of pUC19$(T1T2)_2$.


pTL61T
E H P B
RNase III
lacZ
ori
bla
rrnB (T1T2)2
E
#247
H
#245
E
#248
EV
#244
N
rrnB (T1T2)2
pUC19
#244
N
#245
E
NdeI
S E
B X
P
H
AflIII
lac Z
SapI
ori
bla
pBS SK+
#247
H
#248
EV
NdeI
H
EV
S
B
X
AflIII
f1 origin
lac Z
SapI
ori
bla
Step 1
Step 2/3
pUC19(T1T2)2
NdeI
S E
B X
P
H
AflIII
lac Z
SapI
ori
bla
pTL61T
pUC19
EV/ SapI
NdeI
H
AflIII
pBS.(T1T2)2 (Δplac)
f1 origin
lac Z
ori
bla
pTL61T
pBS SK+
Step 4
142 nt
111 nt
NdeI
S E
B X
P
H
AflIII
(EcoRV/ SapI)
ori
bla
rrnB
(T1T2)2
pTL61T
pUC19
pTL61T
pBS SK+
pUC19
pBC1 (4006bp)
E EcoRI
S SmaI
B BamHI
X XbaI
P PstI
H HindIII
N NdeI
EV EcoRV

Shrimp Alkaline Phosphatase (SAP): USB

T4 DNA ligase: GeneWorks or Promega.

T4 DNA polymerase: New England Biolabs.

T4 Polynucleotide kinase: GeneWorks.

Taq DNA polymerase: Fisher Biotech International (Australia)

Pfu Turbo DNA polymerase: Stratagene.

7.A.5.2. Chemicals

a. Radiochemicals.

Radiochemicals [α-^{32}P]-dCTP, [α-^{32}P]-rUTP and [γ-^{32}P]-dATP of specific activity 3000 Ci/mmol (radioactive concentrations of 10 mCi/ml) were purchased from GeneWorks (Australia).

b. General chemicals.

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

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

Acetic acid: B.D.H. Labs., Australia.

Polyacrylamide solutions: National Diagnostics.

Agarose: Sigma Chemical Co.

Ammonium acetate: B.D.H. Labs., Australia.

Ammonium persulphate (APS): May and Baker Ltd. Stock solutions at 25% (w/v) in H_2O were prepared fresh on the day of use.

Ampicillin (sodium salt): Sigma Chemical Co. Stock solutions (50-100 mg/ml in H_2O) were millipore filtered and stored at –20°C.

β-Mercaptoethanol: Sigma Chemical Co.

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

Boric acid: B.D.H. Labs., Australia.

Bovine serum albumin (BSA): Sigma Chemical Co. Kept as a 10 mg/ml solution in H_2O at –20°C.

Bromophenol blue: B.D.H. Labs., Australia.

Cesium chloride (CsCl): Bethesda Research Labs.

Calcium chloride ($CaCl_2$): Sigma Chemical Co.

Chloramphenicol: Sigma Chemical Co. Stock solutions (30 mg/ml in ethanol) were stored at –20°C.

Chloroform: B.D.H. Labs., Australia.

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

Dithiothreitol (DTT): Sigma Chemical Co. Stored as a 1 M solution in H_2O at -20°C.

Ethidium bromide: Sigma Chemical Co. Stored as a 10 mg/ml solution in H_2O in the dark at 4°C.

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

Formamide: B.D.H. Labs., Australia. De-ionized and stored in the dark at –20°C.

Gelatin: Sigma Chemical Co.

Glucose: Ajax.

Glycerol: B.D.H. Labs., Australia.

Heparin sodium salt: Sigma Chemical Company. Stored as a 100 mg/ml solution in H_2O at –20°C.

Hydrochloric acid (HCl): B.D.H. Labs., Australia.

Isopropanol (IPA): May and Baker Ltd.

Magnesium acetate: B.D.H. Labs., Australia

Magnesium chloride: Ajax.

Magnesium sulphate: B.D.H. Labs., Australia

Methanol: B.D.H. Labs., Australia.

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

O-nitrophenyl-β-D-galactopyranoside (ONPG): Diagnostic Chemicals Ltd. Used as a freshly made 4 mg/ml solution in TZ8 buffer.

Polymyxin B sulphate: Sigma Chemical Co. Stored at 20 mg/ml in H_2O at –20°C.

Potassium acetate: B.D.H. Labs., Australia

Potassium chloride: B.D.H. Labs., Australia

Potassium glutamate: B.D.H. Labs., Australia

Ribonucleotides (rNTPs): Promega

Sodium acetate: B.D.H. Labs., Australia.

Sodium chloride: B.D.H. Labs., Australia

Sodium dodecyl sulphate (SDS): Sigma Chemical Co.

Sodium hydroxide: Ajax.

Sucrose: Ajax

Tris acetate: B.D.H. Labs., Australia.

Xylene cyanol: Sigma Chemical Co.

7.A.6. MEDIA AND BUFFERS.

7.A.6.1. Growth Media.

a. Liquid media.

Luria broth (LB)

1% Bacto-tryptone, 0.5% yeast extract, 1% NaCl, pH 7.0.

SOC medium (SOC)

2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl and H_2O to 1000 ml. This solution was autoclaved, cooled and 10 ml of 1 M $MgSO_4$, 10 ml of 1 M $MgCl_2$ and 20 ml of 1 M glucose added.

YENB medium (YENB)

0.75% Bacto yeast extract, 0.8% Bacto Nutrient Broth, pH7.0.

All media were prepared in glass distilled H_2O and were sterilised by autoclaving for 25 min at 120°C and 120 kPa.

Antibiotics were added to LB at the following concentrations: ampicillin at 100 μg/ml and chloramphenicol at 30 μg/ml.

b. Solid media.

L plates

1.5% Bacto-agar was added to L broth. Plates were poured from 30 ml of the appropriate medium, dried overnight at 37°C and stored at 4°C.

Antibiotics were added to the medium or spread onto plates as follows: ampicillin at 100 μg/ml and chloramphenicol at 30 μg/ml. Plates were poured from 30 ml of the appropriate medium, dried overnight at 37°C and stored at 4°C. When selecting for putative clones by insertional inactivation of the *lacZ* gene or putative promoter-*lacZ*

fusions, transformants were spread on LB plates containing the appropriate antibiotics and supplemented with 0.02 mg/ml X-Gal.

Soft agar overlay

0.7% Bacto-agar, 10mM $MgCl_2$. Used for λ platings.

7.A.6.2. Buffers and solutions.

Cloned *Pfu* DNA polymerase reaction buffer (10 x)

200mM Tris-HCl pH 8.8, 20mM $MgSO_4$, 100 mM KCl, 100 mM $(NH4)_2SO_4$, 1% Triton X-100, 1mg/ml nuclease-free BSA (Stratagene).

Formamide stop/load buffer

80% (v/v) deionised formamide, 20 mM EDTA, 0.1% (w/v) bromophenol blue, 0.1 % (w/v) xylene cyanol

Glycerol loading buffer (10 x)

50% (v/v) glycerol, 0.40% (w/v) Bromophenol Blue, 0.20% (w/v) Xylene Cyanol, 10 mM EDTA.

Ligation buffer (10 x)

500 mM Tris pH 7.5, 100 mM $MgCl_2$, 10 mM ATP, 100mM DTT. (NEB)

Phage Storage Buffer (PSB)

10 mM Tris-HCl, pH 7.1, 10 mM $MgSO_4$, 100 mM NaCl, 0.05% gelatin. Used for preparation and storage of λ.

Polynucleotide kinase buffer (10x)

500 mM Tris pH 7.9, 100 mM $MgCl_2$, 100 mM, β-mercapto-ethanol, 100 mM EDTA. Stored at -20° C. (GeneWorks)

TAE (10 x)

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

Taq DNA polymerase reaction buffer (10 x)

670 mM Tris-HCl pH 8.8, 166 mM $(NH_4)_2SO_4$,

0.45% Triton X-100, 2 mg/ml gelatin (Biotech International)

TBE (10 x)

0.89 M Tris-HCl, 0.89 M boric acid, 2.7 mM EDTA, pH 8.3.

Transcription buffer (1 x)

20 mM Tris-acetate pH 8.0, 3 mM Mg acetate, 200 mM K glutamate, 1 mM DDT, 5% glycerol and 0.3 units/µl RNase inhibitor. Buffer was stored at –20°C and DDT and RNase inhibitor was added immediately prior to use.

TZ8

100 mM Tris-HCl pH 8.0, 1 mM $MgSO_4$, and 10 mM KCl, used for lacZ assays.

7.A.7. DNA MARKERS.

DNA size markers were all purchased from GeneWorks (Australia) and 500 ng were routinely loaded on an agarose gel.

HpaII digest of pUC19 DNA at 500 ng/µl. Fragment sizes in bp: 501, 489, 404, 331, 242, 190, 147, 111, 110, 67, 34, 34, 26.

EcoRI digest of phage SPP-1 DNA at 500 ng/µl. Fragment sizes in bp: 7840, 6960, 5860, 4690, 3370, 2680, 1890, 1800, 1450, 1330, 1090, 880, 660, 480, 380.

7.A.8. DNA purification kits.

All kits were used according to the manufacturers specifications, using the buffers, solutions and materials contained within the kit.

QIAquick Nucleotide Removal kit: Qiagen

QIAquick Gel Extraction Kit: Qiagen

QIAprep Spin Miniprep Kit: Qiagen

QIAfilter Plasmid Midiprep Kit: Qiagen

Ultra Clean PCR Clean-up DNA purification kit: MO BIO laboratories

7.B Methods

7.B.1 BACTERIAL AND PHAGE PROCEDURES.

7.B.1.1. Storage of bacterial and phage stocks.

Bacterial stocks for short term storage were maintained on the appropriate plates at 4°C. Long term storage of bacterial cultures was at –80°C, after addition of glycerol to a final concentration of 10%. Low titre λ phage stocks were stored in PSB mixed with a few drops of chloroform and stored at 4°C. Phage stocks for longer term storage were kept at –80°C after addition of glycerol to a final concentration of 20%.

7.B.1.2. Growth of bacterial strains.

Stationary phase bacterial cultures were prepared by inoculating broth with a single colony of bacteria from a plate stock, or a loopful of bacteria directly from a glycerol stock, and incubating overnight with aeration at 37°C. Log phase cultures and indicator bacteria were prepared by diluting a fresh stationary culture 50-200 fold into sterile broth and incubating with aeration at the appropriate temperature, until the required cell density was reached. 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.

7.B.1.3. Preparation and Transformation of calcium chloride competent cells.

Bacterial cells competent for DNA transformation were prepared by inoculating 50-100 ml of fresh LB with the appropriate bacterial strain and incubating at 37°C, with aeration, to an A_{600} = 0.4-0.6. The culture was chilled on ice for 10 min and the cells harvested by centrifugation (5 min, 5000 rpm, 4°C, JA10). The cells were then resuspended in 10 ml cold $CaCl_2$ solution (100 mM $CaCl_2$, 10% glycerol) and left on ice for a further 60 min. The cells were again harvested and resuspended in 2-4 ml cold

$CaCl_2$ solution. Cells were either used fresh after an overnight incubation on ice at 4°C or 400 μl aliquots were transferred to pre-chilled tubes, snap frozen on a dry ice ethanol bath and stored at –75°C until use.

To transform competent cells 100 µl of cells were aliquoted into Eppendorf tubes on ice and 1-10 µl DNA (usually 5-10 ng DNA) added. The mix was incubated on ice for 20 min, transferred to a 42°C heating block for 2 min, then 200 µl of SOC medium was added and the mixture incubated at 37°C for 1-2 hours then spread directly onto selective agar plates prewarmed to 37°C. Plates were incubated overnight at 37°C. As a transformation control (to determine cell competence) 5 ng of pBluescript was added to one cell aliquot.

For use of the *lacZ* blue/white colony screening of pBluescript plasmids or pBC2 plasmids, 20 μl of 100 mM IPTG and 40 μl of 20 mg/ml X-gal was spread onto the plates and allowed to dry before plating the transformation mix.

Recombinant colonies were verified by colony PCR of patched cells using primers flanking the insert, or an internal and flanking pair of primers (to determine insert orientation). Alternatively, restriction digestion was used to analyse cloning results. Recombinant colonies were then spread onto appropriate plates and single colonies isolated, and used to innoculated a 4 ml overnight culture. Cultures were used for glycerol stocks and sequencing of the clone.

7.B.1.4. Preparation and transformation of electrocompetent cells.

Cells competent for electroporation were prepared by inoculating 500 ml fresh YENB with the appropriate culture and grown overnight at 37°C. The culture was chilled on ice for 10 min and the cells harvested by centrifugation (5 min, 5000 rpm, 4°C, JA10). The medium was discarded and the cells were then washed twice in 100 ml cold H_2O, harvesting cells as before. Finally cells were again harvested and resuspended in 2 ml of cold 10% glycerol. 40 μl aliquots were transferred to pre-chilled tubes and stored at –75°C until use.

For electrotransformation either 1μl of DNA in low ionic strength buffer (usually a ligation mix was diluted 1:50 into MQ) were added to 40μl of electrocompetent cells, mixed and incubated on ice for 1-5 min. Cells were then transferred to cold 0.2-cm electroporation cuvettes (BioRad) and electroporated in a BioRad MicroPulser according to the manufacturers instructions. The pulse produced had a time constant of about 4.7 ms. Cuvettes were removed from the chamber and 1 ml of SOC or YENB buffer was immediately to the cuvette. The resuspended cells were transferred to an eppendorf tube and incubated for 1-3 hours at 37°C and finally plated on selective medium plates and incubated overnight at 37°C.

7.B.1.5. Low titre λ phage stocks.

Low titre plug stocks of λ phage were obtained by placing 1-3 plaques into 100μl PSB plus a few drops of chloroform which was then vortexed. Phage were eluted for 30 minutes at room temperature and then used immediately or stored at 4°C. Larger volumes of low titre phage stocks were made by scraping the top agar from a confluent indicator plate infected with the appropriate phage into 3 ml of PSB (a confluent plate was made by plating 10^5 phage on a lawn of indicator bacteria and incubating overnight at 37°C). After 15 min at room temperature, chloroform was added to the solution, it was vortexed and then centrifuged at 3000 rpm for 5 min. The supernatant was removed and a few drops of chloroform added before storage at 4°C.

7.B.1.6. Plating and assaying phage

λ phage were plated or assayed for plaque forming units (pfu) by mixing 10-100 μl of phage diluted in PSB with 0.2 ml of log phase indicator bacteria and then adding 3 ml of molten (0.7%) soft agar overlay to the mixture and pouring onto LB plates. The agar was allowed to solidify and the plates were inverted and incubated overnight at 37°C. Plaques were counted and scored as plaque forming units per ml (pfu/ml). If the bacterial lawn contained a plasmid, LB plates supplemented with the appropriate antibiotic were used.

7.B.1.7. Construction of chromosomal *lacZ* fusions.

Single-copy chromosomal promoter-*lacZ* fusions were obtained by *in vivo* homologous recombination with λRS45 (or λRS45ΔYA) followed by lysogenisation (Simons *et al.*, 1987). λRS45 (or λRS45ΔYA) and the *lacZ* reporter vectors used share portions of the amino terminus of both the ampicillin gene and the *lacZ* gene thus allowing the promoter insert to be recombined into the phage. Recombinant phage were obtained by spotting 10 μl of a low-titre stock of phage λRS45 or λRS45ΔYA onto a lawn of *RecA*$^+$ indicator bacteria (usually MC1061.5) which had been transformed with the plasmid carrying the *lacZ* reporter construct to be assayed. Plug stocks were obtained and recombinant phage were purified from non-recombinant phage by plating (and replating blue plaques) on an appropriate indicator strain (either C-2420, MC1061.5 or NK7049) on the basis of colour in presence of 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal). Blue lysogens were isolated by spreading the centre of a purified, recombinant, turbid plaque, picked using a toothpick, onto LB plates spread with X-gal, and isolated blue colonies were purified by restreaking. Single blue colonies were tested for the presence of single or multiple copy λ lysogens using PCR with λ primers (primer numbers 154, 155 and 156) by the method of (Powell *et al.*, 1994). Detection of a single 501bp PCR fragment indicated a single lysogen while detection of two PCR fragments of 501 bp and 379 bp indicated a multiple lysogen. Single lysogen were kept and assayed.

7.B.2. DNA MANIPULATION.

7.B.2.1. Plasmid miniprep.

Small scale plasmid DNA preparations were performed using QIAprep Spin Plasmid Miniprep Kit (Qiagen) according to the manufacturers instructions. Elution of bound DNA from the silica spin columns was performed using either 30 or 50 μl of Qiagen elution buffer. Eluted DNA was routinely stored at –20°C.

7.B.2.2. Large scale plasmid purification.

Large scale preparations of plasmid DNA were obtained either by using the QIAfilter Midiprep Kit (Qiagen) or by alkaline extraction followed by sedimentation on a CsCl gradient. Midiprep DNA was prepared according to the manufacturers instructions and

purified DNA was finally resuspended in 100-150 µl of Qiagen elution buffer and stored at –20°C.

For alkaline extraction, a 200-500 ml saturated culture was pelleted (5K 10' 4°C) and resuspended in 4 ml of lysozyme solution (15% glucose, 50 mM Tris-Cl pH8.0, 5mM EDTA, 4 mg/ml lysozyme). Following incubation at room temperature for 10', 8 ml of 0.2M NaOH, 1% SDS was added, the solution was mixed gently and placed on ice for 10'. The pH was neutralized by addition of 5 ml of 3M NaAc pH4.6 and the solution placed on ice for a further 10' before cell debris was pelleted by centrifugation (16K 30' 4°C). Nucleic acids were precipitated with addition of 5 ml isopropanol (10' room temperature) and pelleted by centrifugation (10K 40' 4°C). (For starting cultures of 100 ml above volumes were halved).

For purification over a CsCl gradient the nucleic acid pellet was well drained and resuspended in 1.4 ml of water. 1.5 g of CsCl and 120µl of 10 mg/ml Ethidium bromide were dissolved in the solution before it was transfered to a 2.2 ml quickseal polyallomer tube. After sealing the tube, DNA was banded in a Beckmann TL-100 at 80 K for 12 hours at 20°C.

If DNA was required the same day, the dried nucleic acid pellet was resuspended in 720 µl of water and 1.26 g of CsCl was dissolved in this solution. Following addition of 120 ml of 10 mg/ml Ethidium bromide, 500 µl of this DNA solution was carefully layered under 1.4 ml of a 65% CsCl solution in a 2.2 ml quickseal polyallomer tube. The DNA was banded by centrifugation at 100K for 3 hrs at 20°C.

The DNA band obtained after centrifugation was removed and purified from Ethidium bromide and CsCl by isopropanol extraction. The DNA solution was extracted with an equal volume of saturated isopropanol (isopropanol saturated with 5M NaCl, 10mM Tris pH8.0, 1mM EDTA pH8.0) at least 3 times or until all visible traces of colour were removed and subsequently precipitated by addition of 2 volumes of water and 6 volumes of 95% ethanol, incubation at –20°C for 20' and centrifugation 12K 15' 4°C. The DNA obtained was pelleted twice from 70% ethanol before being used. For DNA

to be used in *in vitro* transcription assays further purification was achieved by using Qiagen PCR purification kit.

7.B.2.3. Ethanol precipitation.

Ethanol precipitation of plasmid DNA was routinely used to remove salt and other contaminants, or when changing enzymatic buffers. To a solution of DNA, 0.1 volumes of 3M NaAc (pH 5.2) and 2 volumes 95% cold ethanol (RNase-free) were added (with mixing after each addition), and incubated on ice for 5-20min, or immediately centrifuged (12K, 15' 4°C Eppendorf centrifuge or 16K, 20' 4°C JA20 rotor). The supernatant was removed and the pellet sometimes resuspended in 70% ethanol (v/v) and centrifuged as previously or rinsed in 70% ethanol (v/v) and dried in vacuo for 10'. The DNA was finally resuspended in Qiagen elution buffer (10 mM Tris-Cl, pH 8.5) or water and stored at 4°C or –20°C.

7.B.2.4. Agarose gel electrophoresis.

DNA purity, size and quantity were determined by agarose gel electrophoresis on 1-2% agarose in horizontal minigels. DNA samples were combined with loading buffer, and electrophoresis performed in 1x TAE buffer at 90-120V. DNA bands were visualized by staining with low concentration ethidium bromide, and photographed under short wavelength UV light. Comparison of band intensity (and therefore ethidium bromide staining) with that of molecular weight markers of known concentrations, allowed approximate DNA concentrations to be determined.

7.B.2.5. Determination of DNA concentration.

Accurate determination of DNA concentration for plasmid solutions to be used in *in vitro* trancription assays was performed spectrophotometric measurement of the absorption at 260nm. 5-10 µl of DNA solution was added to a total of 200 µl water and A_{260} was measured using a Cary 3 Bio UV- visible spectrophotometer. Calculations of DNA concentration were made on the assumption that 1 A_{260} unit = 50 µg/ml of dsDNA.

7.B.2.6. Isolation of DNA fragments from agarose gels.

Restricted DNA fragments or plasmids required for cloning were excised from an agarose gel by cutting with a scalpel blade (using ethidium bromide visualization under long wavelength UV light). DNA was isolated from the agarose using the QIAquick Gel Extraction Kit (Qiagen) as specified by the manufacturers.

7.B.2.7. Restriction digestion.

Restriction digests were performed in conditions recommended by the manufacturers, in 10-50μl for 1hr to overnight at the recommended temperature. Digestion was checked by agarose gel electrophoresis.

7.B.2.8. Reactions with alkaline phosphatase.

To reduce background religation of digested vectors during cloning, linearised vectors with compatible ends were treated with Shrimp alkaline phosphatase (SAP) by 1hr incubation (37°C) with 1-2 Units SAP in 1x restriction enzyme buffer. The enzyme was inactivated by heating to 65°C for 20'. SAP reactions were deemed efficient when the number of colonies obtained after transformation of the same concentration of cut vector DNA with or without the addition of ligase were the same.

7.B.2.9. Blunt ending 5' and 3' overhangs

T4 DNA polymerase was used to remove 3' overhangs or fill-in 5' overhangs. Reactions were performed in the buffers used for restriction digests but supplemented with 50 μg/ml BSA and 100 μM of each dNTP. 1-3 units of T4 DNA polymerase was added and reactions incubated at 16°C for 20 min then the enzyme was heat inactivated by incubating at 75°C for 10 min.

7.B.2.10. Polymerase Chain Reaction (PCR).

Routine PCRs were performed in 10 μl reactions containing 1x PCR reaction buffer, 2 mM $MgCl_2$, 0.2 mM dNTP mix, 2ng of each primer and 0.5 Units of Taq polymerase. Plasmid DNA was used as a template at a concentration of 1-5ng, or a colony was

picked and swirled into the reaction mixture, for colony screening (the cycle heats the cell and releases plasmid DNA). The PCR cycle used was 2 min 94°C, then 30 cycles of 5 sec 94°C, 5 sec 50°C and 45 sec 74°C in a Rapidcycler from Idaho Technology and reactions were contained in thin-walled 0.2 ml PCR tubes. $MgCl_2$ and cycling reaction conditions were optimised for each primer pair to produce high yields of specific product. If PCR products were to be used in further enzymatic reactions such as sequencing, purification was achieved by pooling a number of PCR reactions and using the UltraClean PCR purification kit.

PCR reactions using *Pfu turbo* were used for preparation of inserts for cloning. 20 µl reactions contained 1 x cloned *Pfu* reaction buffer, 0.2 mM dNTP mix, 0.4 µM of each primer and 2 units of *Pfu* plus template and $MgSO_4$. Template was ~1ng of plasmid DNA and the concentration of $MgSO_4$ used was optimised for each primer pair but usually between 2-6 mM. The PCR cycle used for extension products under 1 kb in length was 2 min 94°C, then 30 cycles of 10 sec 98°C, 10 sec 50°C and 60 sec 74°C in a RapidCycler from Idaho Technology and reactions were contained in thin-walled 0.2 ml PCR tubes. For longer products a longer extension time was used (~2 min). Products were purified using the UltraClean PCR purification kit prior to use in cloning reactions.

7.B.2.11. Preparation of annealed oligos for cloning

To prepare the inserts for cloning of tA in chapter 3 long double stranded DNA was prepared from two complimentary oligos containing the relevant sequences. A solution of 2 ng/µl of each oligo was placed in a heating block at 80°C then the block was switched off and allowed to cool slowly to room temperature.

7.B.2.12. Insert preparation by primer extension

Inserts for $tA^{=}$ clones were constructed by extension of a primer (#57) that anneals to the end of a long oligo containing $tA^{=}$ sequence (#401). Primer extension reactions were equivalent to *Pfu* turbo PCR reactions using 0.4 µM of each oligo.

7.B.2.13. DNA Ligations.

Ligations were performed in 10μl or 20ul reactions containing 1x ligation buffer, 0.5-2 Units T4 DNA ligase and an approximate 3:1 ratio of insert to vector (100-200 ng). The mix was incubated for 1-3 hr at room temperature or 4-16hrs at 16°C . Half of the ligation mix was used to transform into calcium chloride competent cells, or 0.1-0.4 μl was used to transform electrocompetent cells.

7.B.2.14. Sub-cloning of PCR products

Some PCR products required sub-cloning into the blue-white colour selection cloning vector pBluescriptSK+. PCR inserts prepared by *Pfu* polymerase produce blunt ends. PCR inserts were therefore ligated with pBS which had been digested with EcoRV. Competent DH5α cells were transformed with the ligation mix and after blue/white colour selection white transformants were tested for the presence of the correct insert by PCR then sequenced. A miniprep of the correct sub-clone was obtained and digested with appropriate restriction enzymes to release the appropriate sub-cloned fragment. The fragment was then isolated by agarose gel electrophoresis and cloned into the appropriate vector.

7.B.2.15. Site-directed mutagenesis

The QuickChange method of Stratagene was used to generate desired promoter mutations. Complementary primers were designed which contain the desired mutations and enough wild-type sequence flanking the mutations to allow stable annealing of the primers to opposite strands of a double-stranded vector containing the promoter to be mutated. The oligos were annealed to the vector and extended by *Pfu Turbo* DNA polymerase in a reaction containing 1 x *Pfu* DNA polymerase reaction buffer, 5-50ng of dsDNA template, 125 ng of each primer, 1 μl of 10mM dNTP mix, 1 ml of *Pfu turbo* DNA polymerase (2.5 U/μl) and distilled water to a volume of 50 μl. Extension was performed by first overlaying the reaction with 30 μl of mineral oil then cycling the reaction in a MJ Research, Inc. PTC-100™ Programmable Thermal Controller using 30 sec 95°C, then 15 cycles of (30 sec 95°C, 1 min 65°C, 15 min 68°C) followed by rest at 4°C. To digest the template DNA, 2 μl of the methylation dependent DpnI restriction enzyme (5 U/μl) was added directly to the amplification reaction and then incubated

overnight at 37°C. The reaction was purified and concentrated by ethanol precipitation and a proportion of the reaction transformed into electrocompetent DH5α cells. Transformants carrying the correct mutation were identified by PCR and restriction digest for the appropriate promoter mutations and confirmed by sequencing.

7.B.2.16. Big Dye Sequencing reactions

Sequencing reactions were performed using Perkin-Elmer ABI PRISM Big Dye Version 3. Templates for reactions were mostly purified PCR products and occasionally miniprep DNA was used. Reactions contained 2 μl Big Dye v3 Ready Mix, 6 μl dilution buffer (200mM Tris-HCl (pH 9) and 5mM $MgCl_2$), 60-180 ng of template and 3.2 pmol of primer in a volume of 20 μl. Reactions were cycled using 20 min 96°C, then 25 cycles of 30 sec 96°C, 30 sec 50°C and 4 min 74°C in a RapidCycler from Idaho Technology and reactions were contained in thin-walled 0.2 ml PCR tubes.

Completed reactions were precipitated using isopropanol. 80 μl of 75% isopropanol was added to the reaction and tubes vortexed briefly. Extension products were precipitated for 15 min at room temperature, then placed in a microcentrifuge and centrifuged at maximum speed for 20 min. The supernatant was removed by pipetting then pellets were washed with 250 μl of 75% isopropanol. After centrifuging for a further 5 min, and removing supernatant, samples were dried in the 37°C room for 1 hour. Samples were analysed at the Institute for Medical and Veterinary Science Sequencing Centre. Sequences were analysed by examination of chromatograph files using EditView 1.0.1 and the use of DNA Strider 1.3 to compare sequence files.

7.B.2.17. Labelling DNA markers

Radiolabelled DNA markers were used to analyse the size of labelled RNA transcripts from *in vitro* transcription assays. 1 μg of pUC19/ HpaII markers were labelled in a 10 μl reaction mixture containing 1x Klenow buffer, 0.5 units of Klenow fragment, 0.05 mM dNTPs (dA, dG, dT) and 10 μCi [α-^{32}P]-CTP, which was incubated at 37°C for 15' then heated to 70°C for 5' to inactivate the enzyme. After diluting the reaction 2 fold,

1 volume of formamide load buffer was added. Labelled marker was stored in a lead pot at 4°C. When required 1 µl of marker was loaded onto a sequencing gel after denaturing the DNA (95°C for 5').

7.B.2.18. Preparation of loading control used during *in vitro* transcription.

The loading control used for accurate quantitation of *in vitro* transcription reactions was prepared by PCR using primers SK and USP and the template was pBluescript SK+ plasmid DNA. Prior to the PCR reaction, the SK primer was end labelled using T4 polynucleotide kinase (PNK). Reactions were performed in a 5 µl volume containing100 ng/µl SK primer, 5 units of PNK and 5 µCi [γ-^{32}P]-dATP in a buffer of 70mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$ and 5 mM DTT, and incubated at 37°C for 30 min. PNK was then inactivated by incubation at 70°C for 20 min. Labelled primer was then used in a routine PCR reaction with USP and template. Reactions were purified using a PCR purification kit and eluted in elution buffer and stored at –20°C util required. During *in vitro* transcription the loading control was added to loading/stop buffer prior to termination of the reactions at a sufficient concentration to allow visualisation and quantitation after electrophoresis of the reactions (usually ~0.2 to 0.5 µl were loaded, depending on the activity of the label).

7.B.3. LacZ assays.

Kinetic LacZ assays were done in 96 well microtitre plates by an extensively modified method of Miller (1972) as described by Dodd *et al.* 2001. Fresh colonies on selective LB plates were resuspended in LB and used to inoculate 200 µl of LB +antibiotic +IPTG. Dishes were sealed and incubated for ~16 h without shaking. These cultures were subcultured by diluting 2 µl into 98 µl fresh medium and incubated with rotation to an OD_{600} of 0.2–1.2 (log phase). OD_{600} was measured using a Labsystems Multiskan Ascent plate reader with a 620 nm filter; the OD_{620} values were converted to OD_{600} (1 cm path length) values using an empirically derived relationship and adjusted for light-scattering non-linearity according to (Bipatnath *et al.*, 1998). Cells were chilled and then permeabilized with polymyxin B (Schupp *et al.*, 1995) by adding either 10 µl culture + 40 µl LB (when assaying strong promoters, *pR* and *pe*) or 50 µl culture (when assaying weaker promoters *pL* and *pc*) to 150 µl lysis buffer in a microtitre dish.

Lysis buffer was TZ8 (100 mM Tris-HCl pH 8.0, 1 mM $MgSO_4$, 10 mM KCl) + 2.7 µl/ml 2-mercaptoethanol and 50 µg/ml polymyxin B. The presence of detergents and chelating agents (Schupp *et al.*, 1995) did not improve the assay. Use of pH 8 for the assay buffer rather than the pH 7 used by Miller (1972) improved display of *o*-nitrophenol in the absence of Na_2CO_3 added to stop the reaction. Assays were at 28° and were begun by addition of 40 µl *o*-nitrophenyl-β-D galactoside (4 mg/ml in TZ8). The plate reader was used to incubate the reactions and take A_{414} readings every 2 min for 1 h. Enzyme activity was determined as the slope of the line of best fit of A_{414} versus time (readings with $A_{414} > 2.5$ were ignored). Enzyme activity was found to be directly proportional to the OD_{600} of the culture and the volume of culture added to the assay (V - in µl). LacZ units were calculated as 200,000 x $(A_{414}/min)/(A_{600} \times V)$ and were roughly equivalent to standard Miller units. Background units (mostly less than 1) from strains carrying the relevant *lacZ* recombinant prophage without promoter inserts was calculated for each plate and subtracted from the units calculated for all other strains.

The method of calculating average lacZ units for a given strain, and how fold interference was calculated from the measured lacZ units, is explained in Chapter 2.

7.B.4. *In vitro* transcription assays

In vitro transcription assays conditions were based on those used by (Ryu *et al.*, 1994); all reactions were performed in a 37°C room and reactions and reagents were pre-warmed prior to use. The plasmid DNA template used in the reactions was prepared from DH5α cells using either Qiagen midiprep kits or purification over a CsCl gradient, and DNA concentrations were quantified by spectrophotometry. RNAP (50 nM final concentration) was first incubated with supercoiled DNA (2 nM final concentration) for various times (as indicated in the figure legends of the transcript patterns) in transcription buffer (20mM Tris-acetate pH 8.0, 3mM Mg acetate, 200 mM K glutamate, 1mM DDT, 5% glycerol and 0.3 units/µl RNase inhibitor) to allow open complex formation. Elongation was initiated by the addition of rNTP/heparin solution giving a final concentration of 0.2 mM each of rATP, rCTP and rGTP, 0.02mM rUTP, 0.5 mCi/ml [α^{32}P]-UTP and 0.5 mg/ml heparin salt. Elongation was terminated at

various times (indicated in the figure legends of transcript patterns) by the addition of an equal volume of formamide stop/load buffer containing the loading control.

To analyse the transcript patterns, samples were heated to 80°C for 5 minutes and then 10 µl were loaded onto a 6% polyacrylamide gel (Sequagel 6, National Diagnostics) and subjected to electrophoresis. Polyacrylamide gels were pre-electrophoresed for 30', and the wells were flushed prior to loading. Electrophoresis was performed for 2-2.5 hr in 1x TBE at a constant temperature of 50 °C. After electrophoresis, the gel was transferred to Whatman filter paper and vacuum dried (60-70 °C, 1 hr). The dried gel was placed into an erased phosphorimager imaging plate (Fuji Photo Film Co.) and exposed overnight. Transcripts were observed using a BioRad FX phoshorimager. Using the software Quantity One the relative amount of each full length transcript accumulated was determined by quantitating the volume for a small rectangle containing the relevant band, then subtracting the background volume of the lane (obtained using an equivalent sized rectangle placed in a nearby region of the same lane which contains no major transcripts). Final values were then normalised first to the value of the loading control for each lane (obtained in a similar way) and then to the uridine content of each full length transcript. The uridine content for full-length transcripts were 61 uridines for the 242 nt $pR(pL^{-})$ and $pR(pL^{+})$ transcript, 76 uridines for the 288 nt $pL(pR^{=})$ transcript, 77 uridines for the 288 nt $pL(pR^{+})$ transcript, 71 uridines for the 288 nt $pc(pe^{-})$ transcript and 27 uridines for the 108 nt RNA1 transcript. Variations of this method used for single round assays, multiple round assays, and assaying clearance rates or rates of open complex formation are described in the figure legends of the relevant transcript patterns. The transcript patterns shown were first contrast adjusted using Quantity One and prepared for presentation using Canvas 6.

REFERENCES

References

Adams, L.W. and Hatfield, G.W. (1984) Effects of promoter strengths and growth conditions on copy number of transcription-fusion vectors. *J. Biol. Chem.*, **259**, 7399-7403.

Adelman, K., La Porta, A., Santangelo, T.J., Lis, J.T., Roberts, J.W. and Wang, M.D. (2002) Single molecule analysis of RNA polymerase elongation reveals uniform kinetic behavior. *Proc. Natl. Acad. Sci. USA*, **99**, 13538-13543.

Adhya, S. and Gottesman, M. (1982) Promoter occlusion: transcription through a promoter may inhibit its activity. *Cell*, **29**, 939-944.

Atlung, T. and Ingmer, H. (1997) H-NS: a modulator of environmentally regulated gene expression. *Mol. Microbiol.*, **24**, 7-17.

Beck, C.F. and Warren, R.A.J. (1988) Divergent promoters, a common form of gene organization. *Microbiological Reviews*, **52**, 318-326.

Bertrand-Burggraf, E., Hurstel, S., Daune, M. and Schnarr, M. (1987) Promoter properties and negative regulation of the *uvrA* gene by the LexA repressor and its amino-terminal DNA binding domain. *J. Mol. Biol.*, **193**, 293-302.

Bipatnath, M., Dennis, P. and Bremer, H. (1998) Initiation and velocity of chromosome replication in *Escherichia coli* B/r and K-12. *J. Bacteriol.*, **180**, 265-273.

Brantl, S. (2002) Antisense-RNA regulation and RNA interference. *Biochem Biophys Acta*, **1575**, 15-25.

Brantl, S. and Wagner, E.G.H. (1997) Dual function of the *copR* gene product of plasmid pIP501. *J. Bacteriol.*, **179**, 7016-7024.

Bushman, F.D. and Ptashne, M. (1986) Activation of transcription by the bacteriophage 434 repressor. *Proc. Natl. Acad. Sci. USA*, **83**, 9353-9357.

Cashel, M., Hsu, L.M. and Hernandez, V.J. (2002) Changes in conserved region 3 of *Escherichia coli* σ^{70} reduce abortive transcription and enhance promoter escape. *JBC Papers in Press*, **M211430200**.

Charlier, D., Piette, J. and Glansdorff, N. (1982) IS3 can function as a mobile promoter in *E. coli*. *Nucl. Acid. Res.*, **10**, 5935-5947.

Choy, H.E. and Adhya, S. (1993) RNA polymerase idling and clearance in *gal* promoters: Use of supercoiled minicircle DNA template made *in vivo*. *Proc. Natl. Acad. Sci. USA*, **90**, 472-476.

Consortium., FANTOM and the RIKEN genome Exploration Research Group Phase I & II Team. (2002) Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature*, **420**, 563-573.

Contursi, C., Minchiotti, M. and Di Nocera, P.P. (1993) Functional dissection of two promoters that control sense and antisense transcription of Drosophila melanogaster F elements. *J. Mol. Biol.*, **234**, 988-997.

Cramer, P., Bushnell, D. and Kornberg, R. (2001) Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science*, **292**, 1863-1876.

Cunin, R., Eckhardt, T., Piette, J., Boyen, A., Pierard, A. and Glansdorff, N. (1983) Molecular basis for modulated regulation of expression in the arginine regulon of *Eschericia coli* K-12. *Nucl. Acid. Res.*, **11**, 5007-5017.

Davenport, R.J., Wuite, G.J., Landick, R. and Bustamante, C. (2000) Single-molecule study of transcriptional pausing and arrest by *E. coli* RNA polymerase. *Science*, **287**, 2497-2500.

Deutschle, U., Gentz, R. and Bujard, H. (1986) *lac* repressor blocks transcribing RNA polymerase and terminates transcription. *Proc. Natl. Acad. Sci. USA*, **83**, 4134-4137.

Dodd, I.B. and Egan, J.B. (2002) Action at a distance in CI repressor regulation of the bacteriophage 186 genetic switch. *Molecular Microbiology*, **45**, 679-710.

Dodd, I.B., Kalionis, B. and Egan, J.B. (1990) Control of gene expression in the temperate coliphage 186. VIII. Control of lysis and lysogeny by a transcriptional switch involving face-to-face promoters. *J. Mol. Biol.*, **214**, 27-37.

Dodd, I.B., Perkins, A.J., Tsemitsidis, D. and Egan, J.B. (2001) Octamerization of λ CI repressor is needed for effictive repression of P_{RM} and efficient switching from lysogeny. *Genes Dev.*, **15**.

Dodd, I.B., Reed, M.R. and Egan, J.B. (1993) The Cro-like Apl repressor of coliphage 186 is required for prophage excision and binds near the phage attachment site. *Mol. Microbiol.*, **10**, 1139-1150.

Dorman, C.J. (1995) DNA topology and the global control of bacterial gene expression: implications for the regulation of virulence gene expression. *Microbiology*, **141**, 1271-1280.

Dujon, B. (1996) The yeast genome project: what did we learn? *Trends in Genetics*, **12**, 263-270.

Dvir, A. (2002) Promoter escape by RNA polymerase II. *Biochim Biophys Acta*, **1577**, 208-223.

Ebright, R.H. (2000) RNA polymerase: structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II. *J. Mol. Biol.*, **304**, 687-698.

Eddy, S.R. (2002) Computational genomics of noncoding RNA genes. *Cell*, **109**, 137-140.

Elias-Arnanz, M. and Salas, M. (1997) Bacteriophage ϕ29 DNA replication arrest caused by codirectional collisions with the transcription machinery. *EMBO J.*, **16**, 5775-5783.

Elias-Arnanz, M. and Salas, M. (1999) Resolution of head-on collisions between the transcription machinery and bacteriophage ϕ29 DNA polymerase is dependent on RNA polymerase translocation. *EMBO J.*, **18**, 5675-5682.

Elledge, S.J. and Davis, R.W. (1989) Position and density effects on repression by stationary and mobile DNA-binding proteins. *Genes Dev.*, **3**, 185-197.

Ellinger, T., Behnke, D., Bujard, H. and Gralla, J.D. (1994) Stalling of *Escherichia coli* RNA polymerase in the +6 to +12 region *in vivo* is associated with tight binding to consensus promoter elements. *J. Mol. Biol.*, **239**, 455-465.

Erie, D.A. (2002) The many conformation states of RNA polymerase elongation complexes and their roles in the regulation of transcription. *Biochimica et Biophysica Acta*, **1577**, 224-239.

Esposito, D., Wilson, J.C.E. and Scocca, J.J. (1997) Reciprocal regulation of the early promoter region of bacteriophage HP1 by the Cox and CI proteins. *Virology*, **234**, 267-276.

Eszterhas, S.K., Bauhassira, E.E., Martin, D.I.K. and Fiering, S. (2002) Transcriptional interference by independently regulated genes occurs in any relative arrangement of the genes and is influenced by chromosomal integration position. *Mol. Cell. Biol.*, **22**, 469-479.

Ferrigno, O., Virolle, T., Djabari, Z., Ortonne, J.-P., White, R.J. and Aberdam, D. (2001) Transposable B2 SINE elements can provide mobile RNA polymerase II promoters. *Nature Genet*, **28**, 77-81.

Gafny, R., Cohen, S., Nachaliel, N. and Glaser, G. (1994) Isolated P2 rRNA promoters of *Escherichia coli* are strong promoters that are subject to stringent control. *J. Mol. Biol.*, **243**, 152-156.

Gnatt, A., Cramer, P., Fu, J., Bushnell, D. and Kornberg, R. (2001) Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. *Science*, **292**, 1876-1882.

Goosen, N. and van de Putte, P. (1995) The regulation of transcription initiation by integration host factor. *Mol. Microbiol.*, **16**, 1-7.

Gusarov, I. and Nudler, E. (1999) The mechanism of intrinsic transcription termination. *Molecular Cell*, **3**, 495-504.

Guthold, M., Zhu, X., Rivetti, C., Yang, G., Thomson, N.H., Kasas, S., Hansma, H.G., Smith, B., Hansma, P.K. and Bustamante, C. (1999) Direct observation of one-dimensional diffusion and transcription by *Escherichia coli* RNA polymerase. *Biophysical Journal*, **77**, 2284-2294.

Hawley, D. and McClure, W.R. (1983) Compilation and analysis of *Esherichia coli* promoter DNA sequences. *Nucl. Acid. Res.*, **11**, 2237-2255.

Heggeler-Bordier, B., Wahli, W., Adrian, M., Stasiak, A. and Dubochet, J. (1992) The apical localization of transcribing RNA polymerase on supercoiled DNA prevents their rotation around the template. *EMBO J.*, **11**, 667-672.

Hippel, P.H. (1994) Passing lanes for polymerases? *Current Biology*, **4**, 333-336.

Horowitz, H. and Platt, T. (1982) Regulation from tandem and convergent promoters. *Nucl. Acid. Res.*, **10**, 5447-5466.

Hsu, L.M. (1996) Quantitative parameters for promoter clearance. *Methods Enzymol.*, **273**, 59-71.

Hsu, L.M. (2002) Promoter clearance and escape in prokaryotes. *Biochimica et Biophysica Acta*, **1577**, 208-223.

Hsu, L.M., Vo, N.V. and Chamberlain, M.J. (1995) *Escherichia coli* transcript cleavage factors GreA and GreB stimulate promoter escape and gene expression *in vivo* and *in vitro*. *Proc. Natl. Acad. Sci. USA*, **92**, 11588-11592.

Ingelbrecht, I., Breyne, P., Vancompernolle, K., Jacobs, A., Van Montagu, M. and Depicker, A. (1991) Transcriptional interference in transgenic plants. *Gene*, **109**, 239-242.

Jagura-Burdzy, G. and Thomas, C.M. (1997) Dissection of the switch between genes for replication and transfer of promiscuous plasmid RK2: basis of the dominance of *trfAp* over *trbAp* and specificity in controlling the switch. *J. Mol. Biol.*, **265**, 507-518.

Jain, V.K. (1996) Vectors with bidirectional reporter genes for studying divergent promoters. *Methods Enzymol.*, **273**, 59-71.

Julien, B. and Calendar, R. (1995) Purification and characterization of the bacteriophage P4 δ protein. *J. Bacteriol.*, **177**, 3743-3751.

Kapranov, P., Cawley, S.E., Drenkow, J., Bekiranov, S., Strausberg, R.L., Fodor, S.P.A. and Gingeras, T.R. (2002) Large-scale transcriptional activity in chromosomes 21 and 22. *Science*, **296**, 916-919.

Kimura, M. and Yamaguchi, I. (1998) Convergent transcription units and their promoters at both ends of *Pot2*, an inverted repeat transposon from the rice blast fungus. *J. Biochem.*, **124**, 268-273.

Kobayashi, M., Nagata, K. and Ishihama, A. (1990) Promoter selectivity of *Escherichia coli* RNA polymerase: effect of base substitutions in the promoter –35 region on promoter strength. *Nucl. Acid. Res.*, **18**, 7367-7372.

Koop, A.H., Hartley, M.E. and Bourgeois, S. (1987) A low-copy-number vector utilizing β-galactosidase for the analysis of gene control elements. *Gene*, **52**, 245-256.

Korzheva, N. and Mustaev, A. (2001) Transcription elongation complex: structure and function. *Curr. Opin. Microbiol.*, **4**, 119-125.

Kruglyak, S. and Tang, H. (2000) Regulation of adjacent yeast genes. *Trends in Genetics*, **16**, 109-111.

Leirmo, S., Harrison, C., Cayley, D.S., Burgess, R.R. and Record, M.T.J. (1987) Replacement of potassium chloride by potassium glutamate dramatically enhances protein-DNA interactions *in vitro*. *Biochemistry*, **26**, 2095-2101.

Levinthal, M. and Nikaido, H. (1969) Consequences of deletion mutations joining two operons of opposite polarity. *J. Mol. Biol.*, **42**, 511-520.

Liang, S.-T., Bipatnath, M., Xu, Y.-C., Chen, S.-L., Dennis, P., Ehrenberg, M. and Bremer, H. (1999) Activities of constitutive promoters in Eschericia coli. *J. Mol. Biol.*, **292**, 19-37.

Liang, S.-T., Dennis, P.P. and Bremer, H. (1998) Expression of *lacZ* from the promoter of the *Escherichia coli spc* operon cloned into vectors carrying the W205 *trp-lac* fusion. *J. Bacteriol.*, **180**, 6090-6100.

Linn, T. and Pierre, R. (1990) Improved vector system for constructing transcriptional fusions that ensure independent translation of *lacZ*. *J. Bacteriol.*, **172**, 1077-1084.

Lisser, S. and Margalit, H. (1993) Compilation of *E.coli* mRNA promoter sequences. *Nucl. Acid. Res.*, **21**, 1507-1516.

Liu, B. and Alberts, B.M. (1995) Head-on collision between a DNA replication apparatus and RNA polymerase transcription complex. *Science*, **267**, 1131-1137.

Liu, B., Wong, M.L. and Alberts, B.M. (1994) A transcribing RNA polymerase molecule survives DNA replication without aborting its growing RNA chain. *Proc. Natl. Acad. Sci. USA*, **91**, 10660-10664.

Liu, B., Wong, M.L., Tinker, R.L., Geiduschek, E.P. and Alberts, B.M. (1993) The DNA replication fork can pass RNA polymerase without displacing the nascent transcript. *Nature*, **366**, 33-39.

Liu, L.F. and Wang, J.C. (1987) Supercoiling of the DNA template during transcription. *Proc. Natl. Acad. Sci. USA.*, **84**, 7024-7027.

Lloubes, R., Baty, D. and Lazdunski, C. (1986) The promoters of the genes for colicin production, release and immunity in the ColA plasmid: effects of convergent transcription and Lex A protein. *Nucl. Acid. Res.*, **6**.

Mattick, J.S. (2001) Non-coding RNAs: the architects of eukaryotic complexity. *EMBO reports*, **21**, 986-991.

Miller, J.H. (1972) *Experiments in molecular genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Miller, J.H., Reznikoff, W., Silverstone, A.E., Ippen, K., Signer, E.R. and Beckwith, J.R. (1970) *J. Bacteriol.*, **104**, 1273-1279.

Minchiotti, G., Contursi, C., Graziani, F., Gargiulo, G. and Di Nocera, P.P. (1994) Expression of *Drosophilia melanogaster* F elements *in vivo*. *Mol. Gen. Genet.*, **245**, 152-159.

Mooney, R.A., Artsimovitch, I. and Landick, R. (1998) Information processing by RNA polymerase: recognition of regulatory signal during RNA chain elongation. *J. Bacteriol.*, **180**, 3265-3275.

Muskhelishvili, G., Buckle, M., Heumann, H., Kahmann, R. and Travers, A.A. (1997) FIS activates sequential steps during transcription initiation at a stable RNA promoter. *EMBO J.*, **16**, 3655-3665.

Nasser, W., Rochman, M. and Muskhelishvili, G. (2002) Transcriptional regulation of fis operon involves a module of coupled promoters. *EMBO J.*, **21**, 715-724.

Neufing, P.J. (1997) Establishment of lysogeny in coliphage 186. *Biochemistry*. University of Adelaide, Adelaide, SA.

Neufing, P.J., Shearwin, K.E. and Egan, J.B. (2001) Establishing lysogenic transcription in the temperate coliphage 186. *J. Bacteriol.*, **183**, 2376-2379.

Nicholson, A.W. (1999) *FEMS Microbiol. Rev.*, **23**, 371-390.

Nigumann, P., Redik, K., Maetlik, K. and Speek, M. (2002) Many human genes are transcribed from the antisense promoter of L1 retrotransposon. *Genomics*, **79**, 628-634.

Nomura, T., Aiba, H. and Ishihama, A. (1985a) Transcriptional organization of the convergent overlapping *dnaQ-rnh* genes of *Escherichia coli*. *J. Biol. Chem.*, **260**, 7122-7125.

Nomura, T., Fujita, N. and Ishihama, A. (1985b) Promoter selectivity of *E. coli* RNA polymerase: analysis of the promoter system of convergently transcribed *dnaQ-rnh* genes. *Nucl. Acid. Res.*, **13**, 7647-7663.

Opel, M.L., Arfin, S.M. and Hatfield, G.W. (2001) The effects of DNA supercoiling on the expression of operons of the *ilv* regulon of *Escherichia coli* suggest a

physiological rationale for divergently transcribed operons. *Mol. Microbiol.*, **39**, 1109-1115.

Owens, E.M. and Gussin, G.N. (1983) Differential binding of RNA polymerase to the pR_M and *pR* promoters of bacteriophage lambda. *Gene*, **23**, 157-166.

Padidam, M. and Cao, Y. (2001) Elimination of transcriptional interference between tandem genes in plant cells. *BioTechniques*, **31**, 328-334.

Portin, P. (2002) Historical development of the concept of the gene. *J. Med. Phil.*, **27**, 257-286.

Powell, B.S., Court, D.L., Nakamura, Y., Rivas, M.P. and Turnbough, C.L.j. (1994) Rapid confirmation of single copy lambda prophage integration by PCR. *Nucl. Acids Res.*, **22**, 5765-5766.

Prescott, E.M. and Proudfoot, N.J. (2002) Transcriptional collision between convergent genes in budding yeast. *Proc. Natl. Acad. Sci. USA.*, **99**, 8796-8801.

Ptashne, M. (1992) *A Genetic Switch: Gene Control and Phage λ*. Blackwell Scientific Publications & Cell Press, Cambridge, Massachusetts.

Record, T., Reznikoff, W., Craig, M., McQuade, K. and Schlax, P. (1996) Eschericia coli RNA polymerase (σ70), promoters, and the kinetics of the steps of transcription initiation. In Neidhardt, F., Curtiss, R., Ingraham, J., Lin, E., Low, K., Magasanik, B., Reznikoff, W., Riley, M., Schaechter, M. and Umbarger, H. (eds.), *Eschericia coli and Salmonella : Cellular and molecular biology*, Washington, D.C., pp. 792-821.

Reed, M.R. (1994) Apl: a multifunctional repressor and excisionase essential for prophage induction and temperate coliphage 186. *Biochemistry*. University of Adelaide, Adelaide, SA.

Reed, M.R., Shearwin, K.E., Pell, L.M. and Egan, J.B. (1997) The dual role of Apl in prophage induction of coliphage 186. *Mol Microbiol.*, **23**, 669-681.

Reynolds, R., Bermudez-Cruz, R.M. and Chamberlain, M.J. (1992) Parameters affecting transcription termination by *Eschericia coli* RNA polymerase. *J. Mol. Biol.*, **224**, 31-51.

Richardson, J.P. and Greenblatt, J. (1996) Control of RNA chain elongation and termination. In Neidhardt, F., Curtiss, R., Ingraham, J., Lin, E., Low, K., Magasanik, B., Reznikoff, W., Riley, M., Schaechter, M. and Umbarger, H. (eds.), *Eschericia coli and Salmonella : Cellular and molecular biology*, Washington, D.C., pp. 822-848.

Richet, E. and Raibaud, O. (1991) Supercoiling is essential for the foramtion and stability of the initiation complex at the divergent *malEp* and *malKp* promoters. *J. Mol. Biol.*, **218**, 529-542.

Rivetti, C., Guthold, M. and Bustamante, C. (1999) Wrapping of DNA around the *E. coli* RNA polymerase open complex. *The EMBO Journal*, **18**, 4464-4475.

Rojo, F. (2001) Mechanisms of transcriptional repression. *Curr. Opin. Microbiol.*, **4**, 145-151.

Rougeulle, C. and Heard, E. (2002) Antisense RNA in imprinting: spreading silence through Air. *Trends in Genetics*, **18**, 434-437.

Ryder, E.F. and Robakiewicz, P. (1998) Statistics for the molecular biologist: group comparisons. *Current Protocols in Molecular Biology*. John Wiley & Sons, Inc., Vol. 4, pp. A.3I.1-A.3I.22.

Ryu, S., Garges, S. and Adhya, S. (1994) An arcane role of DNA in transcription activitation. *Proc. Natl. Acad. Sci. USA*, **91**, 8582-8586.

Saha, S., Haggard-Ljungquist, E. and Nordström, K. (1987a) The *cox* protein of bacteriophage P2 inhibits the formation of the repressor protein and autoregulates the early operon. *EMBO J.*, **6**, 3191-3199.

Saha, S., Lundqvist, B. and Haggard-Ljungquist, E. (1987b) Autoregulation of bacteriophage P2 repressor. *EMBO J.*, **6**, 809-814.

Salgado, H., Santos-Zavaleta, A., Gama-Castro, S., Millan-Zarate, D., Diaz-Peredo, E., Sanchez-Solano, F., Perez-Rueda, E., Bonavides-Martinez, C. and Collado-Vides, J. (2001) RegulonDB (version 3.2): transcriptional regulation and operon organization in *Escherichia coli* K-12. *Nucl. Acid. Res.*, **29**, 72-74.

Sasse-Dwight, S. and Gralla, J.D. (1989) $KMnO_4$ as a probe for lac promoter DNA melting and mechanism in vivo. *J. Biol. Chem.*, **264**, 8074-8081.

Schlax, P., Capp, M.W. and Record, T. (1995) Inhibition of transcription initiation by *lac* repressor. *J. Mol. Biol.*, **245**, 331-350.

Schupp, J.M., Travis, S.E., Price, L.B., Shand, R.F. and Keim, P. (1995) Rapid bacterial permeabilization reagent useful for enzyme assays. *Biotechniques*, **19**, 18-20.

Sen, R., Nagai, H. and Shimamoto, N. (2000) Polymerase arrest at the λP_R promoter during transcription initiation. *J. Biol. Chem.*, **275**, 10899-10904.

Shanblatt, S.H. and Revzin, A. (1984) Kinetics of RNA polymerase-promoter complex formation: effects of nonspecific DNA-protein interactions. *Nucl. Acid. Res.*, **12**, 5287-5303.

Shearwin, K.E. and Egan, J.B. (1996) Purification and self-association equilibria of the lysis-lysogeny switch proteins of coliphage 186. *J. Biol. Chem.*, **271**, 11525-11531.

Shepherd, N., Dennis, P. and Bremer, H. (2001) Cytoplasmic RNA polymerase in *Eschericia coli*. *J. Bacteriol.*, **183**, 2527-2534.

Shibata, S. and Lee, J.T. (2003) Characterization and quantitation of differential Tsix transcripts: implications for Tsix function. *Hum. Mol. Genet.*, **12**, 125-136.

Shih, M.C. and Gussin, G.N. (1983) Mutations affecting two different steps in transcription initiation at the phage lambda P_{RM} promoter. *Proc. Natl. Acad. Sci. USA*, **80**, 496-500.

Shumann, G., Zundorf, I., Hofmann, J., Marschalek, R. and Dingermann, T. (1994) Internally located and oppositely oriented polymerase II promoters direct convergent transcription of a LINE-like retroelement, the *Dicotystelium* repetitive element, from *Dictyostelium discoideum. Mol. Cell. Biol.*, **14**, 3074-3084.

Simons, R.W., Hoopes, B.C., McClure, W.R. and Kleckner, N. (1983) Three promoters near the termini of IS10: pIN, pOUT, and pIII. *Cell*, **34**, 673-682.

Simons, R.W., Houman, F. and Kleckner, N. (1987) Improved single and multicopy *lac*-based cloning vectors for protein and operon fusions. *Gene*, **53**, 85-96.

Simons, R.W. and Kleckner, N. (1983) Translational control of IS10 transposition. *Cell*, **34**, 683-691.

Soukup, G.A. and Maher, L.M. (1998) Selection and characterization of RNAs that relieve transcriptional interference in *Esherichia coli. Nucl. Acid. Res.*, **26**, 2715-2722.

Spiegelman, W.G., Reichardt, L.F., Yaniv, M., Heinemann, S.F., Kaiser, A.D. and Eisen, H. (1972) Biderectional transcription and the regulation of phage λ repressor synthesis. *Proc. Natl. Acad. Sci. USA*, **69**, 3156-3160.

Stanssens, P., Remaut, E. and Fiers, W. (1986) Inefficient translation initiation causes premature transcription termination in the *lacZ* gene. *Cell*, **44**, 711-718.

Steiner, K. and Malke, H. (1997) Primary structure requirements for in vivo activity and bidirectional function of the transcription terminator shared by the oppositely oriented *skc/rel-orf1* genes of *Streptococcus equisimilis* H46A. *Mol. Gen. Genet.*, **255**, 611-618.

Sukhodolets, M.V., Cabrera, J.E., Zhi, H. and Jin, D.J. (2001) RapA, a bacterial homolog of SWI2/SNF2, stimulates RNA polymerase recycling in transcription. *Genes and Development*, **15**, 3330-3341.

Susa, M., Sen, R. and Shimamoto, N. (2002) Generality of the branched pathway in transcription initiation by *Escherichia coli* RNA polymerase. *J. Biol. Chem.*, **277**, 15407-15412.

Szoke, P.A., Allen, T.L. and deHaseth, P.L. (1987) Promoter recognition by *Escherichia coli* RNA polymerase: effects of base substitutions in the –10 and –35 regions. *Biochemistry*, **26**, 6188-6194.

Tjaden, B., Saxena, R.M., Stolyar, S., Haynor, D.R., Kolker, E. and Rosenow, C. (2002) Transcriptome analysis of *Escherichia coli* using high-density oligonucteotide probe analysis. *Nucl. Acid. Res.*, **30**, 3732-3738.

Ucker, D.S. and Yamamoto, K.R. (1984) Early events in the stimulation of mammary tumor virus RNA synthesis by glucocorticoids: Novel assays of transcription rates. *J. Biol. Chem.*, **259**, 7416-7420.

van Rijn, P.A., Goosen, N., Turk, S.C.H.J. and van de Putte, P. (1989) Regulation of phage Mu repressor transcription by IHF depends on the level of early transcription. *Nucl. Acid. Res.*, **17**, 10203-10212.

Vassylyev, D.G., Sekine, S., Laptenko, O., Lee, J., Vassylyev, M.N., Borukhov, S. and Yokoyama, S. (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature*, **417**, 712-719.

Vogel, U. and Jensen, K.F. (1994) Effects of guanosine 3',5'-bisphosphate (ppGpp) on rate of transcription elongation in isoleucine-starved Eschericia coli. *J. Biol. Chem.*, **269**, 16236-16241.

Ward, D.F. and Murray, N.E. (1979) Convergent transcription in bacteriophage λ: interference with gene expression. *J. Mol. Biol.*, **133**, 249-266.

Whitelaw, E. and Martin, D.I.K. (2001) Retrotransposons as epigenetic mediators of phenotypic variation in mammals. *Nature Genetics*, **27**, 361-365.

Woody, S.T., Fong, R.S. and Gussin, G.N. (1993) Effects of a single base-pair deletion in the bacteriophage λ P_{RM} promoter. *J. Mol. Biol.*, **229**, 37-51.

Wu, J., Grindlay, J., Bushel, P., Mendelsohn, L. and Allan, M. (1990) Negative regulation of the human ε-globin gene by transcriptional interference: role of an alu repetitive element. *Molecular and cellular biology*, **10**, 1209-1216.

Xu, J. and Koudelka, G.B. (2000) Mutually exclusive utilization of P_R and P_{RM} promoters in bacteriophage 434 O_R. *J. Bacteriol.*, **182**, 3165-3174.

Yamamoto, K., Ogasawara, H., Fujita, N., Utsumi, R. and Ishihama, A. (2002) Novel mode of transcription regulation of divergently overlapping promoters by PhoP, the regulator of two-component system sensing external magnesium availability. *Mol. Microbiol.*, **45**, 423-438.

Yang, M.-T., Scott II, H.B. and Gardner, J.F. (1995) Transcription termination at the *thr* attenuator. *J. Biol. Chem.*, **270**, 23330-23336.

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

Youderian, P., Bouvier, S. and Susskind, M.M. (1982) Sequence determinants of promoter activity. *Cell*, **30**, 843-853.

Young, B.A., Gruber, T.M. and Gross, C.A. (2002) Views of transcription initiation. *Cell*, **109**, 417-420.

Zhang, G., Campbell, E., Minakhin, L., Richter, C., Severinov, K. and Darst, S. (1999) Crystal structure of Thermus aquaticus core RNA polymerase at 3.3 A resolution. *Cell*, **98**, 811-824.

"A child of five could understand this. Fetch me a child of five!"

Graucho Marx