This is to certify that the

dissertation entitled

"Studies on the Interaction of E. coli RNA Polymerase with Lactose Promoter DNA"

presented by

Donald D. Lorimer

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Biochemistry

Arnold Revzin
Major professor

Date March 22, 1989

STUDIES ON THE INTERACTION OF *E. COLI* RNA POLYMERASE WITH LACTOSE PROMOTER DNA

By

Donald Duane Lorimer

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Biochemistry

1989

ABSTRACT

STUDIES ON THE INTERACTION OF E. COLI
RNA POLYMERASE WITH LACTOSE PROMOTER DNA

By

Donald Duane Lorimer

To gain a better understanding of gene expression in bacteria, I have performed experiments on the initiation of transcription at the E. coli lactose operon. The lac promoter region contains overlapping binding sites for RNA polymerase. Polymerase binds at the P1 site in the presence of cAMP and the catabolite activator protein (CAP) and initiates transcription at nucleotide +1; in the absence of CAP it binds at P2 and initiates at -22. Solutions of E. coli RNA polymerase and linear lac DNA fragments were probed by gel electrophoresis binding assays, runoff transcription experiments, and exonuclease III digestions. The results indicate that mixing RNA polymerase with wild type lac promoter fragments leads to formation of more than one kind of complex - an "open", transcriptionally competent complex at the P2 site, and a "closed" complex, sensitive to the competitor heparin, at the P1 promoter. Addition of CAP to

closed P1 complexes caused the conversion to open complexes at P1. Other data imply that there may be additional stable complexes as well (P3?). Under all conditions a large fraction of DNA fragments does not form heparin-insensitive complexes with RNA polymerase, perhaps because of interference from alternative binding sites at the promoter.

An attempt was made to prepare DNA molecules that would bind polymerase only at P1. A mutation in P2 at position -29 was introduced by site directed mutagenesis. A series of deletions with endpoints at +25, +19, +14, +1, and -7 was constructed and fused to the $P2^-$ mutation. Deletion of DNA downstream of -7 had no effect on the level of <u>in vitro</u> transcription. However, the deletion sequences did not bind polymerase better than wild type DNA, indicating that if P3 exists it is not downstream of -7. An alternate explanation based on differential polymerase activities is proposed. Polymerase preparations typically contain many (>50%) inactive molecules, which might form complexes that are incapable of transcription yet are stable enough to block productive complex formation at P1 and P2. Thus a better understanding of the structure of the enzyme itself may be needed to elucidate the initiation process.

For Nancy

ACKNOWLEDGMENTS

I would like to thank all those who have provided me with their time and support. First, thanks are due to Dr. Arnold Revzin for allowing me to pursue my research in a free and stimulating environment. Thanks also to my guidance committee - Drs. Veronica Maher, Jon Kaguni, Zachary Burton, and Larry Snyder for their assistance through the years. I would also like to acknowledge the early influence of Dr. Ron Davis and the late Dr. Barry Chelm. I also thank my labmates, past and present - Drs. Stephanie Shanblatt, Stephen Brooks, and Mark Garner as well as Jianli Cao, John Ceglarek, Lloyd LeCureux, Diane Cryderman, and Mark Sinton - for their friendship and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I.

INTRODUCTION AND LITERATURE REVIEW

INTRODUCTION

The elucidation of DNA structure by Watson and Crick in 1953 led to the central dogma of molecular genetics: genetic information flows from DNA to RNA to protein. The discovery of the different forms of RNA (transfer, ribosomal, and messenger) and the roles which they play allowed for the determination of the genetic code and clarification of the translation process. The discovery of a DNA-dependent RNA polymerase in 1959 led the way for the examination of the transcription of DNA into RNA.

It is convenient to divide the transcription process into three parts, initiation, elongation, and termination. While regulation can occur at any of these stages, it is now clear that control is exerted, in many instances, at the initiation step. Transcription is often regulated by proteins which act near the starting point for RNA synthesis. Positive regulators, or activator proteins, enhance initiation, negative regulators, or repressors, diminish initiation. The activity of the control proteins (or of RNA polymerase itself) can, in turn, be influenced by other, smaller molecules, whose availability is a function of the physical state of the cell or its environment. The

interplay of the components of the initiation process forms an intricate system which, when disturbed, can cause disfunction and disease.

This dissertation examines the interaction of RNA polymerase and an activator protein with the control region of the lactose operon of E. coli. Understanding the molecular mechanisms of initiation requires knowledge of where, and at what level, RNA polymerase binds to lac promoter DNA. These issues have been addressed by applying a variety of techniques to solutions containing purified proteins and wild type or mutant lac DNA fragments.

LITERATURE REVIEW

The Lactose Operon

A paradigm for the control of gene expression at the level of transcription is the lactose operon of the bacterium *Escherichia coli*. The *lac* operon is composed of four genes: the *lacI* gene, which encodes the *lac* repressor; the *lacZ* gene, which codes for ß-galactosidase; the *lacY* gene, which encodes the *lac* permease; and the *lacA* gene, which encodes an acetyltransferase enzyme. The *lac* system is controlled positively by the catabolite activator protein (CAP) and negatively by the *lac* repressor. The activities of repressor and CAP are regulated *in vivo* by allolactose and adenosine 3',5'-cyclic monophosphate (cAMP), respectively. A wealth of known mutations has enhanced study of the interaction of the components of this system at the molecular level. The ease with which the DNA and proteins can be isolated and manipulated has made the *lac* operon the most studied and best known genetic control system.

The description of the lactose operon by Jacob and Monod in 1961 (1) followed more than fifty years of physical

and genetic study on the ability of E. coli to adapt to its environment. Monod had shown earlier (2) that when E. coli cells were grown in a medium containing glucose and, for example, lactose, the glucose in the medium was preferentially utilized, leaving the lactose level relatively constant. When the glucose in the medium was spent, the lactose was consumed following the appearance of the lac enzymes. The model proposed by Monod accounted for the repression of the lac enzymes in the absence of lactose but did not explain the phenomenon of "catabolite repression", the low levels of expression when glucose was present in the medium (3). It was eventually discovered that cells grown without glucose had elevated intracellular levels of cAMP, and that addition of cAMP to a medium containing glucose and lactose led to synthesis of the lac enzymes (4,5). Subsequent experiments yielded a protein which could bind cAMP and stimulate lac expression (6); this protein was called the catabolite activator protein.

The final genetic feature of the system to be described was the nucleotide sequence which specifies the starting point of transcription. Jacob coined the word promoter to describe this necessary element (7). While Jacob did not know whether the promoter resided with the mRNA or the DNA, he was able to show that such a feature was necessary. We now know that the promoter is determined by the DNA sequence around the start point of transcription. The lac operon

displays two overlapping and mutually exclusive promoters *in vitro*: P1, from which CAP-dependent transcription begins at +1, and P2, which initiates efficiently at -22 in the absence of CAP (8,9). The *lac* P2 promoter does not appear to be active *in vivo*.

RNA Polymerase

The DNA-dependent RNA polymerase from *E. coli* is composed of 4 different subunits, α, β, β', and σ, in a ratio of 2:1:1:1 (10,11). The core polymerase, which contains the polymerization activity, resides in the $\alpha_2\beta\beta'$portion of the enzyme. The σ subunit is required for proper recognition of the promoter (10).

The exact amino sequences of the subunits have been deduced by sequence analysis of the genes, from which the molecular weights have been precisely determined: α = 36,511 D; β = 150,615 D; β' = 155,159 D; and σ = 70,262 D; thus, the mass of the holoenzyme, $E\sigma^{70}$, is 449,058 D (12).

During the initiation process at least two events involving polymerase must occur: (1) the enzyme recognizes the promoter sequence and binds to the double helical DNA, forming a "closed" complex and, (2) the enzyme "melts" into the two strands of DNA forming an "open" complex allowing for the copying of DNA from one strand into RNA. In fact, this is now known to be more complicated, with various

intermediates identified; these intermediates involve different conformations of the enzyme. During formation of the open complex a stretch of single stranded DNA is formed in the region from about -10 to +3 (13). After the enzyme has catalyzed polymerization of about 10 nucleotides of RNA, the sigma factor dissociates from the complex and elongation is catalyzed by core. After termination of transcription the core enzyme is released and may reassociate with a free σ factor (14).

An actively transcribing polymerase adds nucleotides to a growing RNA molecule at a rate of about 50 per sec at 37°C (15). Mechanistically, there must be two sites of action on the polymerase: a primer terminus site and an elongation site (16). In a reaction similar to ribosome movement in translation, a nucleotide enters the polymerase-DNA-RNA complex at the elongation site; there follows a nucleophilic attack on the α-phosphate of the incoming nucleotide by the 3′ hydroxyl at the end of the nascent RNA chain, leading to formation of the phosphodiester bond. After formation of the bond, the polymerase moves one base pair down the DNA so that the nucleotide added in the elongation site now resides in the primer terminus site and the process is repeated. At the start of this process both of the sites are initially empty and the enzyme must bind two free nucleoside triphosphates. The primer terminus site has a slight preference for purines, yet there appears to be no such

differential affinity at the elongation site (16). The elevated affinity for purines during initiation may be related to the fact that a majority of messenger RNAs start with an A or G.

Inhibitors of transcription have allowed study of the roles of the various subunits. Two well known classes of antibiotics act on the ß subunit: the rifamycins and the streptolydigins. Rifampicin, a commonly used rifamycin, acts at the initiation step by blocking the formation of the first phosphodiester bond (17). After the first two nucleotides are polymerized, rifampicin is not an effective inhibitor, implying that the antibiotic blocks entry of the initiating nucleotide. The streptolydigins block elongation presumably by binding to the elongation site (15). Polymerase mutations which confer resistance to these antibiotics reside in the ß subunit, indicating that this subunit may contain the nucleotide binding sites. Heparin, an initiation inhibitor used throughout this study, acts by competing with DNA for polymerase. The effect of heparin is largely electrostatic in nature. The ß′ subunit is the most basic of the subunits and is itself able to bind DNA, thus it is a likely candidate for the heparin binding site (18).

The α subunit exists in solution as a dimer. During infection of E. coli by phage T4, the α subunit is ADP-ribosylated (19). The modified polymerase has a reduced

affinity for promoters indicating an involvement of the α subunit in the DNA binding function of the enzyme.

Several σ factors are known in E. coli (for a review of σ factors see (20)). The form of the enzyme which acts at most promoters is $E\sigma^{70}$ (the superscript 70 refers to the molecular weight of sigma of 70,000 D). Two other σs have been described; σ^{32} which operates at heat shock promoters (21), and ntrA (σ^{54}) which is required for some nitrogen-regulated genes (22). These alternate sigmas confer specificity to promoters with sequences different from those used by $E\sigma^{70}$. The presence of these alternate σs is determined by the physiological state of the cell, thus they are found only in certain circumstances.

Sigma factors confer promoter specificity on the polymerase. For example, core polymerase can bind (at 0.1 M KCl) to any DNA sequence, forming a nonspecific complex with a binding constant of $\sim 10^{11}$ M^{-1}. Holoenzyme binds nonspecifically with a smaller binding constant of $\sim 10^{7}$ M^{-1}, yet binds specifically to a promoter with a constant of $\sim 10^{14}$ M^{-1}. Thus, σ^{70} acts to destabilize nonspecific binding by holoenzyme, while stabilizing specific complexes.

The Catabolite Activator Protein

Stimulation of transcription at lac and many other operons is achieved by a complex between the catabolite

activator protein and cAMP. Purified CAP has been shown to enhance lac transcription in vitro approximately 30 fold (23). The stimulatory action of CAP is strictly cAMP dependent.

CAP is a dimer of identical subunits each with a molecular weight of 23,619 D. It is a basic protein with a pI of 9.1. Each subunit has a cAMP binding site associated with it, yet it appears that only one cAMP is required for proper stimulation of transcription in vitro (24).

Initiation from the CAP sensitive P1 promoter site at both the lac and gal operons begins by definition at +1. Stoichiometry studies revealed that one CAP molecule is sufficient to stimulate lac expression, yet two CAPs appear to function at gal (24,25). It has been proposed that CAP acts at these promoters in two ways: by enhancing open complex formation, and by blocking polymerase from binding at other sites (9,26). The mechanism by which CAP stimulates open complex formation is as yet unclear. The known binding sites for CAP are located at different distances from the polymerase binding sites. At lac, for example, CAP binds between bases -55 and -70 relative to the P1 start site at +1 (27), while at gal the primary CAP site is between -34 and -49 (28,29), about 20 base pairs closer to the P1 initiation site. The disparity in distance between these two binding sites indicates that the action of CAP at lac and gal is possibly quite different.

As is the case for RNA polymerase and other genome regulatory proteins, CAP has a general affinity for DNA in addition to binding tightly at specific sites. CAP is known to cause alterations in DNA structure. Salemme observed by electron microscopy a general shortening of DNA upon nonspecific binding of CAP (30). He proposed that CAP could work at promoters by forming a solenoidal structure which takes up the excess DNA between RNA polymerase and CAP allowing the two proteins to interact over a long distance. Recently, gel mobility assays have shown that CAP induces a bend of about 135° in lac DNA (31,32). Models of polymerase and CAP binding at lac indicate that such a bend could bring the two proteins in contact. These models assume that a direct interaction between polymerase and CAP is necessary for stimulation, though it is not known with certainty whether that is the case. The formation of a bend or loop has been proposed as part of the mechanism of action of the lac and gal repressors and may play a role at other operons as well.

CAP also acts at lac and gal by blocking alternate promoters. Each of these operons has a second polymerase binding site, called P2. At lac, P2 initiation occurs from -22 (8,9) while initiation from gal P2 occurs at -5 (33). The P2 promoters are not stimulated by CAP; on the contrary, the binding sites for polymerase and CAP overlap in such a way as to be mutually exclusive. Thus, a CAP molecule bound

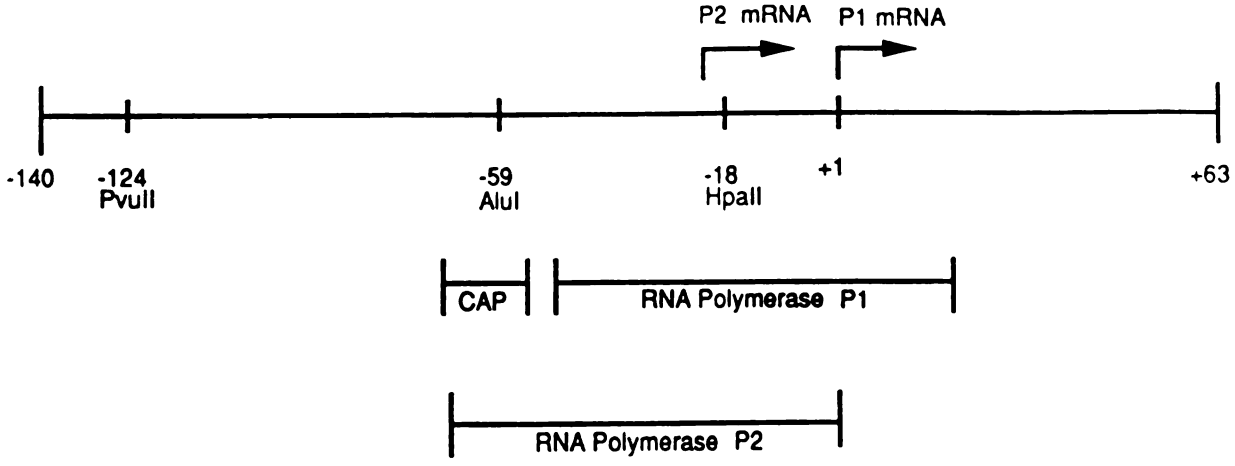


Figure 1. The lac promoter region showing the relevant start sites for transcription and the binding sites for CAP and RNA polymerase.

at its site will necessarily direct polymerase away from P2 and perhaps channel it to the P1 site. Since the gal P2 promoter is active in vivo, the CAP/P2 competition can play a regulatory role. However, lac P2 is apparently inactive in vivo so its physiological role is unclear. In principle, if CAP prevents polymerase binding to a nonproductive site it could lead to stimulation of transcription without any direct contact with polymerase. In Chapter II we show that simple P2 occlusion does not account for CAP activity at lac in vitro. RNA polymerase forms a stable, long lived closed complex at P1 which requires CAP binding for activation. This observation implies that CAP directly stimulates lac transcription, probably via contacts with polymerase.

CAP acts at other operons as well. Two other well studied CAP-stimulated promoters are malT and araBAD. The malT promoter region has a single polymerase binding site with very poor sequence homology to normal E. coli promoters (34). The CAP site is centered about 70 base pairs upstream of the initiation point. Assays of in vitro activity indicate that open complex formation is not greatly enhanced by CAP (35). Kinetic analyses revealed that CAP affects steps after the binding of polymerase, perhaps enhancing polymerase escape from the malT promoter. Stimulation of the araBAD operon requires another protein, araC, which works with CAP to achieve maximum stimulation (36). In the absence of the inducer, arabinose, the araC protein seems to block entry of polymerase to the promoter (37). However, in the presence of arabinose, araC bound at the same site from which repression occurs is now able to stimulate transcription (38). A change in conformation of the araC protein accompanies the change in activity. The requirement for stimulation by araC is absolute, but CAP provides an additional ten-fold stimulation. CAP also activates an adjacent promoter P_c, from which the structural gene for araC is transcribed. CAP acts at both P_{BAD} and P_c from the same site (39).

Thus the molecular details whereby CAP stimulates transcription at these operons are quite different than those at lac or gal. While the exact mechanism of CAP

function is not known, it is becoming clear that the process is more complicated than previously thought. It is interesting to study these systems further, looking for similarities and differences.

Promoters

The sequence of DNA which specifies the start point of transcription is called the promoter. Analysis of more than 200 *E. coli* promoters has revealed two conserved sequences centered around -10 and -35 (40,41). The -10 region, with a consensus sequence of TATAAT, is separated from the -35 consensus sequence of TTGACA by 17±1 base pairs. Several promoters have exact agreement with one or the other sequences, but no naturally occurring promoter has been found with perfect homology in both regions. The artificial promoter, *tac*, which comprises the *trp* -35 region and the mutant, *lac*UV5 -10 sequence, does have both consensus sequences perfectly matched, with a 16 base pair spacer (42). The activity of the *tac* promoter is about 11 times higher than the UV5 promoter and about 3 times greater than the *trp* promoter. *In vivo*, the *tac* promoter is difficult to regulate. The promoter constructs used contain the *lac* operator; however, the ability of the *lac* repressor to regulate transcription is partially defeated by the increased strength of the promoter. Plasmid vectors which

utilize this promoter are known to be "leaky", that is, under conditions when repression should occur, some expression is observed. It is also possible to imagine a promoter so strong as to prevent escape of the polymerase due to very tight binding. Thus, it seems that when promoter strength is optimized for biological function *in vivo*, it is not necessarily maximized.

The role of the two consensus sequences beyond promoter recognition is unclear. Promoters with poor homology to the -35 consensus, especially in the TTG portion, are often positively controlled by other factors (43). Alternatively, the heat shock promoters have a -35 sequence similar to the σ^{70} promoters yet have very different -10 sequences, which implies that sigma factor may guide polymerase to the promoter by recognition of the -10 region (44).

The spacer DNA appears to be important also. No mutations have been found in the spacer, thus it is likely that no specific contacts occur in the spacer other than casual ones near the consensus sequences. However, when the spacer is replaced by DNA with a conformation different from B-form, promoter strength can decrease several fold (45). In addition, single base deletions or insertions in the spacer of promoters can cause severe alteration of promoter strength (46). Thus, the spacer may be thought of as correctly positioning the -10 and -35 sequences for proper polymerase binding.

The roles of the upstream (to the left of -35) and downstream (to the right of -10) sequences are not well understood. It has been shown that DNA upstream of the -35 sequence can be deleted and replaced by heterologous DNA without affecting the strength of the lacUV5 promoter *in vitro* and *in vivo* (47,48).

Bujard and coworkers have shown that the strength of some promoters is influenced by downstream sequences (49,50). By systematically changing the sequence between +1 and +20 in promoters from bacteriophages T5, T7, and lambda, these authors observed a more than 10-fold variation in promoter strength. In addition, their studies identified the potentially important sequence TTGA from +7 to +10 followed by a string of purines from around +11 to +20 in the early promoters of phage T5. Specific replacement of the TTGA sequence resulted in a 30% reduction in promoter utilization *in vivo*. Several mutations between +1 and +10 have been isolated in the lac promoter region (8). These mutations apparently have no effect on the utilization of the P1 promoter; rather, they influence the repression of the operon, as expected, since they fall in the operator sequence. Some of the mutants also activate new promoters which initiate at sites away from +1.

In bacteria the first nucleotide in most RNA molecules is an A or a G (41). As mentioned above, the polymerase template site has a slight preferential affinity for

purines, thus, initiation with an A or a G may be part of an optimization process for *in vivo* promoter function. A systematic investigation of the effect of the initiating nucleotide on the transcription process has not been performed. Plasmid vectors used for expression of foreign genes in bacteria generally have a common start point followed by a polylinker for convenient insertion of various DNA sequences. The mRNAs produced from these vectors have some RNA at their 5'-ends which is derived from the vector.

A renewed interest in RNA structure-function relationships has made necessary isolation of sizable quantities of RNA molecules with precisely defined sequences for physical analysis. For example, many mammalian mitochondrial mRNAs appear to start with an AUG. It would be of value to prepare such mRNAs, *in vitro*, for ribosome binding studies. Systems for synthesizing RNA *in vitro* have been devised, based on the strong promoters from phages SP6 and T7. These promoters are transcribed by their corresponding polymerases with very high specificity. However the polymerases have strong requirements for specific sequences at and near to +1 (51-53). The SP6 enzyme for instance, will not transcribe any RNA that does not begin with a G (52). In Chapter III we show that such stringent requirements are not the case at the *lac* promoter. Thus a mutated *lac*UV5 promoter, in conjunction with *E. coli* RNA polymerase can in principle transcribe efficiently DNAs

having any desired sequence near +1. The addition of another tool for *in vitro* RNA synthesis will prove very useful in some situations.

Methods of Analysis

Analysis of promoter-polymerase interactions can be done using a variety of physical methods: (1) gel mobility shift assays, (2) runoff transcription assays, and (3) protection of the DNA from nuclease digestion or chemical degradation, also called "footprinting". Each of these techniques gives different information about the system; when used together, much insight can be obtained.

The gel mobility shift assay provides direct information on the level of binding of a DNA-binding protein to its target DNA molecule (54). Briefly, a mixture of DNA and protein is layered atop a polyacrylamide gel and electrophoresis performed. After a sufficient length of time, the gel is developed and the concentration of free DNA determined. By difference, the amount of DNA in complex can be calculated. If the association between the protein and the DNA is sufficiently strong that negligible dissociation occurs during the electrophoresis run, then the level of the bound form can be determined directly. From the concentration of bound and free DNA, the ratio of bound to total (bound+free) can be made; the resulting value is the

fractional binding or promoter occupancy. Kinetic data in the form of association or dissociation rate constants can be determined by measuring complex formation or decay as a function of time.

In many natural promoter regions there is more than one polymerase binding site. In these cases the mobility shift assay may not distinguish between polymerases bound at one or the other sites. Runoff transcription assays can be used to identify and quantify stable binding at different binding sites. To do this, binding reactions are prepared and incubated at the desired temperature. A mixture of nucleoside triphosphates is then added and transcription is allowed to occur. When the polymerase reaches the end of the DNA fragment it "runs off" or dissociates from the DNA, releasing the nascent transcript. The length of the transcript is equivalent to the distance from the start point of transcription to the end of the fragment. Polymerase initiating at distinct sites will give transcripts separable by electrophoresis in denaturing gels. If transcription is performed with a radiolabelled nucleotide, the bands of RNA on the gel can be visualized by autoradiography. If quantitation is desired, the transcripts can be excised from the gel and the radioactivity determined by scintillation counting.

A method similar to runoff transcription is the abortive initiation assay (55). In this technique a short

RNA primer which is equivalent to the first few nucleotides of the transcript is incubated with the DNA and polymerase. Extension of the RNA primer in the complex is catalyzed by bound polymerase; if only the next single nucleotide is added, the elongated primer is released and the cycle is repeated. The products can be separated by thin layer chromatography and quantified. Data obtained in this manner complement those from runoff transcription assays.

The combination of gel shift and transcription assays can be used to determine stoichiometry and kinetic parameters necessary to describe the interaction of RNA polymerase with promoters. To characterize the complexes further, footprinting with nucleases or chemicals is used to identify the regions of DNA where proteins are bound. In this approach, a protein is mixed with DNA (labelled at one end) and, following equilibration, complexes are exposed to a nuclease such as DNaseI (a nonspecific endonuclease) or chemical reagents such as iron(II)-EDTA (56,57). The reaction is allowed to proceed only to a limited extent. The presence of a bound protein will cause some part of the DNA to be resistant to degradation relative to a naked DNA control; the digested DNA can be displayed on a denaturing polyacrylamide gel for further analysis. If a high proportion of DNA molecules in the binding reactions are free of bound protein (e.g., for some reason the promoter occupancy is low) it may be difficult to distinguish the

protected regions against the background from the free DNA. In this situation the footprinted complexes can be separated from the free DNA by gel shift electrophoresis, excised from the gel and eluted prior to analysis on the denaturing gel. Alternatively, digestion by exonuclease III (a 3′ to 5′ processive exonuclease) can be used to determine the boundaries of a bound protein. Because the digestion is generally done to completion, there is little interference from the free DNA. Restriction endonucleases can also be used to probe DNA-protein interactions. A protein which, when bound, covers a restriction site will interfere with the action of the enzyme giving rise to incomplete digestion.

Kinetic analyses and equilibrium studies are usually done at an excess of polymerase to promoter fragment; this simplifies the rate equations, and can eliminate effects of any inactive molecules in the preparation. Studies done at equilibrium, such as footprints, are also done at polymerase excess to ensure complete binding.

A minimal model for polymerase binding to a promoter to form an open complex was proposed many years ago (58).

$$R + P \underset{k_{-1}}{\overset{k_1}{\Longleftrightarrow}} (RP)_c \underset{k_{-2}}{\overset{k_2}{\Longleftrightarrow}} (RP)_o$$

where

R = RNA polymerase

P = promoter DNA

$(RP)_c$ = "closed" complex of polymerase and DNA

$(RP)_o$ = "open" complex of polymerase and DNA

$k_1/k_{-1} = K_B$ = equilibrium constant for polymerase binding to form a closed complex

$k_2/k_{-2} = K_2$ = equilibrium constant for isomerization step.

Later refinements of this model have included additional steps involving conformational changes of the protein; for this discussion only the simple, two step model will be considered.

The rate equation to be solved is:

$$\frac{d(RP)_o}{dt} = k_2[(RP)_c] - k_{-2}[(RP)_o] \cdot$$

The general solution to this differential equation is (59):

$$[(RP)_o] = [P_t](1-\exp(-k_{obs}t))$$

where P_t = total promoter concentration ($[P] + [(RP)_c] + [(RP)_o]$). If one assumes that k_{-2} is $<< k_2$ (i.e., that isomerization to open complexes is essentially irreversible) then

$$k_{obs} = \frac{k_1[R]k_2}{k_1[R]+k_{-1}+k_2}.$$

The reciprocal of k_{obs} provides a measure of the time required for open complex formation, τ_{obs}:

$$\tau_{obs} = \frac{1}{k_{obs}} = \frac{1}{k_2} + \frac{k_{-1} + k_2}{k_1[R]k_2}.$$

Thus a plot of τ_{obs} vs 1/[R] will yield $1/k_2$ at the y-intercept, and if one assumes that $k_{-1}>>k_2$ (i.e. a rapid equilibrium for closed complex formation), the slope becomes

$$\text{slope} = \frac{k_{-1}}{k_1} \cdot \frac{1}{k_2} = \frac{1}{K_B} \cdot \frac{1}{k_2}.$$

The evaluation of τ_{obs} at various polymerase concentrations can be performed as follows. Proteins and DNA are mixed, and the extent of open complex formation determined (by gel retardation or runoff transcription) as a function of time after quenching with heparin; for example, samples might be taken at 30 sec intervals after the reaction is initiated. The heparin removes nonspecifically bound polymerase molecules; it also prevents reinitiation from occuring in transcription assays. The ratio, F, of $[(RP)_o]_t$ to $[(RP)_o]_\infty$ is determined for each time point

sampled. (The value of $[(RP)_o]_\infty$ is empirically determined; usually it can be approximated as the maximum level of open complex detected after a sufficiently long wait.) If pseudo-first order kinetics are obeyed (i.e. polymerase excess), then, a plot of ln(1-F) vs time will have a slope equivalent to $-k_{obs} \equiv -1/\tau_{obs}$. Similar analyses are performed at various polymerase concentrations and a secondary plot of τ_{obs} vs 1/[R] is made, from which K_B and k_2 are derived.

Dissociation kinetics are evaluated by determining the level of stable complexes as a function of time after the addition of heparin. A plot of -ln(F) vs time will have a slope equal to k_{-2}. Here, F is the fraction of open complex remaining at time t.

Taken together, τ analyses and dissociation kinetics can shed light on the interaction of RNA polymerase with promoter DNA by revealing possible reaction mechanisms and suggesting potential intermediates. In principle, the role of activators such as CAP can be assayed by these methods as well. However as the number of intermediates grows the analyses become quite difficult and accuracy is reduced.

CHAPTER II.

SOLUTIONS OF RNA POLYMERASE PLUS LINEAR WILD TYPE *E. COLI lac* DNA FRAGMENTS CONTAIN A MIXTURE OF STABLE P1 AND P2 PROMOTER COMPLEXES

CHAPTER II

INTRODUCTION

Studies of the molecular events involved in the regulation of transcription have shown these processes to be more complicated than originally envisioned. Many operons have multiple RNA polymerase binding sites which are overlapping and mutually exclusive. The interaction of polymerase with overlapping promoters may provide an extra level of control. Also, the question of how CAP acts at catabolite sensitive operons is not fully resolved.

The ready availability of purified proteins and DNA fragments makes feasible the physical study of these systems. In light of the fact that polymerase may bind to more than one site, I have undertaken to study solutions of RNA polymerase, cAMP/CAP, and *lac* promoter fragments. In this chapter I report a number of unusual findings.

MATERIALS AND METHODS

Proteins. RNA polymerase holoenzyme (60,61) and CAP (62) were prepared as previously described (54). The polymerase was 45% active in specific binding to the *lac*UV5 promoter using a gel binding assay (54) and was about 30%

active when analyzed by the method of Chamberlin et al. (63). Based on the amount required to stimulate fully lac P1 transcription in vitro, the CAP was estimated to be 25% active. Protein concentrations were determined spectrophotometrically (see reference 54).

lac Promoter Fragments. All fragments are numbered relative to the P1 mRNA start site designated +1. DNA fragments 211 bp long (-140 to +63 plus EcoRI linkers) containing the wild type or mutant lacUV5 promoter were purified from recombinant pMB9 plasmids which were generously provided by Forrest Fuller. The UV5 sequence also has the L8 mutation, which reduces affinity of CAP for its site. Restriction fragments 130 bp in length which contained only the P1 promoter were generated by partial digestion of the 211 bp fragment with AluI. The fragment extending from -59 to the EcoRI linker at +67 was eluted from a polyacrylamide gel, a HindIII linker was ligated to the blunt -59 end, and the fragment was cloned between the HindIII and EcoRI sites in pBR322.

Restriction fragments were purified by separation in polyacrylamide gels followed by electroelution and chromatography through DEAE Sephadex (64). DNA fragments used in binding gels and exonuclease III protection experiments were 5'-end labelled with [γ-^{32}P] ATP and T4 polynucleotide kinase following treatment with calf intestinal phosphatase (64). Concentrations of DNA were

determined spectrophotometrically using ϵ_{260}=13,000 $(M,bp)^{-1}$ (65) and are expressed as molar in promoter fragments.

Binding Reactions. All binding reactions were performed in a buffer which contained 20 mM Tris-HCl (pH 7.9 at 25°C), 10 mM $MgCl_2$, 0.1 mM EDTA, 0.1 mM DTT and 100 mM KCl. Reactions contained 50 nM DNA fragments, 250 nM RNA polymerase and where applicable 200 nM CAP and 20 μM cAMP. Polymerase was incubated with DNA fragments, with or without CAP, for 30 min at 37°C. If CAP was added before RNA polymerase it was allowed to react for 30 min.

Binding Assays by Gel Retardation. Protein-DNA complexes were formed as described above using radiolabelled fragments. Following incubation for the desired length of time, heparin was added to 100 μg/ml. Thirty sec later 1/10 volume of 25% Ficoll, 0.1% bromphenol blue and 0.1% xylene cyanol was added and the reactions were immediately loaded onto a 0.1 x 10 cm vertical 5% polyacrylamide gel (54) in TBE buffer (90 mM Tris, 90 mM H_3BO_3, 2.5 mM EDTA). The gel was electrophoresed at 15 V/cm for 45 min, dried onto Whatman 3MM paper and autoradiographed with Kodak XAR-5 film.

Transcription Assays. Transcription was assayed in 25 μl binding reactions as described above. After the proteins were incubated for the appropriate times, heparin was added to a final concentration of 100 μg/ml followed 30 sec later by the addition of ATP, GTP and CTP to final concentrations

of 200 μM each and [α-^{32}P] UTP (10 μCi, 3000 Ci/mmole) plus cold UTP to 50 μM. After 15 min, the reactions were terminated by adding an equal volume of stop buffer (0.02 M EDTA, 0.3 M sodium acetate, 100 μg/ml tRNA), immediately extracted with phenol, and ethanol precipitated. The pellets were dissolved in 5 M urea, 0.02 M Tris-HCl (pH 8.0), 0.1 mM EDTA, heated to 95°C for 2 min, quick chilled on ice and electrophoresed at 50 watts on 8%, 0.4 x 34 x 40 cm polyacrylamide, 7 M urea gels in TBE buffer (66). Transcripts were visualized by autoradiography and quantified by scintillation counting of excised bands. In some experiments, radiolabelled DNA was added to the stop buffer to serve as an internal standard during precipitation of the RNA products.

Exonuclease III Protection. The 211 base pair and the 130 base pair restriction fragments were labelled at one or the other 5′ ends by standard procedures (64). Binding reactions were performed as described above (using 50 μl reaction volumes). Heparin was added in some cases to a final concentration of 100 μg/ml followed 30 sec later by the addition of 10-50 units of exonuclease III. After further incubation for 30 min at 37°C the reactions were terminated by the addition of an equal volume of stop buffer, immediately phenol extracted and ethanol precipitated. Products were dissolved and electrophoresed as for transcripts. Lengths of band sizes were estimated by

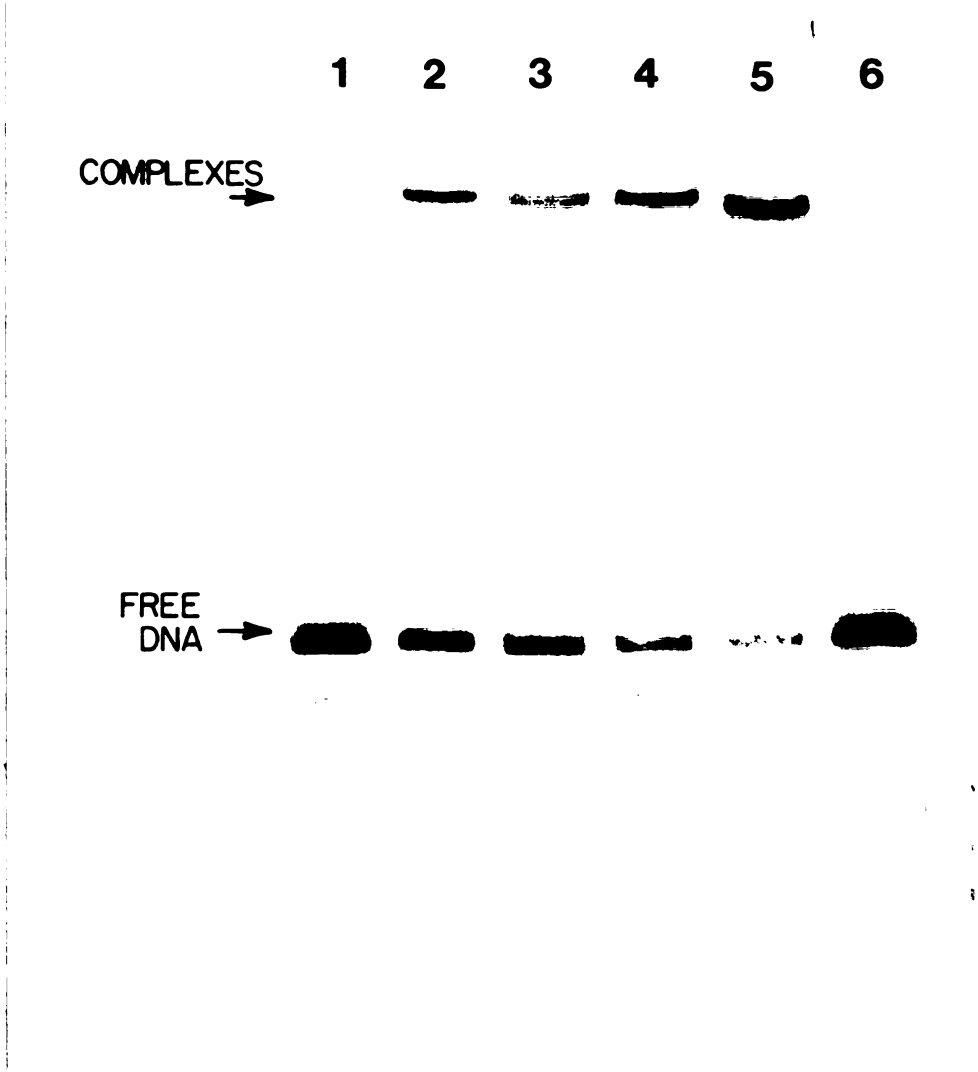
their migration relative to size markers of [^{32}P] end-labelled HpaII fragments of pBR322.

RESULTS

Binding Gels. A gel electrophoresis technique (54) was used to assay the level of stable, heparin-insensitive complexes prepared under these conditions (Figure 2). Incubating only RNA polymerase with DNA fragments containing the wild type lac promoter region resulted in about 53% of the DNA being found in open complexes in this experiment. If CAP was added first, then polymerase, the degree of open complex formation was about 45%; adding polymerase, then CAP, led to about 63% complexes. In all cases, 30 min incubation times were used after each protein addition to insure that these relatively slow reactions were complete. These findings are consistent with data in the literature (67,68) which indicate that only about 50% or less promoter occupancy is achieved on linear wild type lac DNA fragments (although a much higher fraction of DNA is found in complex if the mutant lacUV5 promoter is used [Figure 2, lane 5]).

Transcription Assays. Runoff transcripts enable one to distinguish the types of stable complexes formed under different experimental regimes. The lac DNA fragment used should yield a mRNA of about 69 nucleotides if transcription begins at +1 (P1) or about 91 nucleotides if initiation is at -22 (P2) (8,9). These are readily distinguished on a

Figure 2. Gel binding assay of protein-DNA complexes formed at *lac* promoters. Lane 1, wild type *lac* DNA alone; lane 2, wild type *lac* DNA plus RNA polymerase; lane 3, wild type *lac* DNA plus CAP, then RNA polymerase; lane 4, wild type *lac* DNA plus RNA polymerase, then CAP; lane 5, *lac*UV5 DNA plus RNA polymerase; lane 6, *lac*UV5 DNA alone.

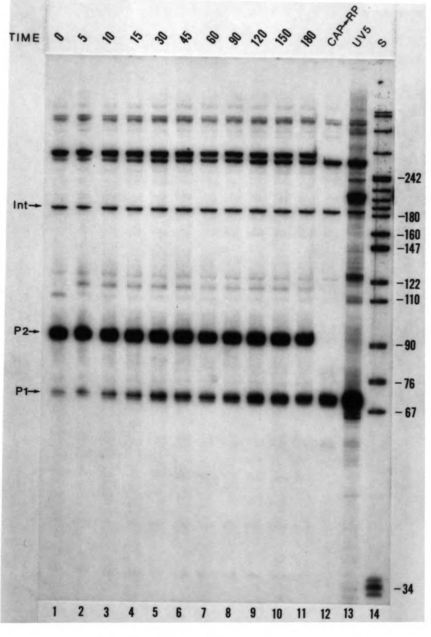

1
2
3
4
5
6
COMPLEXES
FREE
DNA

denaturing gel (Figure 3), from which it is seen that virtually all transcripts are from P2 in the absence of CAP (lane 1 and data not shown), while only the P1 start site is used if CAP is added first (lane 12). However, if polymerase is added first, then CAP, both transcripts are made. Similar results have been presented by Meiklejohn and Gralla (26).

The kinetics of formation of transcriptionally competent complexes at P1 and P2 can be followed by mixing DNA and proteins, quenching the reaction with heparin at various times, then quantifying the numbers of transcripts at each point. To determine dissociation rates, complexes are incubated in the presence of heparin for various lengths of time, then nucleotides are added and runoff transcripts quantified. Any polymerase molecules which dissociate are sequestered by the heparin and cannot rebind. The same sorts of kinetics data can be derived using gel electrophoresis to determine the number of heparin-resistant complexes. Binding gels do not distinguish between open complexes at P1 and P2, but for the appropriate cases the gel and transcription experiments give the same results, indicating that the level of transcription provides a measure of the concentration of complexes (i.e. of promoter occupancy).

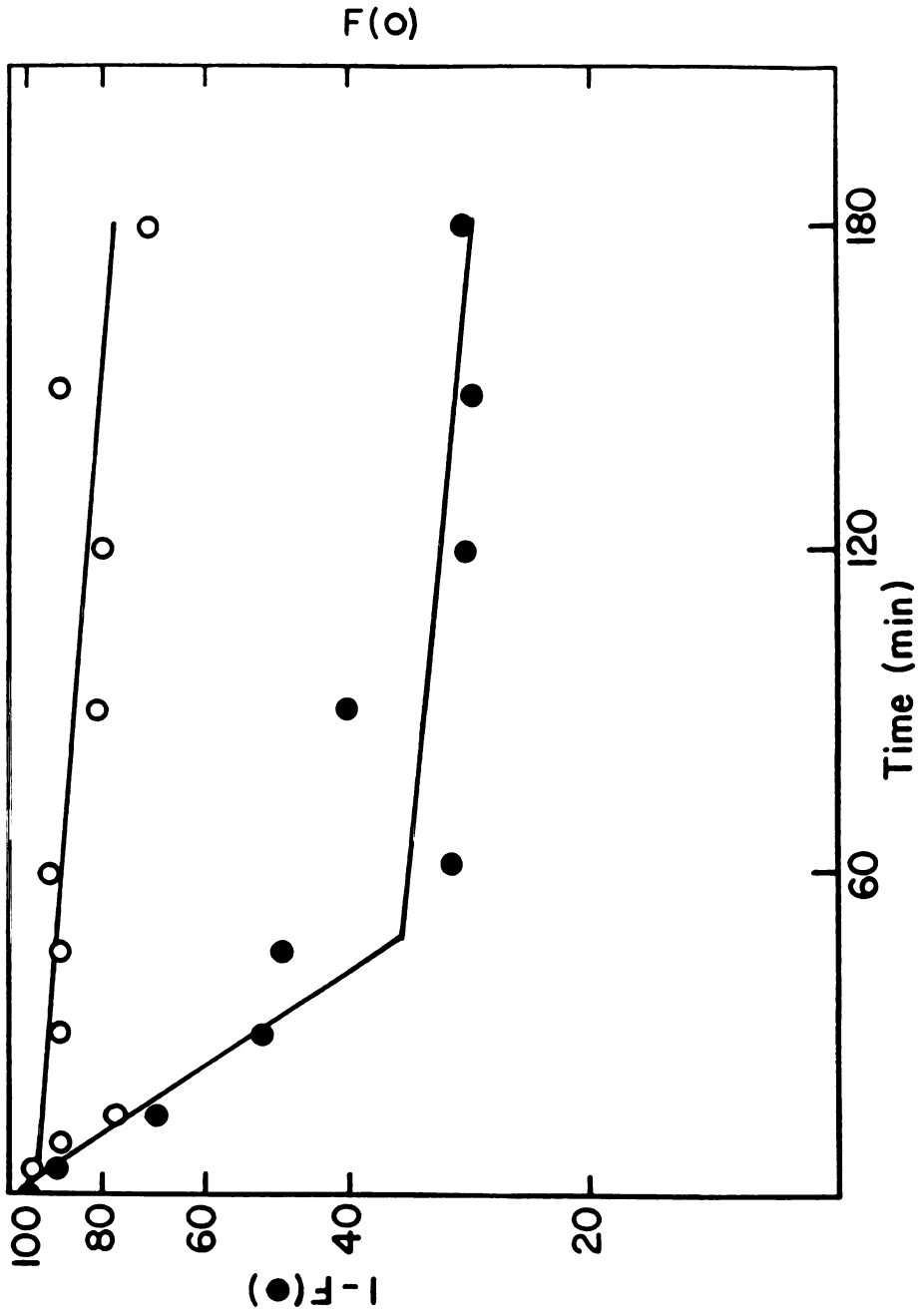
In the transcription experiment for which data are depicted in Figure 3, formation of P1 complexes was measured

Figure 3. Transcription in vitro from lac promoters. Lanes 1 through 11 show transcripts from solutions in which wild type lac DNA and RNA polymerase were incubated at 37°C for 30 min prior to the addition of CAP. CAP was allowed to react for 0 to 180 min, as indicated, before the addition of heparin and nucleoside triphosphates. For lane 12, CAP was incubated with wild type lac DNA for 30 min prior to the addition of RNA polymerase and another 30 min incubation. Lane 13 shows transcription from the mutant lacUV5 promoter in the absence of CAP. Size standards of pBR322 digested with HpaII are displayed in Lane 14. "Int" refers to a [^{32}P]-labeled 188 base pair DNA fragment used as an internal standard.


TIME
0
5
10
15
30
45
60
90
120
150
180
CAP-RP
UV5
S
Int
P2
P1
242
180
160
147
122
110
90
76
67
34
1
2
3
4
5
6
7
8
9
10
11
12
13
14

after addition of CAP to polymerase-promoter complexes. Analysis of the data (Figure 4) reveals the striking result that, while the preformed P2 complexes dissociate rather slowly, transcription from P1 increases relatively rapidly at first ($t_{1/2}$~7.5 min), then slows. The rate of P1 formation at long times parallels the rate of P2 dissociation. Dissociation results for solutions in which CAP was incubated with polymerase-promoter complexes for 30 min before heparin was added are shown in Figure 5. Off-rates from these data are $t_{1/2}$~105 min for P1, 102 min for P2. If CAP is added first, so that only P1 open complexes are formed, the same half time for association (8 min) is seen; these P1 complexes dissociate with $t_{1/2}$~92 min (data not shown). Omitting CAP results in stable complexes only at P2. These also dissociate with a half-life of 90 min. Taking into account the difference in numbers of uridine residues in P1 and P2 mRNAs, data derived from the experiments shown in Figures 2 and 3 can be converted to yield the actual numbers of transcripts synthesized from each start point. The faster component of the P1 curve in Figure 4 is complete after about 45 min. For an initial total input DNA level of 1.25 pmoles, we find 0.33 pmoles of P1 transcript at this time. At time zero, 0.52 pmoles of P2 transcript are made, while 0.44 pmoles of P2 mRNA are synthesized after 45 min of incubation in the presence of CAP. Thus at most a small fraction (0.08 pmoles) of the

Figure 4. First order kinetics analysis of lac promoter-RNA polymerase complexes. Transcripts from the gel shown in Figure 3 were quantified by scintillation counting. Data for P2 dissociation are plotted as ln F vs. time, where F is the number of transcripts at the indicated time divided by the number of transcripts at time zero. Data for P1 open complex formation were corrected for the small amount of P1 transcription found in the absence of CAP, then were plotted as ln(1-F) vs. time. Here F is the ratio of the number of transcripts at the indicated time to an arbitrary limiting number of transcripts; this arbitrary quantity is used because the level of P1 transcription at infinite time is not known (since at long times formation of complexes at P1 is tied to dissociation of enzyme from P2). Comparison of the rate of P1 complex formation with the dissociation rate from P2 cannot be made using fractions of reaction, but must involve actual numbers of transcripts as described in the text.


F(o)
1-F(●)
100
80
60
40
20
60
120
180
Time (min)

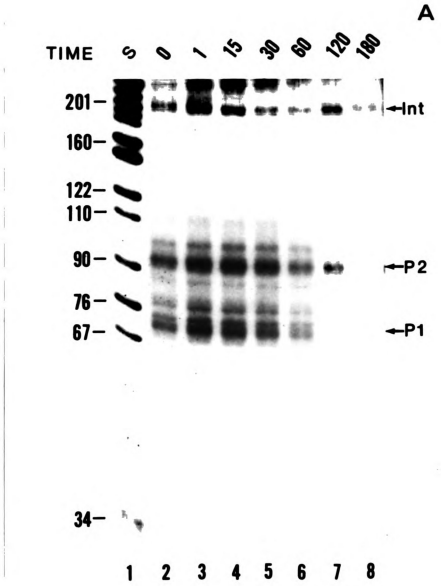
0.33 pmoles of P1 transcript can arise from DNA molecules at which RNA polymerase was initially bound at P2.

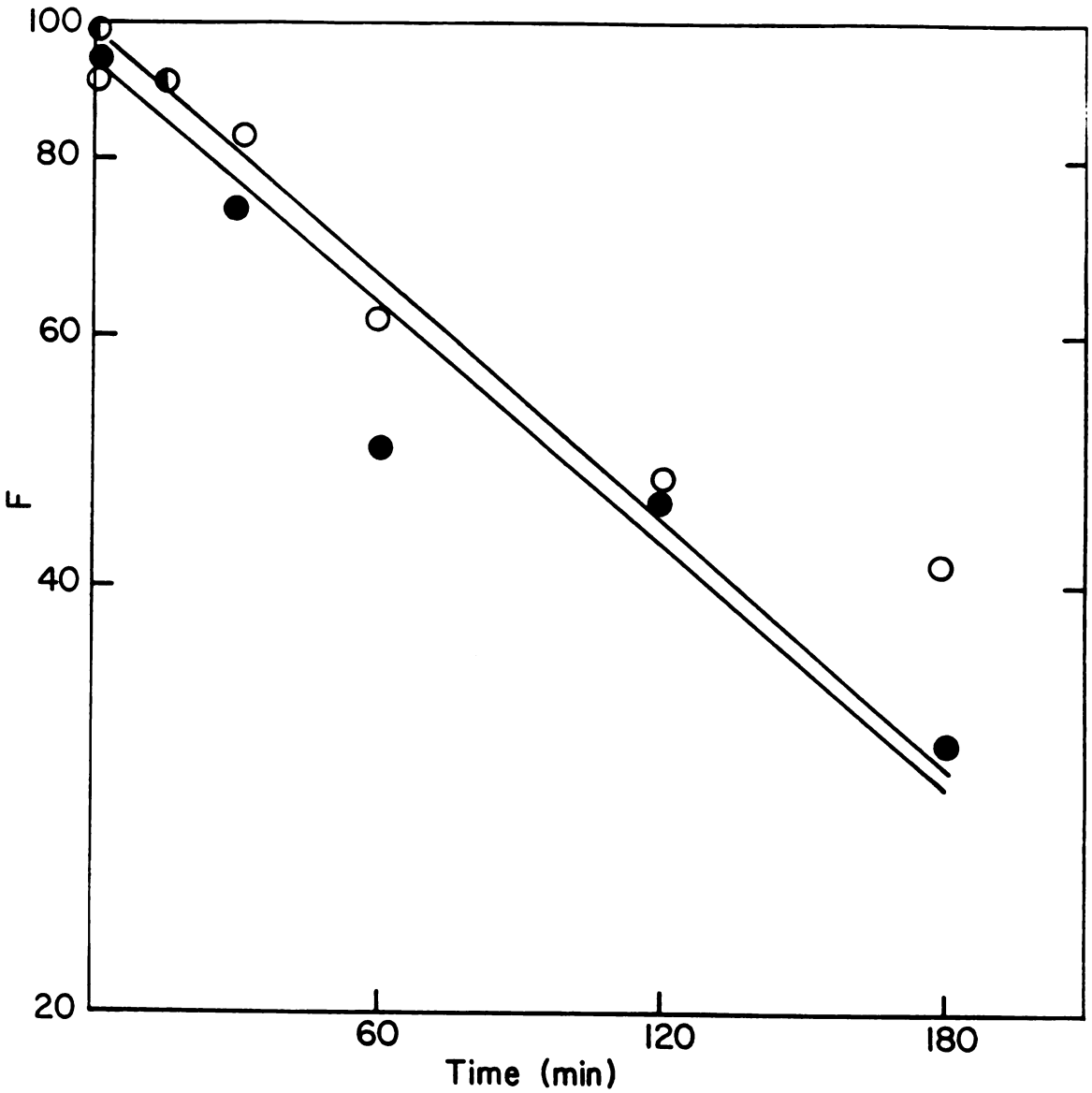
These results are consistent with data from the binding gel shown in Figure 2. In this particular experiment, of the 1.25 pmoles of DNA in each reaction, 0.66 pmoles of DNA had RNA polymerase bound at P2 when enzyme was added alone, while 0.79 pmoles of DNA had polymerase at either P1 or P2 when CAP was added after polymerase. Correcting for the number of P2 complexes which dissociate during the 30 min incubation with CAP (see upper curve, Figure 4), we calculate that, of the 0.79 pmoles of open complexes, 0.59 pmoles (or 75%) had polymerase at P2, while 0.20 pmoles (25%) were P1 complexes. This is in reasonable accord with transcription results, from which we infer that, for the experiment depicted in Figures 2 and 3, initially 70% of the complexes were at P2 while 30% had RNA polymerase bound in such a way that it readily melted in at P1 upon addition of CAP.

At longer times the rate at which enzyme dissociates from P2 is about 60% as fast as the rate at which transcripts arise from P1. At present we do not have an unequivocal explanation for this result. The effect could arise from a low level of CAP-stimulated "clearance" of enzyme from P2 (or from "P3", see below).

Exonuclease III Digestions. The transcription assays imply that stable complexes at both P1 and P2 are found in

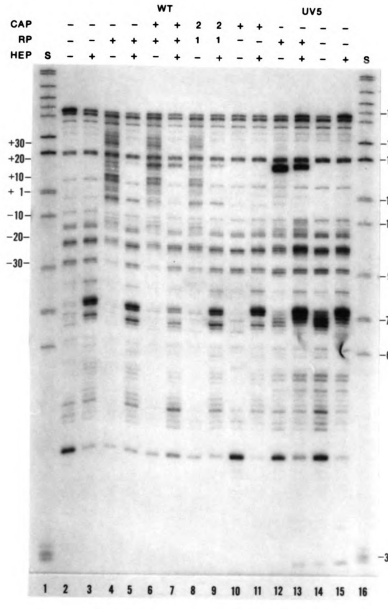
Figure 5. Dissociation kinetics of P1 and P2 open complexes. (A) Complexes were formed by incubating wild type lac DNA with RNA polymerase for 30 min at 37°C, then adding CAP. Thirty minutes later, heparin was added; and solutions were incubated for 0 to 180 min as indicated (lanes 2-8); then nucleoside triphosphates were added and transcription was allowed to proceed for 15 min. Size standards of [^{32}P]-labeled pBR322 digested with HpaII are in lane 1. "Int" refers to a [^{32}P]-labeled, 211 base pair DNA fragment added as an internal standard. (B) First order dissociation kinetics plot for transcripts from the gel in panel (A). Bands were quantified by scintillation counting, corrected for the number of input counts using the internal standard, and plotted as ln F vs time, where F is the ratio of transcripts at the indicated time to the level of transcripts at time zero.


A
TIME
S
0
1
15
30
60
120
180
201
160
122
110
90
76
67
34
Int
P2
P1
1
2
3
4
5
6
7
8


100
80
60
40
20
F
60
120
180
Time (min)

solutions in which RNA polymerase is added to wild type lac DNA. This conclusion was substantiated by use of exonuclease III digestions to determine the locations of promoter-bound proteins. Lanes 12 and 13 of Figure 6 show that at the mutant, CAP-independent lacUV5 promoter, RNA polymerase extends to about +18. The same result is found whether heparin is added or is absent. This provides a reference for polymerase in open complexes which initiate at +1. It is expected that enzyme at the wild type lac P1 site will bind in the same way, since the DNA contacts made by polymerase at the wild type and UV5 promoters are very similar (69). Indeed, in lanes 6 and 7 of Figure 6 one sees that when CAP is added prior to RNA polymerase, a band appears at the P1 position, irrespective of the presence of heparin. When polymerase is added alone and the complexes challenged with heparin a different band is seen, from which it can be deduced that enzyme bound at P2 protects to about -2 (lane 5). When the exo III digestion is done in the absence of heparin, a band at P2 appears, as well as a group of bands at the P1 position (lane 4) for polymerase-DNA complexes (no CAP). Lanes 8 and 9 represent exo III digestions of complexes found when RNA polymerase is added first, then CAP; bands at both P1 and P2 are seen. Corresponding results are found using the lower strand (data not shown).

Figure 6. Exonuclease III digestions of RNA polymerase-CAP-*lac* DNA complexes. The DNAs were 5'-end labeled at the *Pvu*II site at -123, hence data are for the upper strand. Lanes 1 and 16 contain size markers, some lengths of which are indicated at the right. The "+" and "-" symbols denote whether a component is present or absent. For lanes 6 and 7, CAP was added before RNA polymerase; the reverse order of addition was used for lanes 8 and 9. The protected fragments when enzyme is bound at P1 and P2 are 141 and 121 bases in length, respectively.


WT
UV5
CAP
RP
HEP
S
+30
+20
+10
+ 1
−10
−20
−30
190
160
147
122
110
90
76
67
34
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

The autoradiogram shown in Figure 6 is typical of those obtained with exonuclease III from various sources (International Biotechnologies, Inc., Bethesda Research Laboratories, New England BioLabs). The multiple bands found even with DNA alone could not be eliminated by alterations in digestion time or enzyme concentration. Peterson and Reznikoff (67) were able to obtain rather fewer extraneous bands in similar experiments, although their fraction of counts in P1 or P2 bands was low (~10% of input). We performed additional controls using blunt-ended DNA fragments. These gave the same results as found with DNA molecules having partially single stranded ends, in exonuclease III digestions and in binding gel studies as well. The integrity of the DNA fragments was checked by electrophoresis in denaturing gels--no evidence for nicking was seen.

Studies with Fragments Cut at -59. From the transcription assays and exo III data presented above we infer that solutions made by adding RNA polymerase to lac DNA contain some DNA molecules with enzyme bound at P1, some with enzyme at P2; the protein at P1 is heparin-sensitive, and perhaps is in closed complex. Further evidence that this is so derives from experiments using fragments containing lac DNA sequences which extend from the AluI site at -59 to +63. Binding gels (not shown) reveal that no heparin-insensitive complexes form on these fragments.

Furthermore, they do not serve as templates for transcription (Figure 7). One would not expect CAP-dependent initiation from P1, since much of the binding site for the activator [-55 to -70] (70) has been eliminated, but Figure 7 shows that on these truncated DNA molecules no open complexes form at P2 either. In principle, then, adding RNA polymerase to lac DNA cut at -59 (no heparin) might serve as a source of P1 closed complexes, which could be further characterized without interference from the P2 promoter. Exonuclease III digestions of such complexes are shown in Figure 8. A prominent band (or bands) is seen at the P1 position near -45 on the lower strand, along with an additional band as well indicating protection at a second site about 12 nucleotides away. Corresponding results are found using the upper strand (data not shown).

DISCUSSION

The experiments described here are aimed at explaining the observation that in solutions of RNA polymerase and/or CAP plus linear wild type lac promoter fragments a substantial fraction of the DNA molecules do not form heparin-stable complexes. Our interpretation of the data is that at least two types of complexes (and perhaps more) are formed when polymerase interacts with lac DNA: (1) open, heparin-insensitive complexes at P2 and, (2) closed, heparin-sensitive complexes at P1. Adding CAP to such solutions leads to relatively rapid melting-in of the pre-

Figure 7. Transcription from truncated wild type and lacUV5 promoters. All reactions contained AluI-AluI DNA fragments (-59 to +36, with blunt ends). Lanes 1-4, wild type lac promoter; lanes 5-7, lacUV5 promoter; lane 8, size markers. After DNA was incubated with RNA polymerase for 30 min, heparin and nucleoside triphosphates were added and transcription was allowed to proceed for 15 min. The presence of excess glycerol at 5% (lane 2), 10% (lanes 3 and 6) or 20% (lanes 4 and 7) did not increase the low number of transcripts from the wild type promoter.

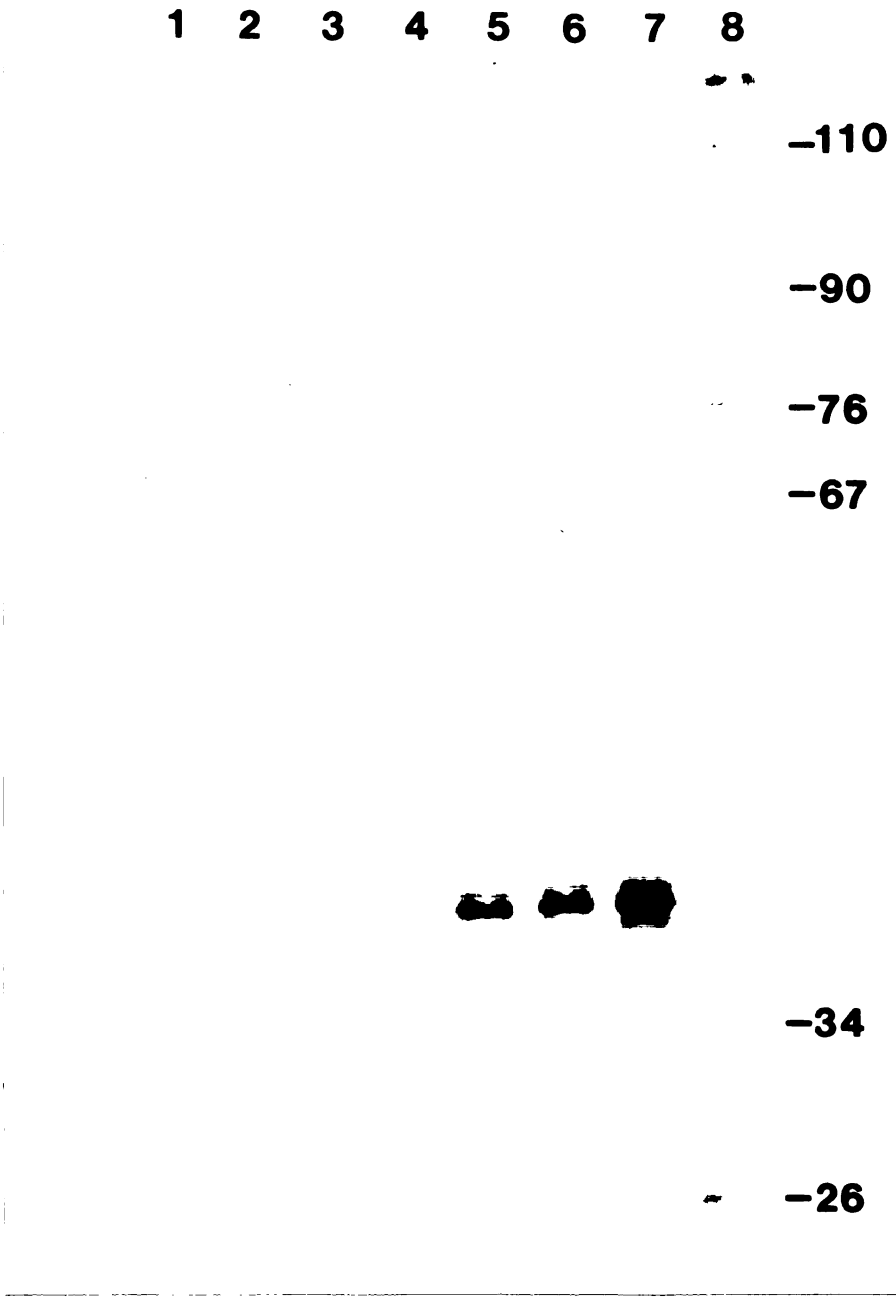
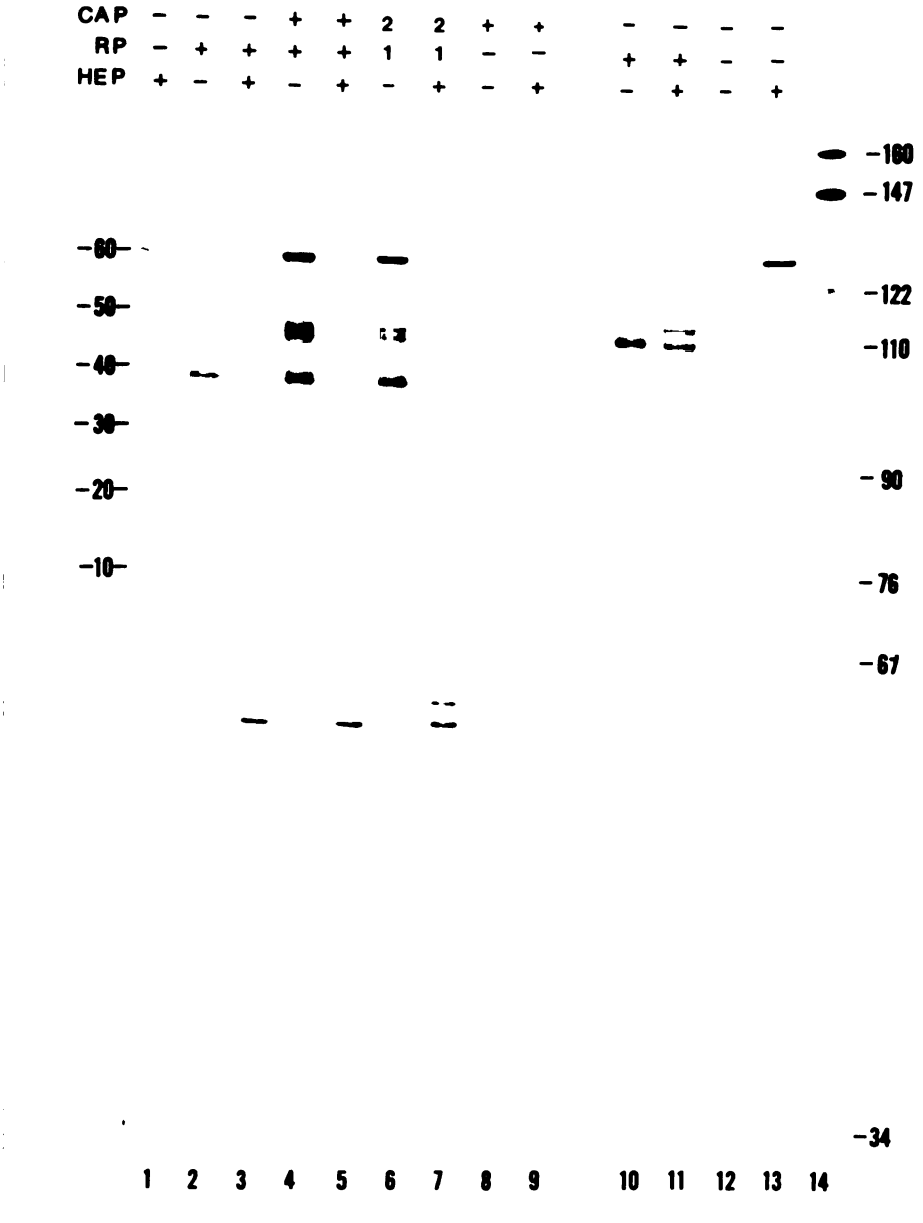

1 2 3 4 5 6 7 8
110
90
76
67
34
26

Figure 8. Exonuclease III digestions of RNA polymerase-CAP-truncated lac DNA complexes. The DNAs extend from a HindIII linker at -59 to the 5'-end labeled EcoRI linker at +67, hence these data are for the lower strand. Size markers are shown in lane 14. Lanes 1-9, wild type lac promoter; lanes 10-13, lacUV5 promoter. The "+" and "-" symbols denote whether a component is present or absent. For lanes 4 and 5, CAP was added first, then RNA polymerase; the reverse order of addition was used for lanes 6 and 7. The protected fragment when enzyme is bound at P1 is 110 bases long, indicating protection to about nucleotide -43.


CAP - - - + + 2 2 + + - - - -
RP - + + + + 1 1 - - + + - -
HEP + - + - + - + - + - + - +
-160
-147
-122
-110
-90
-76
-67
-34
-60
-50
-40
-30
-20
-10
1 2 3 4 5 6 7 8 9 10 11 12 13 14

formed P1 complexes. At the CAP and cAMP concentrations used here, this is not at the expense of P2, since studies of dissociation kinetics by transcription assays reveals that the off-rate of RNA polymerase from P2 is unaffected by the presence of CAP. This finding is consistent with those of McClure and co-workers using supercoiled templates (68), but is at variance with the data of Spassky et al. (70) and of Meiklejohn and Gralla (26), who concluded that CAP displaces RNA polymerase from P2 fairly rapidly. These workers used much higher cAMP and CAP levels relative to the amount of DNA in their reactions. As a control, we repeated the experiment shown in Figure 3, at the concentrations of RNA polymerase (50 nm) and DNA (10 nm) used in reference 27, plus (a) 40 nM CAP, 20 μM cAMP (corresponding to our normal conditions), (b) 300 nM CAP, 20 μM cAMP, and (c) 300 nM CAP, 200 μM cAMP. The $t_{1/2}$ for P1 formation was about 10 min for (a), but was 25 min for case (b), 52 min for case (c). Thus high levels of CAP-cAMP seem to hinder open complex formation at P1 (see also references 68 and 72). However, the dissociation rate from P2 does not seem to be changed at elevated CAP and cAMP concentrations ($t_{1/2}$~100 min). This conflicts with the data of Meiklejohn and Gralla (26); the reason for the discrepancy is not known. We note that 20 μM cAMP is more than sufficient to direct all RNA polymerase to P1 (9) if CAP is added first, and to achieve maximum stimulation of in vitro transcription from the wild type lac

promoter (25). Furthermore, when optical techniques such as fluorescence are applied to this system, high concentrations of CAP and cAMP may be undesirable. Hence our conclusions from experiments at lower levels of CAP and cAMP will be of interest as regulation at lac is studied further.

An unusual, and certainly unexpected, feature of the RNA polymerase complexes at lac P1 and P2 is that each seems to be quite stable. Adding more polymerase or waiting longer times does not change the amounts of P1 and P2 complexes, in spite of the fact that polymerase is melted-in at P2 but not at P1. Furthermore, there may be yet additional types of rather stable RNA polymerase-lac DNA complexes, since we (and others) are unable to force greater than about 60% of the input wild type DNA into heparin-stable complexes, even if CAP is added after RNA polymerase. A putative "P3" site could be indicated by the additional band(s) seen in Figure 8, which can also be discerned on longer DNA fragments around the +35 position (Figure 6, lanes 4,6,8). One might speculate that RNA polymerase has a rather high affinity for a DNA region just to the right of +1, since an A-T to T-A transversion at +1 creates the P^r115 mutation which activates transcription from +12 (73). Exo III digestion of a DNA fragment extending from -59 to +67 also shows dual bands. This is not inconsistent with binding of polymerase at the P^r115 region, since enzyme shifted 12 bp downstream from the P1 position should still

be able to interact normally with a DNA molecule which ends at +36. We note that supercoiled plasmids containing the lac promoter do yield 100% complexes when RNA polymerase is added (68); perhaps the supercoiled DNA does not permit stable "P3" binding.

An alternative explanation for the fact that mixing polymerase with lac DNA does not lead to complete formation of P2 complexes would invoke nonspecific binding of enzyme to the fragment. This is not inconsistent with the data in lane 4 of Figure 6, which shows several groups of bands (no heparin). The nonspecifically bound polymerase could inhibit P2 binding, and would ultimately melt-in at P1 after CAP is added. While we cannot rigorously exclude this hypothesis, we do not favor it primarily because the generalized nonspecific association constant for holoenzyme-DNA binding under these salt conditions is only about 10^6 M^{-1} (74). This implies a rather rapid dissociation rate, hence such complexes should be transient, with polymerase readily finding its way to a stable site (P2) during the 30 min incubation period used. Thus the proposed "nonspecific" complexes would necessarily have some special, high affinity, characteristics (such as implied for "P3" described above). In any event, the additional bands seen in the absence of heparin do not disappear when CAP is added (Figure 6, lanes 6 and 8), even if the activator protein is added first.

Our results seem to put us in conflict with Malan and McClure (68), who concluded from abortive initiation data that CAP enhances the binding of RNA polymerase to form a closed complex, but does not affect the rate of isomerization to an open complex at either the lac wild type or UV5 P1 promoter. We find that in the absence of CAP, RNA polymerase binds tightly to wild type P1 in closed form, yet it melts in hardly at all. When CAP is added, these closed complexes are converted to open ones, indicating a substantial CAP effect on the isomerization process. Note that our transcription assays indicate that the rate of open complex formation at P1 does not depend on the order of addition of proteins; thus P1 complexes formed by polymerase in the absence of CAP are not in an unusual metastable state, but merely await CAP to stimulate their melting-in.

The above data do confirm the findings of McClure and coworkers (9,68) and of Peterson and Reznikoff (67) that CAP can occlude RNA polymerase binding at P2. Our results also support the notion that CAP acts much more directly (9,67,68). It appears that P1 and P2 have roughly equal affinities for polymerase. Thus, even if lac P2 is transcriptionally inactive in vivo, channeling by CAP of polymerase away from P2 to P1 would give only a few-fold stimulation, far less than what is observed (75). And, of course, we show that CAP is indeed required for melting-in of RNA polymerase at P1.

Studies with promoter fragments cut at -59 confirm *in vivo* results from deletion mutant analysis (72); namely, that nucleotides critical for binding at P2 are not present in the *Alu*I-cut DNA, hence RNA polymerase does not bind there. In principle then, adding polymerase to these shortened DNA molecules should lead to but one entity in solution--closed complexes at P1. The exo III data indicate that a sizable fraction of complexes on the short fragments are at P1, but that there may be interference from another, presumably overlapping, binding site. Additional work is needed to characterize this system further.

Differences and similarities between the *lac* and *gal* control regions were touched on in Chapter I and have been discussed by many others as well (70). From the data presented here it appears that when RNA polymerase is given the option of binding to either of the overlapping P1 and P2 *lac* promoters, it interacts to about the same extent with each. At *gal* a similar result is found, except that substantial levels of P1 transcription are seen in the absence of CAP (33,68). So these catabolite sensitive promoters appear to function in analogous ways, but with significant variations on a main theme.

CHAPTER III.

SPECIFIC SEQUENCES DOWNSTREAM OF -7 ARE NOT ESSENTIAL FOR PROPER UTILIZATION OF THE *E. COLI* LACTOSE PROMOTER

CHAPTER III.

INTRODUCTION

The studies described in Chapter II indicated that RNA polymerase will bind to the wild type *lac* promoter region in the absence of CAP, forming an open complex at P2 in addition to a stable, yet heparin-sensitive, closed P1 complex. Polymerase may bind to other sites as well. Exonuclease III protection experiments suggested the presence of a strong binding site downstream of P1 which we called "P3".

The formation of closed complexes suggests a potential source of these intermediates for physical study. In this chapter I report on our attempts to construct DNA fragments which have but one polymerase binding site, P1. In addition, the construction of a set of synthetic *lac*UV5 promoter fragments is described. These constructs may be useful for the synthesis of RNA fragments for physical study.

MATERIALS AND METHODS

Proteins. *E. coli* RNA polymerase holoenzyme and CAP were prepared and characterized as described in Chapter II.

lac Promoter Fragments. All fragments are numbered relative to the start of P1 transcription as being +1. DNA molecules 211 bp long (-140 to +63 plus EcoRI linkers) containing the wild type or the mutant lacUV5 promoters were purified from recombinant pMB9 plasmids provided by Forrest Fuller. A wild type lac promoter fragment which extended from -59 to +63 was prepared as described in Chapter II.

Restriction fragments were purified by separation in polyacrylamide gels (64) followed by electroelution in an IBI Model UEA Electroeluter (International Biotechnologies, Inc., New Haven, CT) using 10 M ammonium acetate as the high salt medium as suggested by the manufacturer. DNA purified in this manner rarely required further treatment. DNA was 5′ end-labelled as with [γ-^{32}P]-ATP and T4 polynucleotide kinase after removal of the 5′ phosphate groups with calf intestinal phosphatase (64). Concentrations of DNA were determined spectrophotometrically using ϵ_{260}=13,000 $(M,bp)^{-1}$ (65) and are expressed as molar in promoter fragments.

Synthetic DNA fragments were made using phosphoramidite chemistry (76) on an Applied Biosystems Model 380B DNA synthesizer (Applied Biosystems, Inc., Rose City, CA). The products were purified by electrophoresis in polyacrylamide gels followed by electroelution. Purified fragments were lyophilized and stored at -80°C in twice-distilled water.

Binding Reactions, Binding Assays by Gel Retardation, and Transcription Assays. All reactions were performed as described in Chapter II.

Construction of a P2⁻ Fragment. A mutation was introduced at position -29 on the wild type lac promoter by site specific mutagenesis (77). A mixture of 23 nt single strand oligonucleotide primers equivalent to the upper strand of the lac sequence from -36 to -14, but which contained either a C, G or A residue at the position opposite the A at -29, was annealed to the complementary, lower strand of lac DNA from the single stranded phage vector M13mp18 (78), and was extended with the Klenow fragment of E. coli DNA polymerase I in the presence of the four deoxynucleoside triphosphates. After incubation for 2 hr at room temperature, T4 DNA ligase was added and allowed to react for another 2 hr to catalyze the circular ligation of the extended molecules prior to transformation of E. coli JM101 cells (64). On the following day plaques were absorbed to nitrocellulose filters and probed at room temperature with the mixed primers which had been 5′ end-labelled with [γ-^{32}P] ATP. Those plaques which retained the probe were then subjected to washes with 6X SSC (1X = 150 mM NaCl, 15 mM sodium citrate and 0.5 mM EDTA, pH 7.2) at increasing temperatures until only a few plaques remained hybridized. For the 23-mer used here, a temperature of 65°C was found to be sufficient to cause dissociation of

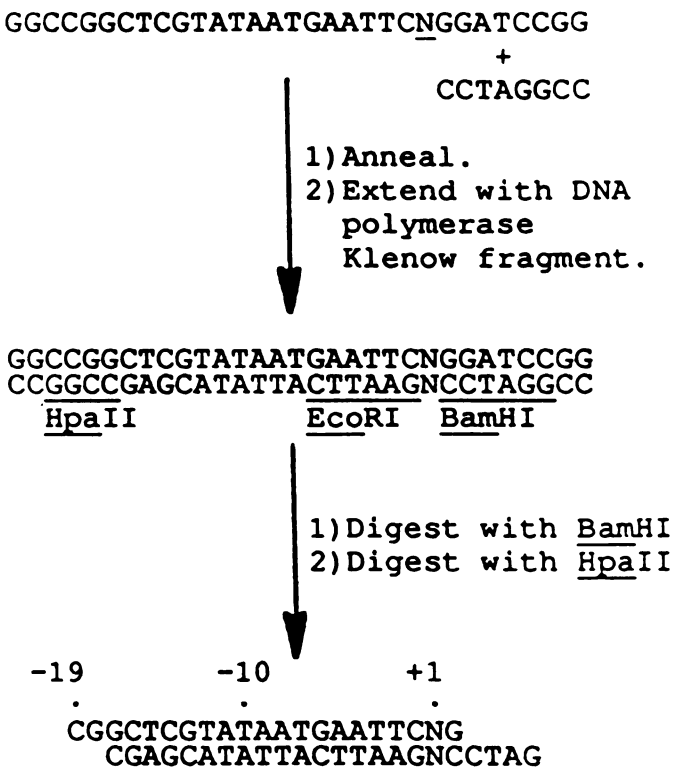
complexes of probes with nonmutated phage DNA (such hybrids would contain a mismatched base pair). Phage DNAs from plaques which were still bound by the probe at 65°C were subjected to sequence analysis (66,79). Inserts with the desired sequences between the EcoRI site at +63 and the PvuII site at -124 were excised and isolated following electrophoresis in a polyacrylamide gel, and were recloned between the HindIII and EcoRI sites of pBR322 after a synthetic HindIII linker was ligated to the blunt PvuII end.

Construction of Deletion Mutants. A 476 bp EcoRI-BamHI DNA fragment which contained the lac sequence from -59 to +63 was isolated from the plasmid pDL1 and used as a substrate for the deletion reactions. The DNA was allowed to react with 0.5 to 2.0 units of Bal31 exonuclease for up to 5 min (64). Aliquots were removed at 30 sec intervals, quenched with 10 mM EGTA, and stored on ice. Portions of these samples were analyzed by polyacrylamide gel electrophoresis; those which contained products more than 400 base pairs in length were pooled and used for cloning. Blunt ends were generated by filling in the ends of the fragments using the Klenow fragment of E. coli DNA polymerase I after which synthetic EcoRI linkers were ligated to both ends of the molecules (64). The products of the above reactions were digested with EcoRI and HindIII (which catalyzes cleavage at -59 in these constructs) to release the lac sequences with the newly created end points.

These fragments were religated into pBR322 previously digested with EcoRI and HindIII and the recombinant plasmids were used to transform E. coli HB101 (64). The deletion end points were determined by sequence analysis of the recloned fragments (66). For the experiments described here we fused the HpaII-EcoRI region of the deletion fragments to a 106 bp HindIII-HpaII segment from the P2$^-$ constructs. Thus we have created fragments which contain wild type lac DNA (except at -29) from -124 to the deletion point, and which include the CAP binding site necessary for proper stimulation of these promoters.

Modified lacUV5 Promoter Fragments. A series of lacUV5 promoter fragments which contained an EcoRI site from -6 to -1, an A, C, G, or T residue at +1, and a BamHI site from +2 to +7 was constructed using DNA fragments that span the HpaII site at -20 to the BamHI site at +2. To make these sequences, a threefold molar excess of a 10 nt primer was annealed at 100°C to the complementary 3'-end of a 31 nt template, cooled to room temperature, and extended using the Klenow fragment of E. coli DNA polymerase I (see Figure 9). The resulting duplex DNAs were cut with HpaII and BamHI and ligated to a 106 bp lac P2$^-$ fragment which extended from the PvuII site at -124 to the HpaII site at -20. A HindIII linker was ligated to the blunt PvuII end and the resultant fragments were cloned between the HindIII and BamHI sites of pBR322. For binding and transcription experiments, DNA

Figure 9. Construction of the modified lacUV5 promoters. The 31 nt template and the 10 nt primer are indicated at the top of the figure. The underlined "N" in the template sequence indicates the A, C, G, or T at position +1 in the final constructs. See Materials and Methods for details of the complete construction procedure.


GGCCGGCTCGTATAATGAATTCNGGATCCGG
+
CCTAGGCC
1)Anneal.
2)Extend with DNA polymerase Klenow fragment.
GGCCGGCTCGTATAATGAATTCNGGATCCGG
CCGGCCGAGCATATTACTTAAGNCCTAGGCC
HpaII
EcoRI
BamHI
1)Digest with BamHI
2)Digest with HpaII
-19
-10
+1
CGGCTCGTATAATGAATTCNG
CGAGCATATTACTTAAGNCCTAG

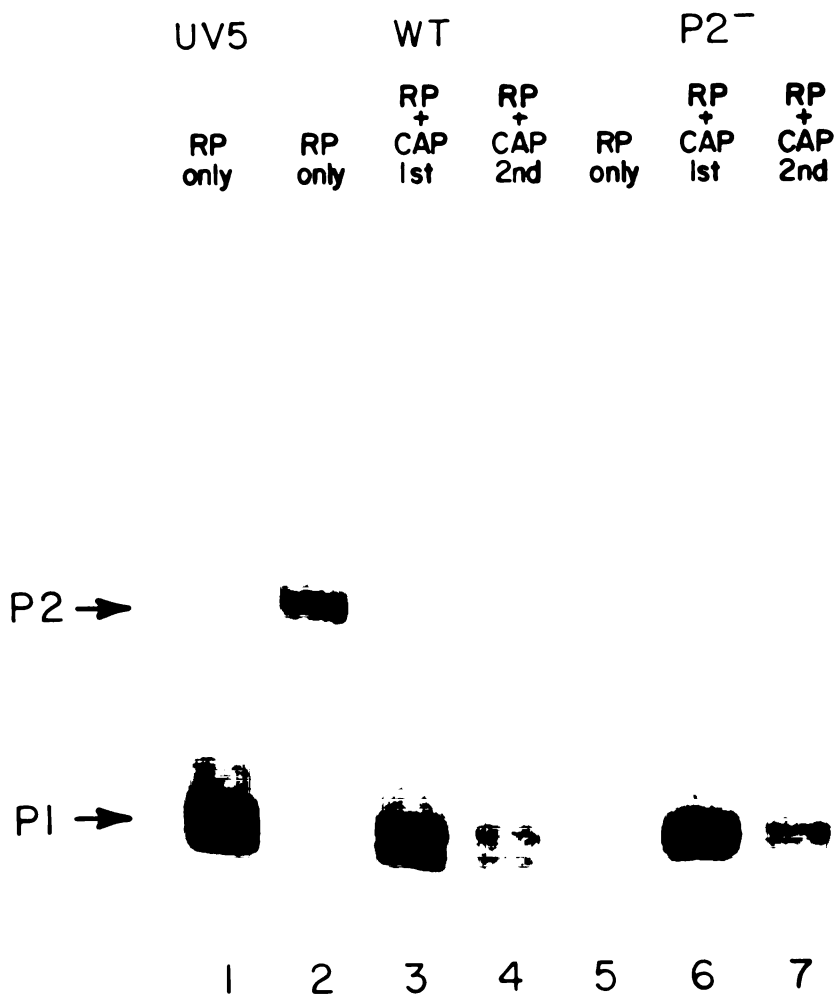
fragments were prepared by cutting the recombinant plasmids with HindIII and SalI to release a 406 bp fragment which contained the desired promoter. The HindIII-SalI fragments were further digested with CfoI to yield a 123 bp fragment which contained the promoter plus 35 bp of DNA from pBR322 downstream of the BamHI site.

RESULTS

Mutation of the P2 Promoter. As a first step toward a more detailed understanding of lac promoter structure and function we decided to eliminate possible interference from the lac P2 promoter region by site specific mutagenesis. We chose the position at -29 because it lies in the P2 -10 region and is distinct from the P1 -35 region and the CAP binding site (8,9). We attempted to make all three possible changes; T to A, C or G. Of more than 50 transformants screened, we recovered no "G" mutants and only one "C" mutant, the rest being "A". The "C" mutant also contained a second mutation in the CAP site rendering it CAP^- as well as $P2^-$ (results not shown). The T-A to A-T transversion at -29 proved to be suitable for our studies.

A runoff transcription assay was used to measure the ability of the putative $P2^-$ fragment to initiate transcription at -22 (Figure 10). Wild type or mutant lac promoter fragments were incubated with RNA polymerase (lanes 1, 2 and 5) or with CAP and polymerase (lanes 3, 4, 6 and 7) prior to the addition of nucleotides. In the absence of

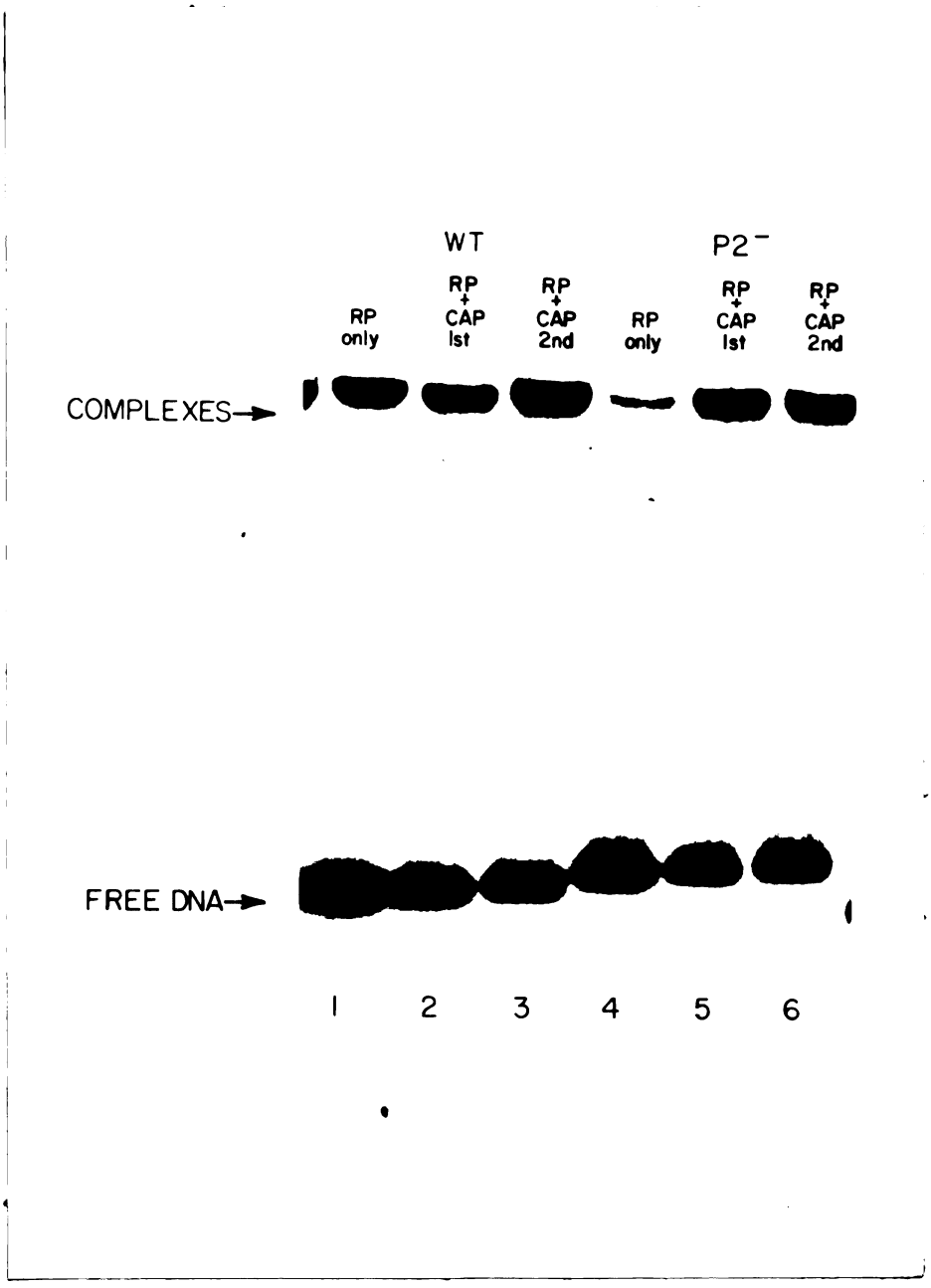
Figure 10. Transcription in vitro from wild type and mutant lac promoters. Lane 1, lacUV5 DNA plus RNA polymerase; lane 2, wild type lac DNA plus RNA polymerase; lane 3, wild type lac DNA plus CAP then RNA polymerase; lane 4, wild type lac DNA plus RNA polymerase then CAP; lane 5, $P2^-$ lac DNA plus RNA polymerase; lane 6, $P2^-$ lac DNA plus CAP then RNA polymerase; lane 7, $P2^-$ lac DNA plus RNA polymerase then CAP. The transcript from P1 is 69 nt long, that from P2 is 91 nt long.


UV5
WT
P2⁻
RP only
RP only
RP + CAP 1st
RP + CAP 2nd
RP only
RP + CAP 1st
RP + CAP 2nd
P2
P1
1
2
3
4
5
6
7

CAP, a P2 transcript of 91 nucleotides was formed as expected in the wild type case but not from the altered $P2^-$ promoter (the mutant UV5 promoter never yields a P2 product). The addition of CAP to the mutant did not affect this result (lanes 6 and 7). This experiment demonstrates that the $P2^-$ mutant promoter does not form an "open", heparin-stable complex capable of initiating transcription in the absence of CAP. However, it does not reveal whether the polymerase molecule might form a nonproductive, yet rather stable "closed" P2 complex which is capable of blocking the P1 site in the absence of heparin.

We used a gel electrophoresis assay (54) to measure the extent of polymerase binding to the wild type and $P2^-$ promoter sequences. In Figure 11 it can be seen that in this experiment incubating only RNA polymerase with the wild type lac promoter results in the formation of complexes on 29% of the DNA fragments (lane 1) while only a few percent of the $P2^-$ fragments form complexes (lane 4), again indicating that the $P2^-$ promoter is incapable of stable polymerase binding in the presence of heparin. Addition of CAP before RNA polymerase leads to heparin-stable (P1) complexes being formed on 32% of the wild type DNA and on 38% of the $P2^-$ fragments upon subsequent addition of the enzyme (lanes 2 and 5, respectively). The addition of CAP to preformed polymerase-DNA complexes increases the total number of complexes in the wild type case to 42% (lane 3);

Figure 11. Gel binding assay of protein-DNA complexes formed at wild type and P2$^-$ mutant lac promoters. Lane 1, wild type lac DNA plus RNA polymerase; lane 2, wild type lac DNA plus CAP then RNA polymerase; lane 3, wild type lac DNA plus RNA polymerase then CAP; lane 4, P2$^-$ lac DNA plus RNA polymerase; lane 5, P2$^-$ lac DNA plus CAP then RNA polymerase; lane 6, P2$^-$ lac DNA plus RNA polymerase then CAP.


WT
P2⁻
RP only
RP + CAP 1st
RP + CAP 2nd
RP only
RP + CAP 1st
RP + CAP 2nd
COMPLEXES
FREE DNA
1
2
3
4
5
6

this is a mixture of P1 and P2 complexes as seen from transcription data (Figure 10, lane 4). For the mutant, 34% of the DNA is found in P1 complexes after CAP is added to a polymerase-DNA solution (Figure 11, lane 6, and Figure 10, lane 7). Thus the binding and transcription data show that the mutation at -29 eliminates P2 utilization yet has no significant effect on initiation from the P1 promoter.

A closed complex at the $P2^-$ site could block both the CAP and P1 binding sites, so that adding CAP after RNA polymerase would result in substantially fewer heparin-stable complexes than when polymerase is added after CAP. Since the level of stable binding seems to be about the same on the $P2^-$ fragment regardless of the order of addition of CAP and polymerase (38% vs 34%), it appears that no tight-binding closed complexes form at the $P2^-$ sequence.

Deletion of DNA downstream of -7. As described in Chapter II, it has long been known that using wild type *lac* promoter fragments it is difficult to find conditions under which all of the DNA forms open, heparin-stable complexes with CAP and RNA polymerase (68). We inferred from exonuclease III protection experiments that there might be a region of DNA downstream of +1, distinct from P1 or P2, which can be covered by a stably-bound RNA polymerase molecule that would then exclude binding at P1 and/or P2; we called this putative site "P3" (47). We reasoned that deletion of DNA to the right of at least +1 would remove the

interfering sequence, thus more open complexes might form at P1. Using Bal31 we generated five inserts with endpoints at +25, +19, +14, +1 and -7. The deletions were joined to $P2^-$ sequences to form what should be $P1^+P2^-P3^-$ promoter fragments. Such fragments might be capable of binding polymerase at P1 at a high level. We used a polyacrylamide gel electrophoresis assay to measure the effect of the deletions on heparin-stable RNA polymerase binding (Figure 12). In this experiment the level of binding on the deletion to +25 was 44% when CAP was added first and 49% when CAP was added second. These levels are comparable to those seen on wild type fragments. When the DNA between +25 and +14 was removed, the level of binding decreased to about one-third of the normal level (Table 1). However, we replaced the deleted DNA with 72 bp of nonlac DNA to see if the decreased binding was due to the loss of specific DNA contacts or simply to the loss of DNA in general. When these longer fragments were tested by gel retardation assays, binding was restored to the wild type level (Table 1).

We used the runoff transcription assay to measure the ability of the deletion fragments (plus 72 bp of nonlac DNA) to support transcription (Figure 13). In all cases transcription begins at the proper start point. Normalizing for the number of radioactive U residues in the RNAs made

Table 1

Degree of Promoter Occupancy of *lac* Deletion Mutants

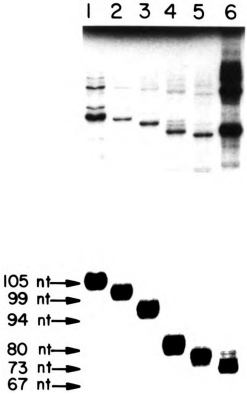
lac promoter end point	Length of added DNA	CAP added 1st	CAP added 2nd
+25	-	44%	49%
	72bp	51%	48%
+19	-	32%	30%
	72bp	49%	48%
+14	-	17%	25%
	72bp	42%	47%
+1	-	nd*	nd*
	72bp	44%	45%
-7	-	nd*	nd*
	72bp	52%	49%

*nd = not determined

Figure 12. Gel binding assay of protein-DNA complexes formed at deleted lac $P2^-$ promoter fragments. Lane 1, deletion to +25 plus RNA polymerase; lane 2, deletion to +25 plus CAP then RNA polymerase; lane 3, deletion to +25 plus RNA polymerase then CAP; lane 4, deletion to +19 plus RNA polymerase; lane 5, deletion to +19 plus CAP then RNA polymerase; lane 6, deletion to +19 plus RNA polymerase then CAP; lane 7, deletion to +14 plus RNA polymerase; lane 8, deletion to +14 plus CAP then RNA polymerase; lane 9, deletion to +14 plus RNA polymerase then CAP. These fragments contain no additional DNA, thus end at the deletion.


+25 DELETION
+19 DELETION
+14 DELETION
RP only
RP + CAP 1st
RP + CAP 2nd
COMPLEXES
FREE DNA
1
2
3
4
5
6
7
8
9

Figure 13. Transcription in vitro from wild type and deleted lac $P2^-$ promoters. All of the fragments in this experiment contain 72 bp of pBR322 DNA attached at the deletion end point. In each reaction CAP was added first then RNA polymerase. Lane 1, deletion to +25; lane 2, deletion to +19; lane 3, deletion to +14; lane 4, deletion to +1; lane 5, deletion to -7; lane 6, wild type lac DNA.


1 2 3 4 5 6
105 nt
99 nt
94 nt
80 nt
73 nt
67 nt

shows that the amount of transcription is about equal in all cases. Replacement of DNA downstream of -7 does not seem to affect utilization of the P1 promoter, nor does the mutation at -29. Thus a $P2^-$ promoter with heterologous DNA downstream of -7 behaves essentially the same as does the wild type lac promoter in CAP-stimulated transcription.

These results do not reveal why we do not find all of the DNA from the $P1^+P2^-P3^-$ constructs in stable complexes with RNA polymerase (in the presence of CAP). If CAP is in excess all of the wild type fragments will form specific CAP-DNA complexes (data not shown). Since our polymerase preparation is approximately 40% active (see Materials and Methods), perhaps the inactive molecules can bind to the P1 and P2 promoters to form tight, yet heparin-sensitive complexes, which are not transcriptionally active. Such complexes would reduce the number of heparin-stable forms present, thereby reducing the total number of complexes observed. Thus the observation that solutions of CAP, RNA polymerase and wild type lac fragments do not yield transcripts from every DNA molecule may arise from properties of the enzyme preparation and not a priori from the DNA sequence.

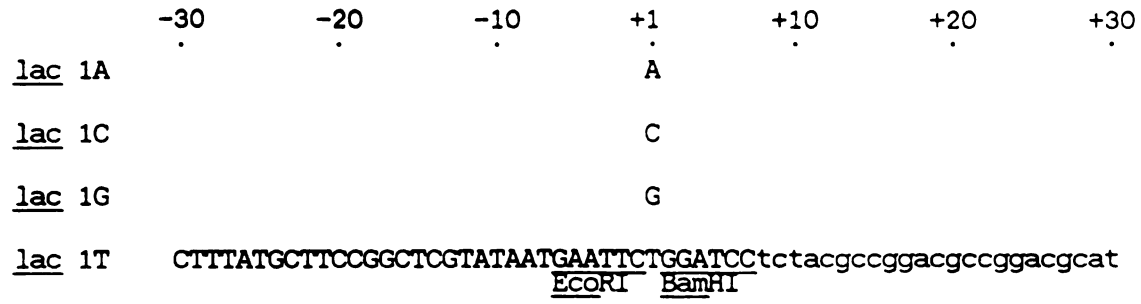
Modified lacUV5 Promoter Fragments. The activity of the deletion fragments suggested the potential usefulness of the lac promoter as a tool for in vitro RNA synthesis. Because a -7 deletion fragment can support RNA polymerase

binding and transcribe efficiently from the wild type lac promoter, we synthesized similar DNA fragments based on the lacUV5 promoter. Such fragments could be of general interest as vehicles for the production of RNAs with precisely defined sequences. The modified promoter fragments comprise lacUV5 sequences to -7, followed by an EcoRI site from -6 to -1 for convenient cloning, followed by one of the four nucleotides at +1, followed by a BamHI site for additional convenience in cloning (see Figure 14A). Because all these fragments are identical, except for the first base, we can systematically assess the contribution of the initiating nucleotide to the transcription process. We used the gel retardation assay to determine the ability of the modified fragments to bind RNA polymerase in vitro. In all cases the level of binding was equivalent to that observed on the normal UV5 fragment (data not shown).

An important factor in the utilization of these promoters is their ability to initiate transcription precisely at +1. We used a runoff transcription assay to determine the level of transcription as well as the point of initiation (Figure 15). In lane 1, transcription was allowed to proceed from the lac 1A fragment in the presence of [γ-^{32}P]-ATP, while in all other lanes transcripts were labelled with [α-^{32}P]-UTP. The nucleotide concentrations used in lanes 1-6 were 50 μM, while the nucleotide concentrations used in lanes 7-11 were 200 μM. The

Figure 14. The DNA sequence from -30 to +30 of the fragments used in this study. (A) The modified lacUV5 promoters. The underlined sequences indicate the EcoRI and BamHI sites. The lower case letters correspond to the DNA added to the right of the BamHI site. The DNA added to these fragments is equivalent to the sequence between the BamHI site at position 375 and the CfoI site at position 414 in pBR322 (13). (B) The wild type and deleted $P2^-$ lac promoters. The underlined sequence is the EcoRI linker added to the right of the deletion end point. The lower case letters correspond to the DNA used to replace the deleted sequences. The replacement DNA is equivalent to the sequence between the EcoRI site at position 4360 and the DdeI site at position 4290 in pBR322 (13).

(A)


-30
-20
-10
+1
+10
+20
+30
lac 1A
A
lac 1C
C
lac 1G
G
lac 1T
CTTTATGCTTCCGGCTCGTATAATGAATTCTGGATCCtctacgccggacgccggacgcat
EcoRI
BamHI

(B)

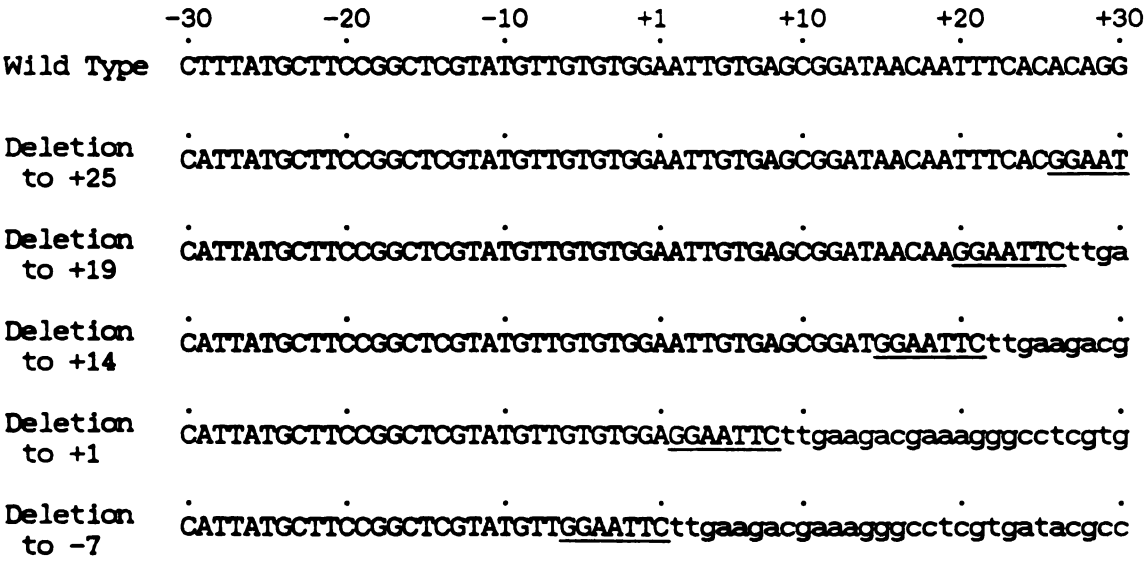
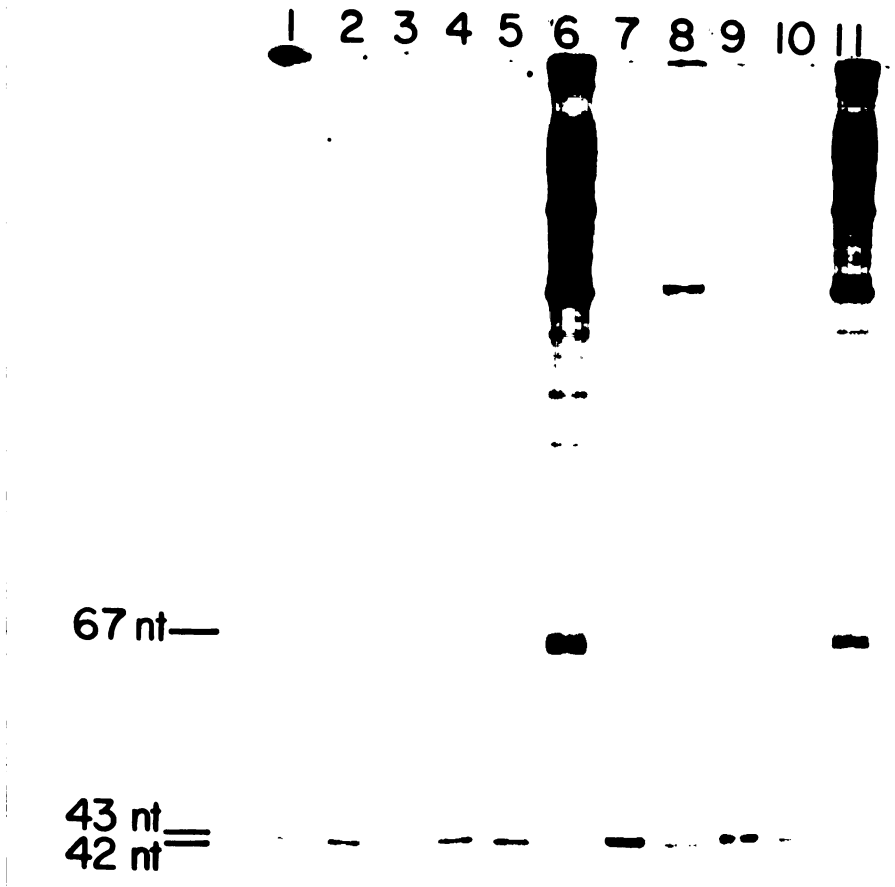

-30
-20
-10
+1
+10
+20
+30
Wild Type
CTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGG
Deletion to +25
CATTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACGGAAT
Deletion to +19
CATTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAAGGAATTCttga
Deletion to +14
CATTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATGGAATTCttgaagacg
Deletion to +1
CATTATGCTTCCGGCTCGTATGTTGTGTGGAGGAATTCttgaagacgaaagggcctcgtg
Deletion to -7
CATTATGCTTCCGGCTCGTATGTTGGAATTCttgaagacgaaagggcctcgtgatacgcc

Figure 15. Transcription in vitro from modified lacUV5 promoters. Lanes 1-6, the concentration of each nucleotide was 50 μM, lanes 7-11 the concentration of each nucleotide was 200 μM. Lanes 1, 2, and 7, the lac 1A promoter; lanes 3 and 8, the lac 1C promoter; lanes 4 and 9, the lac 1G promoter; lanes 5 and 10, the lac 1T promoter; lanes 6 and 11, the normal lacUV5 promoter. The transcripts in lane 1 are labelled with [γ-^{32}P]-ATP while those in lanes 2-11 are labelled with [α-^{32}P]-UTP.


1 2 3 4 5 6 7 8 9 10 11
67 nt
43 nt
42 nt

transcripts synthesized from the lac 1A and 1G promoters (lanes 2, 4, 7 and 9) correspond precisely with those formed in lane 1 indicating that transcription occurs from +1. When transcription is performed using the lac 1C and 1T fragments, some transcripts are observed which apparently start at -1 and +2 (lanes 3, 5, 8 and 10). The lac 1C fragment at 50μM nucleotides appears to make mostly the shorter product (lane 3) while at 200μM (lane 8), slightly more than half of the transcripts seem to be from +1. Transcription experiments at the 500μM nucleotide level yield results identical to those done at 200 μM. Thus, RNA polymerase can initiate transcription with any base at +1, though in some cases a fraction of initiation occurs away from the normal start point. (We note that the difference in length of the transcripts seen in Figure 15 could also be due to alternate termination points, though it is not obvious why transcripts from sequences which differ only in their first nucleotide would terminate differently). Regardless of the level of nucleotides used in these experiments, the overall total level of transcription corresponds precisely to that seen for the normal UV5 promoter after correcting for the number of U residues per transcript.

We have also tested the rates at which the synthetic promoters initiate transcription. Transcription reactions were sampled at early times during the experiments and the

amount of products determined (data not shown). The lacUV5 control reaches its maximum level of transcription with a half-time of 45 sec, while the lac 1A and 1C fragments have similar half-times of 41 sec and 34 sec, respectively. Rates on the lac 1G and 1T fragments are somewhat slower, with half-times of 61 and 99 sec, respectively. There appears to be no correlation between the time required to initiate transcripts and whether alternate start sites can be used. Nevertheless, in all cases, the reactions are easily completed by the end of the usual 20 min assay.

DISCUSSION

In this Chapter we address some fundamental questions concerning promoter structure and utilization. By using a P2 mutation and various deletions we show that specific DNA sequences downstream of the -10 consensus region are not critical for efficient use of the lac P1 promoter by RNA polymerase. When the DNA downstream of -7 is removed and replaced with nonhomologous DNA, both the lac wild type and the stronger lacUV5 promoters appear to behave identically to their native forms. The apparent importance of downstream sequences for some promoters was touched on in Chapter I. Thus, we might have expected that our deletion fragments would be able to bind RNA polymerase less well and transcribe less efficiently than the natural sequences. Previous studies (49,50) were confined mostly to promoters

from the phages T5, T7 and lambda, which are among the strongest promoters that function in E. coli. It was speculated that the strength of these promoters derives in part from specific contacts made in the downstream region which, when changed, can decrease their efficiency. Perhaps the downstream sequences are less important in the case of the weaker lac promoters. Further substantiation for this notion comes from work with the +14, +1 and -7 deletion fragments. Comparison of sequences of the fragments with nonhomologous DNA added (Fig 13B) shows that the sequences between +1 and +20 are quite different yet the levels of transcription from these promoters are identical.

To do physical studies on RNA structure it is necessary to obtain substantial quantities of specific RNA molecules. We have constructed a modified lacUV5 promoter which contains an EcoRI site immediately downstream of the -10 consensus sequence. Such a construct allows the convenient cloning of any DNA sequence, and initiation of transcription at +1 precisely. It is known that promoters which utilize RNA polymerases from phages T7 and SP6 require at least a G at +1 and perhaps other specific DNA sequences downstream for efficient transcription (51-53). We assayed the ability of E. coli RNA polymerase to initiate transcription from modified lacUV5 promoters by using a set of synthetic promoters with each of the four bases at +1. The transcription data demonstrate that initiation is almost

uniformly at +1 when an A or a G is at that position. When a pyrimidine is at +1, the start of transcription becomes less precise. Comparing the start points of more than 200 E. coli and coliphage promoters (40,41) reveals that transcripts begin with an A 53% of the time, a G 36% of the time, a T 7% of the time and a C 4% of the time, indicating a preference for purines at +1. However, since significant activity from +1 occurs at all of our new promoters, such constructs may find use in preparing RNA for other studies. In the presence of heparin approximately one transcript is synthesized per promoter. When no heparin is present, reinitiation allows more transcripts to be made. We have observed up to 25 reinitiations per promoter during a 2 hour incubation. In one experiment, 0.7 μg of a 67 nt RNA transcript was produced from 0.17 μg of input DNA (203 bp). Transcripts from such reactions were isolated by electrophoresis in polyacrylamide gels, and their integrity examined after elution. When care was taken to eliminate RNase contamination, transcripts could be purified intact and in good yield (approximately 75% recovery from the gel).

In summary, the experiments described here address the basic issue of how a promoter is defined. Using deletion mutants of the wild type lac promoter it appears that, at least for that promoter and its derivatives, the -10 and -35 consensus sequences are sufficient to direct RNA polymerase to bind and transcribe. Using modified lacUV5 promoters we

show that transcription can occur at +1 regardless of what nucleotide is there. Thus, it seems that specific sequences outside of the -10 and -35 regions are not required for efficient utilization of the lactose promoter.

CHAPTER IV.

DISSOCIATION KINETICS OF RNA POLYMERASE FROM THE *lac*UV5 PROMOTER

CHAPTER IV.

INTRODUCTION

In Chapters II and III I describe studies of the interaction of RNA polymerase with wild type lac promoter fragments. The inability to form heparin-stable, open complexes on all of the promoter fragments was addressed. In Chapter II it was suggested that polymerase may form a tight, yet heparin-sensitive complex at a site called P3. However, deletion analyses described in Chapter III do not support this hypothesis. Rather, it is suggested that a subset of polymerase molecules with an altered ability to interact with the promoter may be responsible for incomplete promoter occupancy. Here, I report on the dissociation of RNA polymerase from the lacUV5 promoter. An unusual finding is described. This result is put into context with Chapters II and III to suggest that further study of the structure of RNA polymerase would be interesting.

MATERIALS AND METHODS

Proteins and lac Promoter Fragments. E. coli RNA polymerase holoenzyme and CAP were prepared as described in Chapter II.

Binding Reactions, Binding Assays by Gel Retardation, and Transcription Assays. All reactions were performed as described in Chapter II.

Dissociation Assays. Binding reactions were assembled as described in Chapter II. After the incubation period, heparin was added to a final concentration of 100μg/ml. The reactions were allowed to incubate for an additional time from 30 sec to 5 hr prior to being loaded onto a 5% polyacrylamide gel.

RESULTS

Determination of the Dissociation Constant of RNA Polymerase from the lacUV5 Promoter. One of the measures of promoter strength is the rate at which RNA polymerase dissociates from the promoter. As indicated in Chapter I, the dissociation of polymerase from the promoter is slow relative to the rate of isomerization. The rate of dissociation of RNA polymerase from the lacUV5 promoter was assessed at 37°C. The autoradiogram in Figure 16 represents a typical experiment to determine k_{-2} or the "off rate" for polymerase under our normal conditions of a 0.1 M KCl buffer. The fraction of stably bound DNA was determined at each time point and plotted logarithmically vs time (Figure 17). Similar experiments were performed at 0.075 M, 0.14 M, and 0.18 M KCl. Inspection of the curves displayed in Figure 17 indicates that the dissociation of polymerase

Figure 16. Dissociation kinetics of lacUV5 open complexes. Complexes were formed by incubating lacUV5 DNA with RNA polymerase at 37°C. Thirty minutes later, heparin was added and the solutions were incubated for 0 to 90 min (as indicated, lanes 1-7) before being loaded atop a 5% polyacrylamide gel.

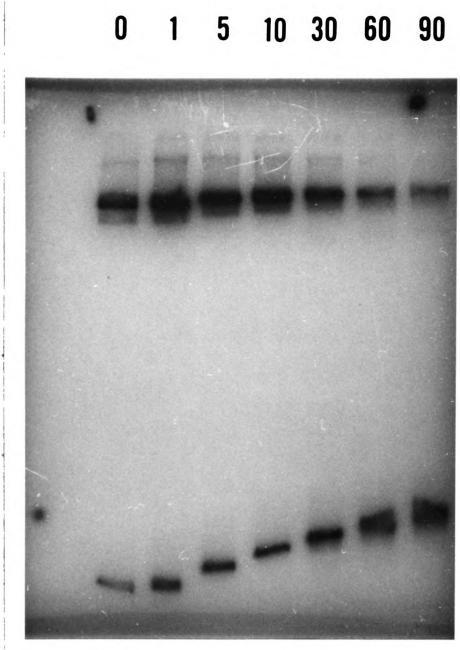
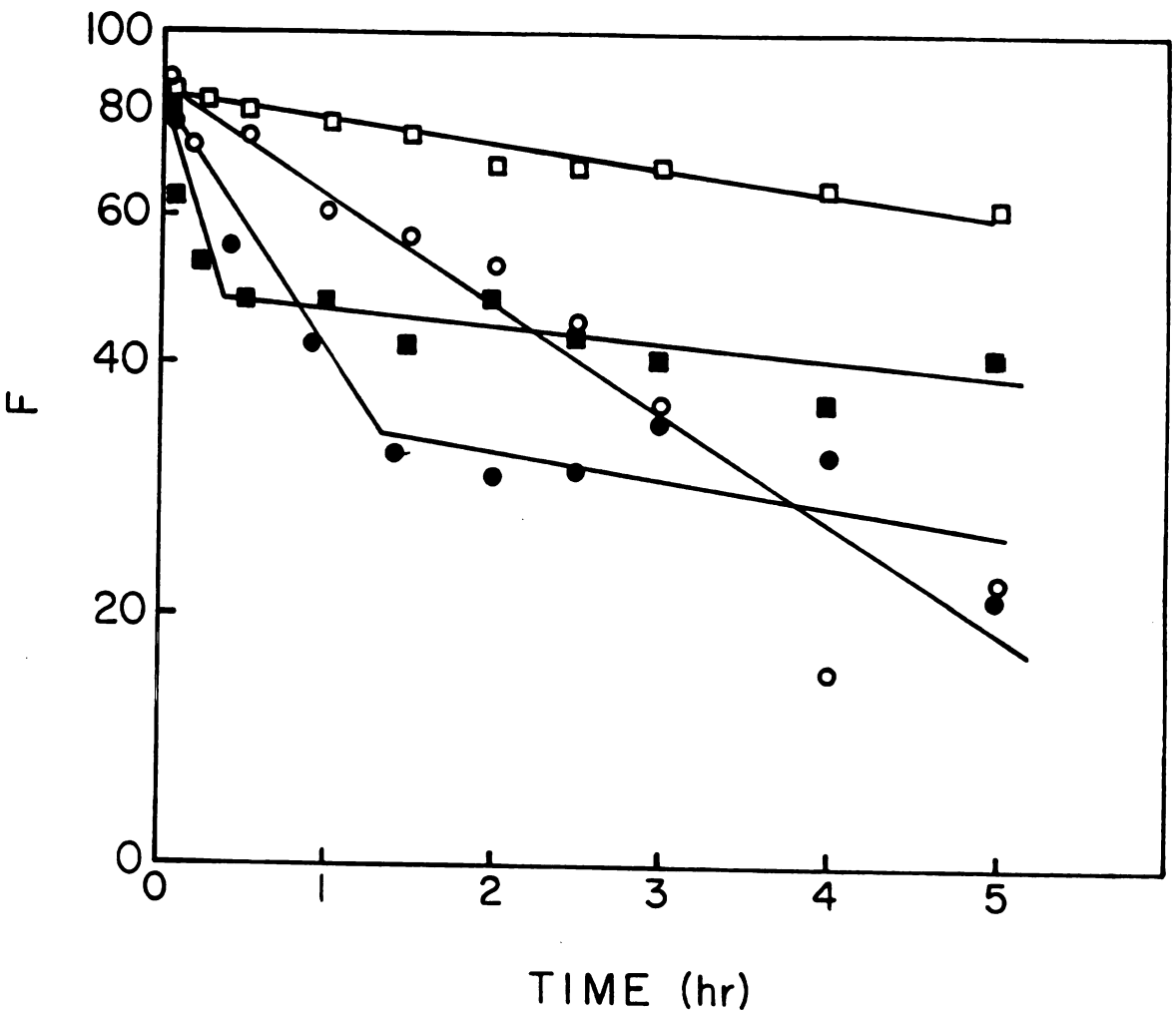

0 1 5 10 30 60 90

Figure 17. First order dissociation kinetics plot of *lac*UV5 open complexes. Complexes from gels such as the one in Figure 16 were quantified by scintillation counting then plotted as ln F vs time, where F is the ratio of complexes at the indicated time to the level of complexes at time zero. Binding and dissociation were performed at 0.075 M KCl (□), 0.100 M KCl (○), 0.140 M KCl (●), and 0.180 M KCl (■).


100
80
60
40
20
0
F
0
1
2
3
4
5
TIME (hr)

from the lacUV5 promoter is more complicated than expected. As the salt concentration of the solution rises from 0.075 M to 0.1 M KCl the value of k_{-2} increases from $1.8x10^{-5}$ sec^{-1} to $8.3x10^{-5}$ sec^{-1} (Table 2). However, at 0.14 M and 0.18 M KCl the dissociation becomes biphasic, with two dissociation rates; an early fast phase (k'_{-2}) is followed by a slow phase (k''_{-2}). The fast phases continue to increase in rate to about $2.0x10^{-4}$ sec^{-1}. Surprisingly, the slow dissociating processes have off rates close to that seen for the lowest salt used ($2.2x10^{-5}$ sec^{-1} and $1.4x10^{-5}$ sec^{-1}). The kinetics of transcript formation were measured at 0.1 M and 0.14 M KCl and found to conform to results seen using gel retardation assays; thus both phases at 0.14 M KCl are transcriptionally active (Ronald Ellis, pers. comm.; data not shown).

Salt Concentration Shift Experiments. The appearance of the biphasic off rate at higher [KCl] was a surprise. To probe this system further binding reactions were assembled at 0.1 M (or 0.14 M) KCl and allowed to incubate in the normal fashion. Simultaneous with the addition of heparin the salt concentration was shifted to 0.14 M (or 0.1 M) KCl. The dissociation rates observed are listed in Table 2. In these experiments we find that the dissociation rate corresponds to the salt concentration in the final solution. Thus, the rate of dissociation is not dependent on the salt

Table 2.

Rate of Dissociation of RNA Polymerase from the *lac*UV5 Promoter at Various Concentrations of KCl.

[KCl] M	k_{-2} (sec^{-1})	k'_{-2} (sec^{-1})	k''_{-2} (sec^{-1})
0.075[1]	1.8×10^{-5}	-	-
0.100[1]	8.3×10^{-5}	-	-
0.100[2]	5.0×10^{-5}	-	-
0.140[1]	-	1.9×10^{-4}	2.2×10^{-5}
0.140[3]	-	1.3×10^{-4}	1.8×10^{-5}
0.180[1]	-	2.0×10^{-4}	1.4×10^{-5}

[1] Binding and dissociation were performed at the indicated [KCl].
[2] Initial binding was performed at 0.140 M KCl.
[3] Initial binding was performed at 0.100 M KCl.

concentration of the initial binding buffer; rather, the [KCl] of the solution after the addition of heparin determines the character of the dissociation process.

DISCUSSION

The dissociation behavior of RNA polymerase at the *lac*UV5 promoter is quite complicated. At lower salt concentrations a single first order dissociation rate constant is seen. At moderately higher levels of salt the off rate becomes biphasic. The data from Figure 17 imply that there are two dissociating species of RNA polymerase. Two simple models can be proposed to account for this behavior: (1) there are two, mutually exclusive, binding sites which bind polymerase and have different dissociation properties or (2) there are two forms of polymerase which bind to DNA and dissociate at different rates. These models will necessarily give rise to different conclusions about the nature of the interaction of RNA polymerase at *lac* or other promoters.

The model of two distinct binding sites is not well substantiated by the experimental data. At lower salt concentrations such as 0.075 M or 0.1 M KCl, the affinity of RNA polymerase for DNA is known to be quite strong, while at higher salt concentrations, the binding becomes less stable. Therefore, one would expect that at high salt a polymerase molecule would have greater freedom to move between the two

putative sites. The observation, however, is that in a high salt medium the polymerase appears to be restricted in its movement (i.e., biphasic behavior is observed). A variation on this model is that the polymerase is not in equilibrium between the two sites, rather the individual off rates vary with the salt concentration. Thus, at low salt, dissociation rates from the two sites would be nearly equal, i.e., $k'_{-2} = k''_{-2}$, but when the [KCl] is raised the constants become unequal; $k'_{-2} > k''_{-2}$. There is no physical evidence for two binding sites at the lacUV5 promoter. Runoff transcription assays indicate that transcription occurs from what appears to be a single start site at all salt concentrations, and footprinting experiments imply that polymerase binds at a unique site (data not shown).

An alternative model hypothesizes two forms of polymerase. At low salt the two fractions of enzyme would have equal dissociation rates from promoter DNA, while at high salt the off rates differ. Both forms of the enzyme appear to be active in transcription. This model accounts for the fact that only a single start site is used. There is no evidence yet as to what the differences may be between the two forms. Phosphorylated polymerase has been detected in 2-D gels after labelling cell lysates with ^{32}P (80). Other modifications of the protein could occur as well; what effect these might have on polymerase function is unknown, but certainly an altered binding ability is possible.

Straney and Crothers have reported some unusual complexes at the lacUV5 promoter (81). At low temperature they see two types of complexes (O_u and O_l) with slightly different mobilities in polyacrylamide gels. The two forms are in equilibrium in solution but not in the gel. Both forms are active in transcription assays yet have different abilities to escape from the promoter. The complexes are indistinguisible by footprinting implying that they share some similarities in binding. Our complexes are not O_u and O_l since we work strictly at 37°C, but Straney and Crothers data imply that our results are certainly worth pursuing.

Chapters II and III of this thesis deal with the fact that under our assay conditions we do not observe complete binding of polymerase to wild type lac promoter fragments. From the studies reported there it was concluded that, while this effect might arise from the presence of a third RNA polymerase binding site, "P3", in addition to P1 and P2, no evidence for a P3 site could be found. Thus, it appears that a subset of inactive polymerase molecules may interfere with stable binding at the wild type promoter. The biphasic dissociation behavior seen for lacUV5-RNA polymerase complexes could support such an explanation. Those polymerase molecules which at higher salt have a short half life at the UV5 promoter could represent a subset of the enzyme molecules which are less stable at the wild type promoter. The dissociation rate of polymerase from the wild

type promoter was measured at 0.1 M KCl and found to be $9.8x10^{-5}$ sec^{-1}. This value is close to that seen for UV5 at this salt. However, the maximum level of stable binding at the wild type promoter was, at most, about 50%. The balance of the promoters might be bound by the fast dissociating polymerase species. The incomplete binding of polymerase at the wild type promoter is unique. In addition to the lacUV5 promoter, the gal P1 and P2 promoters, and the lambda P_r promoter are able to form complexes on virtually all of the fragments. Interestingly, the lac wild type promoter is able to bind polymerase nearly 100% on supercoiled templates (68). Perhaps the less active polymerases are insensitive to supercoiling at this promoter.

Studies of dissociation kinetics have revealed a new level of complexity in the interaction of RNA polymerase with the lac promoter. The presence of two phases of dissociation suggests the possiblity that polymerase may find itself in more than one state in solution. The development of improved separation methods capable of producing a preparation of 100% active protein molecules would be of interest. A more thorough study of the nature of the enzyme when it is bound to the promoter would also clarify this issue.

SUMMARY

SUMMARY

A variety of techniques have been used to study the interaction of RNA polymerase with lactose promoter DNA. During the course of this work a number of interesting observations were made:

(1) RNA polymerase will bind to both the P1 and P2 promoters in the absence of cAMP/CAP.

(2) A closed complex (heparin sensitive) at the P1 site is quite stable with a half life of several hours; the addition of CAP to such complexes brings about the conversion to the open form.

(3) Exonuclease III digestions indicate that RNA polymerase protects DNA downstream of the P1 promoter. We call this other site P3. We proposed that deletion of DNA downstream of +1 could remove this binding site.

(4) The deletion analysis of the lac promoter indicates that P3 is not downstream of +1, rather it may be between P1 and P2 (making its deletion impossible).

(5) The dissociation of RNA polymerase from the mutant lacUV5 promoter is complex. At salt concentrations above 0.1 M KCl the off rate becomes biphasic, indicating more than one binding site for the enzyme or more than one form of polymerase binding at a single site.

(6) The DNA sequences required to achieve proper utilization of the lac promoter were investigated; it appears that DNA regions upstream of the -35 and downstream of the -10 consensus sequences are not necessary to define a promoter.

(7) Finally, the nucleotide used for initiation of transcription at lac can be any of the four possible bases. The efficiency of transcription initiation at the lacUV5 promoter is the same regardless of the first base in the transcript.

These results indicate the complexity of RNA polymerase interactions with promoter DNA sequences. The results of Chapter II indicate that RNA polymerase forms a stable, closed complex at the wild type P1 promoter when no CAP is present. This result is possibly at odds with previously accepted models. Prior studies were interpreted to imply that CAP is required for the proper binding of polymerase during the initiation process. My data indicate that CAP can act after closed complex formation by enhancing the

isomerization step. Moreover, Gralla has recently proposed that at the glnALG promoter RNA polymerase first binds to form a closed complex which is then converted to the open form upon binding of the activator protein, glnG (82). Such a mechanism is analogous to that which I have observed for CAP. Of course overall CAP stimulation may involve action at both the binding and isomerization steps. Additional kinetic analyses may help to clarify this issue.

The issues of promoter structure and utilization raised in Chapter III provide a basis for the possible use of E. coli promoters in the synthesis of RNA molecules for physical analysis. The promoter specificity of polymerases from the phages T7 and SP6 makes them desirable for use in the production of mRNAs in vivo and in vitro. However, the stringent requirements for specific sequences near to +1 renders their utility for the production of specific RNA molecules potentially difficult. The altered lacUV5 promoters which I designed may be of use for the production of a RNA molecule without extraneous sequences supplied by the vector. In practice, the DNA sequence corresponding to any RNA sequence can be synthesized and cloned into the vector at the EcoRI site at -7; the mRNA made from such clones wil start precisely at +1, regardless of the nucleotide there.

Much of the work described here deals with the apparent inability of RNA polymerase to form stable complexes with

the wild type lac promoter on more than about 50% of the DNA fragments. The evidence presented in Chapter II supports the notion of a third binding site referred to as "P3". The interaction of polymerase at P3 would effectively block P1 and P2, thus, reducing the level of binding at these productive sites. It was proposed that the complex formed at P3 is closed and heparin sensitive, yet is stable enough to discourage redistribution of the polymerase molecules. Since the results of Chapter III indicate that deletion of the DNA downstream of -7 cannot abolish P3 binding it was proposed that either P3 is between P1 and P2 ,or that some fraction of the RNA polymerase preparation is unable to form an open complex thus reducing the total number of complexes. Chapter IV summarizes experiments on the dissociation of RNA polymerase from the lacUV5 promoter. The off rate for polymerase is biphasic at high salt concentration indicating the possibility of two polymerase binding sites or two forms of the enzyme. Again, additional knowledge of the conformational states of the polymerase appears to be necessary for complete understanding of the transcription process.

APPENDIX

APPENDIX

The nucleotide sequence of the 203 base pair lac promoter fragment used in this dissertation.

```
                                                                -140
                                                                 .
                                                                CCGATTC
                                                                GGCTAAG

ATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAA
TAATTACGTCGACCGTGCTGTCCAAAGGGCTGACCTTTCGCCCGTCACTCGCGTTGCGTTAATT

          [ -35 ]         P2        [ -10 ]
TGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTG
ACACTCAATCGAGTGAGTAATCCGTGGGGTCCGAAATGTGAAATACGAAGGCCGAGCATACAAC
[      CAP       ]                [ -35  ]       P1       [ -10  ]

TGTGGAATTGTCTCGCCTATTGTTAAAGTGTGTCCTTTGTCGATACTGGTACTAATGCCTAAGTG
ACACCTTAACAGAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATGATTACGGATTCAC

 +63
  .
TGG
ACC
```

LIST OF REFERENCES

References

1. Jacob, F. and Monod, J. (1961) J. Mol. Biol. 3, 318-356.

2. Monod, J. (1947) Growth 11, 223-289.

3. Magasanik, B. (1961) Cold Spring Harbor Symp. Quant. Biol. 26, 248-256.

4. Makman, R.S. and Sutherland, E.Q. (1965) J. Biol. Chem. 240, 1309-1314.

5. Ullman, A. and Monod, J. (1968) FEBS Lett. 2, 57-60.

6. Anderson, W.A., Schneider, A.B., Emmer, M., Perlman, R.L.,and Pastan, I. (1971) J. Biol. Chem. 246, 5929-5937.

7. Jacob, F., Ullman, A., and Monod, J. (1964) C. R. Acad. Sci. 258, 3125-3128.

8. Reznikoff, W.S., Maquat, L.E., Munson, L.M., Johnson, R.C., and Mandecki, W. (1982) in Promoters: Structure and Function, Rodriguez, R.L., and Chamberlin, M.J. eds., pp. 80-95, Praeger Publishers, New York.

9. Malan, T.P. and McClure, W.C. (1984) Cell 39, 173-180.

10. Burgess, R.R., Travers, A.A., Dunn, J.J., and Bautz, E.K.F. (1969) Nature 221, 43-46.

11. Burgess, R.R. (1969) J. Biol. Chem. 244, 6168-6176.

12. Burgess, R.R., Erickson, B., Gentry, D., Gribskov,M., Hager, D., Lesley, S., Strickland, M., and Thompson, N. (1986) in RNA polymerase and the Regulation of Transcription, Reznikoff, W.S., Burgess, R.R., Dahlberg, J.E., Gross, C.A., Record, M.T., Jr., and Wickens, M.P. eds., pp 3-15, Elsevier, NY.

13. Siebenlist, U., Simpson, R.B., and Gilbert, W. (1980) Cell 20, 269-281.

14. Travers, A.A. and Burgess, R.R. (1969) Nature 222, 537-540.

15. Kornberg, A. (1974) DNA Synthesis, pp 329-347, W. H. Freeman and Co., San Francisco, CA.

16. Minkley, E.G. and Pribnow, D. (1973) J. Mol. Biol. 77, 255-277.

17. Wehrli, D. and Staehlin, M. (1971) Bact. Rev. 35, 290-309.

18. Zillig, W., Palm, P., and Heil, A. (1976) in RNA Polymerase, Losick, R. and Chamberlin, M. eds, pp 101-125, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

19. Rohrer, H., Zillig, W., and Mailhammer, R. (1975) Eur. J. Biochem. 60, 227-238.

20. Helmann, J.D. and Chamberlin, M.J. (1988) Ann. Rev. Biochem. 57, 839-872.

21. Landick, R., Vaughn, V., Lau, E.T., van Bogelen, R., Erickson, J.W., and Neinhardt, F.C. (1984) Cell 38, 175-182.

22. Hirschman, J., Wong, P.-K., Sei, K., Keener, J., and Kustu, S. (1985) Proc. Natl. Acad. Sci. USA 82, 7527-7529.

23. Chambers, D.A. and Zubay, G. (1969) Proc. Natl. Acad. Sci. USA 63, 118-122.

24. Garner, M.M. and Revzin, A. (1982) Biochemistry 21, 6032-6036.

25. Shanblatt, S.H. and Revzin, A. (1983) Proc. Natl. Acad. Sci. USA 80, 1594-1598.

26. Meiklejohn, A.L. and Gralla, J.D. (1985) Cell 43, 769-776.

27. Majors, J. (1975) Nature 256, 672-674.

28. Taniguchi, T., O'Neill, M., and deCrombrugghe, B. (1979) Proc. Natl. Acad. Sci. USA 76, 5090-5094.

29. Busby, S., Aiba, H., and deCrombrugghe, B. (1982) J. Mol. Biol. 154, 211-227.

30. Salemme, F.R. (1982) Proc. Natl. Acad. Sci. USA 79, 5263-5267.

31. Liu-Johnson, H.-N, Gartenberg, M.R., and Crothers, D.M. (1986) Cell 47, 995-1005.

32. Wu, H.-M. and Crothers, D.M. (1984) Nature 308, 509-513.

33. Musso, R.E., DiLauro, R., Adhya, S., and deCrombrugghe, B. (1977) Cell 12, 847-854.

34. Buc, H., Amouyal, M., Buckle, M., Herbert, M., Kolb, A., Kotlarz, D., Menedez, M., Rimsky, S., Spassky, A., and Yeramian, E. (1986) in RNA polymerase and the Regulation of Transcription, Reznikoff, W.S., Burgess, R.R., Dahlberg, J.E., Gross, C.A., Record, M.T., Jr., and Wickens, M.P. eds, pp 115-125, Elsevier, NY.

35. Menedez, M., Kolb, A., and Buc, H. (1987) EMBO J. 6, 4227-4234.

36. Sheppard, D.E. and Englesberg, E. (1967) J. Mol. Biol. 25, 443-454.

37. Schleif, R. and Lis, J.T. (1975) J. Mol. Biol. 95, 417-431.

38. Hahn, S., Dunn, T., and Schleif, R. (1984) J. Mol. Biol. 180, 61-72.

39. Hahn, S., Hendrickson, W., and Schleif, R. (1986) J. Mol. Biol. 188, 355-367.

40. Hawley, D.K. and McClure, W.S. (1983) Nucl. Acids Res. 11, 2237-2255.

41. Harley, C.B. and Reynolds, R.P. (1987) Nucl. Acids Res. 15, 2343-2361.

42. deBoer, H.A., Comstock, L.J., and Vasser, M. (1983) Proc. Natl. Acad. Sci. USA 80, 21-25.

43. Raibaud, O. and Schwartz, M. (1984) Ann. Rev. Genet. 18, 178-206.

44. Cowing, D., Bardwell, C.A., Craig, E.A., Woolford, C., Hendrix, R., and Gross, C.A. (1985) Proc Natl. Acad. Sci. USA 82, 2679-2683.

45. Auble, D.T. and deHaseth, P.L. (1988) J. Mol. Biol. 202, 471-482.

46. Stefano, J.E. and Gralla, J.D. (1982) Proc. Natl. Acad. Sci. USA 79, 1069-1072.

47. Lorimer, D.D. and Revzin, A. (1986) Nucl. Acids Res. 14, 2921-2938.

48. Yu, X.-M. and Reznikoff, W.S. (1986) J. Mol. Biol. 188, 545-553.

49. Deuschle, U., Kammerer, W., Gentz, R., and Bujard, H. (1986) EMBO J. 5, 2987-2994.

50. Kammerer, W., Deuschle, U., Gentz, R., and Bujard, H. (1986) EMBO J. 5, 2995-3000.

51. Milligan, J.F., Groebe, D.R., Witherell, G.W., and Uhlenbeck, O.C. (1987) Nucl. Acids Res. 15, 8783-8798.

52. Kang, C. and Wu, C.-W. (1987) Nucl. Acids Res. 15, 2279-2294.

53. Melton, D.A., Krieg, P.A., Rabagliati, M.R., Maniatis, T., Zinn, R., and Green, M.R. (1984) Nucl. Acids Res. 12, 7035-7056.

54. Garner, M.M. and Revzin, A. (1981) Nucl. Acids Res. 9, 3047-3060.

55. McClure, W.R., Cech, C.L., and Johnston, D.E. (1978) J.Biol. Chem. 253, 8941-8948.

56. Schmitz, A. and Galas, D.J. (1979) Nucl. Acids Res. 6, 111-137.

57. Tullius, T.D. and Dombrowski, B.A. (1985) Science 230, 679-681.

58. McClure, W.R. (1980) Proc. Natl. Acad. Sci. USA 77, 5634-5638.

59. Strickland, S.,Palmer, G., and Massey, V. (1975) J. Biol. Chem. 250, 4048-4052.

60. Burgess, R.R. and Jendrisak, J.J. (1975) Biochemistry 14, 4634-4638.

61. Lowe, P.A., Hager, D.A., and Burgess, R.R. (1979) Biochemistry 18, 1344-1352.

62. Boone, T. and Wilcox, G. (1978) Biochim. Biophys. Acta 541, 528-534.

63. Chamberlin, M.J., Nierman, W.C., Wiggs, J., and Neff, N. (1978) J. Biol. Chem. 254, 10061-10069.

64. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) Molecular Cloning-A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

65. Felsenfeld, G. and Hirschman, S.Z. (1965) J. Mol. Biol. 13, 407-427.

66. Maxam, A.M. and Gilbert, W. (1980) Methods Enzymol. 65, 499-560.

67. Peterson, M.L. and Reznikoff, W.S. (1985) J. Mol. Biol. 185, 535-543.

68. Malan, T.P., Kolb, A., Buc, H., and McClure, W.R. (1984) J. Mol. Biol. 180, 881-909.

69. Shanblatt, S.H. and Revzin, A. (1986) J. Biol. Chem. 261, 10885-10890.

70. Spassky, A., Busby, S., and Buc, H. (1984) EMBO J. 3, 43-50.

71. deCrombrugghe, B., Busby, S., and Buc, H. (1984) Science 224, 831-838.

72. Yu, X.-M. and Reznikoff, W.S. (1984) Nucl. Acids Res. 13, 5449-5464.

73. Peterson, M.L. and Reznikoff, W.S. (1985) J. Mol. Biol. 185, 525-533.

74. deHaseth, P.L., Lohman, T., Burgess, R.R., and Record, M.T., Jr. (1978) Biochemistry 17, 1612-1622.

75. Beckwith, J., Grodzicker, T., and Arditti, R. (1972) J. Mol. Biol. 69, 155-160.

76. Beaucage, S.L. and Caruthers, M.H. (1981) Tet. Letters 22, 1859-1862.

77. Zoller, M.J. and Smith, M. (1983) Methods Enzymol. 100, 469-500.

78. Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) Gene 33, 103-119.

79. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.

80. Enami, M. and Ishihama, A. (1984) J. Biol. Chem 259, 526-533.

81. Straney, D.C. and Crothers, D.M. (1985) Cell 43, 449-455.

82. Sasse-Dwight, S. and Gralla, J.D. (1988) Proc. Natl. Acad. Sci. USA 85, 8934-8938.