PH.D.

1993

This is to certify that the

dissertation entitled

Studies of Control of Transcription at the E. coli lac Promoter and the Mechanism of Activation by the Catabolite Activator Protein

presented by

Hongyun Yu

has been accepted towards fulfillment of the requirements for

Ph.D. degree in Genetics

Arnold Revzin
Major professor

Date Oct. 5, 1993

STUDIES OF CONTROL OF TRANSCRIPTION AT THE *E. COLI LAC* PROMOTER AND THE MECHANISM OF ACTIVATION BY THE CATABOLITE ACTIVATOR PROTEIN

By

Hongyun Yu

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Genetics Graduate Program

1993

ABSTRACT

STUDIES OF CONTROL OF TRANSCRIPTION
AT THE *E. COLI LAC* PROMOTER AND THE MECHANISM
OF ACTIVATION BY THE CATABOLITE ACTIVATOR PROTEIN

By

Hongyun Yu

To better understand the control of bacterial gene expression, I have performed experiments on the function of catabolite activator protein (CAP) in the initiation of transcription at the *E. coli* lactose operon. The *lac* promoter contains overlapping binding sites for RNA polymerase. In the presence of CAP-cAMP, RNA polymerase binds at a site called P1 and initiates transcription from nucleotide +1; in the absence of CAP-cAMP, the enzyme binds at an alternate position (P2) and begins transcription from -22 (i.e., 22 nucleotides upstream). Solutions of *E. coli* RNA polymerase and CAP with different *lac* promoter fragments, as well as complexes purified by gel electrophoresis, were probed by gel retardation assays, DNase I footprinting, and exonuclease III digestion. Other techniques used include runoff transcription and abortive initiation experiments. The stoichiometry of CAP in a variety of complexes was also determined. The results indicate that longer promoter fragments have a higher

affinity for CAP. A region around -250 appears to be critical. Complexes of CAP with lac DNA fragments extending from -310 to +60 have two cooperatively bound CAP molecules, while complexes with DNA containing lac sequence from -173 to +60 contain only one CAP. It appears that CAP can interact with lac DNA at around -250, -61, and +15.

All fragments used in these studies are capable of participating in two different stable P1 complexes when incubated with CAP-cAMP and polymerase. A characteristic of these transcriptionally-competent complexes is their resistance to attack by heparin. One such complex (the ternary complex) contains both proteins; the other contains no CAP. Ternary P1 complexes formed with DNA fragments containing the -250 site show elevated transcription efficiency when compared with their CAP-free counterparts. However, there is no significant difference in transcription seen between the two P1 complexes at fragments that do not have the -250 sequence.

A new model for the molecular mechanisms involved in CAP activation of transcription initiation at the lac promoter has emerged from this work: (1) a CAP molecule binds at -61 to ensure that polymerase exclusively binds at P1--there are important protein-protein interactions between this CAP and RNA polymerase; (2) a second CAP molecule bridges the +15 and -250 sites, bringing them into close proximity via a loop structure (by interacting simultaneously with both sites); (3) the loop structure, together with the CAP at

-61, stimulates transcription initiation. These results are consistent with other studies that show that the interactions of transcriptional regulatory proteins at lac, and at other operons as well, are more complicated than one might expect.

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. Jack Preiss, Steven Triezenberg, and Jerry Dodgson for their encouragement through the years. I also thank my labmates -- Dr. Jianli Cao, Mark Sinton, and Diane Cryderman for their friendship and assistance.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION AND LITERATURE REVIEW

INTRODUCTION

The control of gene expression in many systems is achieved at the level of transcription. The lactose operon of Escherichia coli is the prototype transcriptionally regulated system. The operon consists of three genes that code for proteins needed for metabolism of lactose, plus a region of DNA upstream that is required for control of transcription of those genes. Jacob and Monod, and others, used a combination of physical and genetic approaches to deduce the primary molecular events in control of transcription at lac (1-5).

Like many other operons, the lac operon has both negative and positive control elements. When the lac enzymes are not needed (e.g., there is no lactose in the growth medium), its mRNA synthesis is repressed by a protein, the lactose repressor, which binds to a sequence of DNA known as the "operator". If the enzymes are needed, the affinity of repressor for the operator is reduced by the binding of a small inducer molecule (allolactose) to the protein. Allolactose is a metabolic intermediate in lactose metabolism. Thus, if lactose is the only carbon source for E. coli, a repressor-inducer complex is formed, which has a reduced affinity for the operator, and induction can occur.

In this situation, RNA polymerase, which binds to the control region of DNA at the "promoter" region, catalyzes synthesis of the transcript, and the lac enzymes are made.

The existence of a positive regulator at the lactose operon is indicated by the observation that if cells are grown on a medium containing both glucose and lactose, the lac enzymes are not made, even though repressor is not bound to the operator (4). This regulatory factor is known as the catabolite activator protein (CAP); it is also called the cAMP receptor protein (CRP) (3). CAP, in complex with cAMP, is required for efficient initiation of transcription at the lactose operon, and at a number of other operons as well (6). The repression of lac mRNA synthesis by glucose is related to the concentration of cAMP, which is reduced in the presence of glucose (7). At low cAMP levels, the CAP-cAMP complex is not formed and CAP is not active in stimulating transcription.

The molecular mechanisms whereby the control proteins function have been the subject of much study. These proteins control transcription by binding at specific sites on DNA and interacting with components of the transcription apparatus, including promoter sequences, RNA polymerase, and possibly other transcription factors. At lac, repressor binds between RNA polymerase and the genes, hence can prevent transcription by physically blocking polymerase movement, or perhaps by competing with the enzyme for a

partially overlapping binding site. CAP binds upstream of polymerase, and presumably stimulates transcription by contacting the enzyme and causing a conformational change that makes polymerase more efficient at initiation or at escape from the promoter.

Recent studies have revealed new features of bacterial promoters that are of much interest. Many promoters have overlapping, mutually exclusive binding sites for RNA polymerase (8-10); these overlapping promoters may, or may not, be active in initiation of transcription. One aspect of CAP function may involve its excluding RNA polymerase from an inactive site, which would otherwise compete for the enzyme and dilute its effectiveness. There are also a number of cases in which multiple copies of regulatory proteins are involved in control. At gal, for example, two CAP molecules are needed for activation of transcription (11). When the regulatory protein binding site is a distance away from the promoter region [e.g., the NtrC sequence in glnA (12)] or there are multiple binding sites for a specific protein [e.g., AraC in the araBAD region (13)], a looped DNA structure may be formed and play a regulatory role. The lactose operon has three repressor binding sites located at around -85, +11 and +415, (initiation of transcription is at +1); extensive studies have shown that lac DNA looping is involved in repressor binding and repression of transcription (14). Two operator

sequences in the gal operon of E. coli were found to be separated by 114 bp (15, 16). Loop formation between these sites also appears to account for the observed regulatory activity (17). It is interesting that in lac, a second weak CAP binding site has been identified (18), which is centered at +15 and overlaps the primary operator sequence around +11. The function of this CAP site has been unclear, nor is it known whether CAP binds to that site in the cell.

DNA bending is another feature often associated with transcriptional control. Some DNA sequences have a natural curvature, others may be bent when a protein is bound. Clearly the bending of DNA can facilitate loop formation, as well as, potentially, positioning proteins favorably for interactions with each other and with the promoter DNA sequences.

In this thesis are presented data that relate to some unusual CAP-DNA complexes that can form under certain conditions at the lac promoter, and to the formation of DNA loops involving CAP. Thus these studies provide new data on current issues in transcription in prokaryotic systems.

LITERATURE REVIEW

RNA POLYMERASE

The DNA-dependent RNA polymerase holoenzyme from E. coli is composed of four different subunits, α, β, β', and σ, in a ratio of 2:1:1:1 (19, 20). Core polymerase, which contains the polymerization activity, has the structure $\alpha_2\beta\beta'$ (=E). The σ subunit is required for proper recognition of specific promoter sequences (19).

The amino acid sequence of each subunit has been deduced from sequence analysis of the corresponding gene; the molecular weights were determined accordingly as: α = 36,511 D; β = 150,615 D; β'= 155,159 D; and σ = 70,262 D; therefore, the mass of the holoenzyme, $E\sigma^{70}$, is 449,058 D (21). In many bacteria, there is a primary form of RNA polymerase (in E. coli this is $E\sigma^{70}$); alternate forms have the same core composition but different σ factors (22), and recognize different promoters.

The transcription initiation process involves at least two major steps: 1) the holoenzyme (R) recognizes the promoter sequence (P) and binds to the double helical DNA, forming a "closed" complex $(RP)_c$; 2) RNA polymerase "melts" into the two strands of DNA forming a stable "open" complex $(RP)_o$, allowing for the transcribing of DNA from one strand

into RNA.

$$R + P \rightleftarrows (RP)_c \rightarrow (RP)_o.$$

This is an oversimplified description of the initiation process. In fact, a number of intermediate complexes have been identified, implying that the closed complex passes through several stages as the open complex is formed (23, 24). The open complex contains a stretch of single stranded DNA in the region from about -10 to +3 (25), where +1 is the start point of transcription, and positive numbers represent downstream nucleotides.

When nucleoside triphosphates are added to open complexes *in vitro*, RNA polymerase catalyzes the synthesis of phosphodiester bonds between nucleotides that are complementary to the DNA template. At many bacterial promoters, "abortive" transcription occurs (26, 27), in which an oligonucleotide is released and RNA polymerase (without dissociating from the promoter) catalyzes production of a new RNA molecule from +1. "Escape" of polymerase from the promoter, with subsequent elongation of the full-length transcript, begins when σ factor dissociates from the core enzyme when the RNA molecule is 8-12 nt long (28); core polymerase catalyzes the elongation process. It has been reported that *E. coli* σ factor loses contact with the 3′ end of the nascent RNA after synthesis of a tetranucleotide (29). In the elongating transcription complex, the 3′ proximal 8 nucleotides of RNA are paired

with DNA template (30). After termination of transcription the core enzyme is released and may reassociate with a free σ factor (31).

Unlike DNA polymerase, RNA polymerase does not need a primer to initiate polymerization. Mechanistically, however, it is convenient to consider the enzyme to contain two action sites: a "primer terminus" site (at the end of the growing RNA chain) and an "elongation site" (32). Prior to initiation, both these sites are empty though the enzyme can be stably bound at the promoter in an open complex. As initiation begins, each site binds one nucleotide (complementary to the corresponding bases in the DNA template). There follows a nucleophilic attack on the α-phosphate of the nucleotide at the elongation site by the 3′-hydroxyl of the nucleotide at the primer terminus site, 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 that was added from the elongation site now resides in the primer terminus site, and the process is repeated.

The roles of the individual subunits have been elucidated by use of inhibitors of transcription and by studies of polymerase mutations. Two well known classes of antibiotics act on the β 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 (33). After the first two nucleotides are polymerized, rifampicin is no longer an effective inhibitor, implying that the antibiotic blocks entry of the initiating nucleotide. The streptolydigins block elongation presumably by binding to the elongation site (34). Polymerase mutations which confer resistance to these antibiotics reside in the β subunit, indicating that this subunit may contain the nucleotide binding sites. Additional mutants have shown that the β subunit of polymerase is involved both in catalytic function and in DNA binding of the core enzyme (35, 36). Heparin, an initiation inhibitor used in this study, apparently acts by competing with DNA for polymerase (i.e, heparin may behave like a DNA substitute). The β' 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 (37). In the presence of heparin, RNA polymerase will not form new complexes with DNA, either at specific promoter sites or at other nonspecific sequences. Heparin rapidly destabilizes preformed nonspecific complexes but does not affect open complexes that have formed.

The α subunit exists in solution as a dimer, which may be a prototype for the α-α structure in the enzyme. α subunit is ADP-ribosylated during infection of E. coli by bacteriophage T4 (38). The modified polymerase has a reduced affinity for promoters, indicating the involvement

of the α subunit in the DNA binding function of the enzyme either by direct interaction or via allosteric effects on the other subunits. More recent research has shown that the α subunit is important in polymerase interaction with the catabolite activator protein (CAP) (39, 40).

Bacterial cells generally use one primary sigma factor; in E. coli, for example, $E\sigma^{70}$ is involved in transcription of most genes during exponential cell growth (the superscript 70 refers to the molecular weight of sigma, 70,000 D). In addition, the bacterium utilizes alternative sigma factors to direct RNA polymerases to particular regulons. That is, holoenzymes with different σs have different promoter sequence specificities (22, 41). Such specialized RNA polymerases permit the cell to express coordinated sets of genes during different physiological conditions or developmental stages. Examples are σ^{32}, which allows polymerase to function at heat shock promoters (42), and ntrA (σ^{54}) which is required for transcription of some nitrogen-regulated genes (43). The heat shock promoters have DNA consensus sequences around -35 similar to those of σ^{70} promoters (see below), but have very different -10 sequences, implying that sigma factor may guide polymerase to the promoter by recognition of the -10 region (44). σ^{70} functional domains have been characterized by deletion mutations. σ^{70} has a domain required for binding to core polymerase, as well as amino acid sequences needed for -10

recognition. In between these two regions is a segment of σ^{70} that seems to be involved in open complex formation. The -35 recognition sequence is located at the C-terminal end (45). Although σ factor has two putative helix-turn-helix DNA-binding motifs, it is not clear that σ itself binds DNA. It has been proposed that the N-terminal sequences inhibit the ability of intact σ^{70} to bind to DNA when it is not associated with core polymerase.

Sigma factors can be thought of as ligands that confer promoter specificity on core polymerase. For example, core enzyme has a strong affinity for random DNA sequences, forming a nonspecific complex with a binding constant of ~10^{11} M^{-1} at 0.1 M KCl. Holoenzyme binds nonspecifically with a smaller equilibrium constant, ~10^{7} M^{-1}, yet binds specifically to a promoter with a constant of ~10^{14} M^{-1}. Thus, σ^{70} acts to destabilize nonspecific binding by holoenzyme, while stabilizing specific complexes, most likely through allosteric interactions with the individual core subunits (46).

THE CATABOLITE ACTIVATOR PROTEIN

The catabolite activator protein (CAP), when associated with cAMP, stimulates transcription at *lac* and many other operons, which are generally involved in carbon metabolism. Purified CAP has been shown to enhance *lac* transcription *in*

vitro approximately 30-fold (47). The stimulatory action of CAP is strictly cAMP-dependent. In principle, CAP can act at any of the stages involved in initiation, such as by enhancing open complex formation, by reducing the level of (unproductive) abortive transcripts, or by facilitating escape of RNA polymerase from the promoter.

CAP is a dimer of identical subunits each with a molecular weight of 23,619 D. It is a basic protein with an isoelectric point of 9.1. Each subunit is composed of two domains: a large cAMP binding domain and a smaller DNA sequence recognition and binding domain. The DNA binding domain has a helix-turn-helix structure (48, 49) which is common to many DNA binding proteins. Each CAP dimer has two cAMP binding sites, but the $CAP(cAMP)_1$ form is believed to be the preferred active complex; the association constant of $CAP(cAMP)_1$ for the *lac* promoter is significantly higher than that of $CAP(cAMP)_2$ (50). Thus the CAP-cAMP-DNA interaction is reminiscent of systems that show "half-of-the-sites reactivity". CAP undergoes a conformational change upon cAMP binding which readies the complex for DNA binding (for review see 51). However, physiological concentrations of cAMP also stabilize the CAP dimer with respect to subunit exchange in either the presence or the absence of DNA (52).

Specific DNA sequences recognized by CAP vary somewhat among the CAP binding sites at different CAP-sensitive operons. Extensive studies of CAP binding to a number of

these sites show that the CAP dimer covers at least 22 base pairs, and comparison of the nucleotide sequences of known CAP binding sites suggests the symmetric consensus sequence: 5′ AA-TGTGA------TCACA-TT 3′, where - represents any nucleotide. Within the CAP consensus sequence, the 5′ TGTGA 3′ pentanucleotide sequence in each half site is the most crucial element for CAP recognition (53, 54). Changes of 2 bp (one per half site) within TGTGA of the *E. coli lac* CAP site convert it to a FNR regulon sequence (55). A CAP binding sequence containing identical consensus half-sites has been synthesized; it binds CAP 450 times better than the natural *lac* CAP site (56). The crystal structure of a CAP complex with DNA has been solved (57). CAP recognizes its binding site by interacting specifically with both the bases and the phosphate backbone of the DNA consensus sequence (54).

There is no *a priori* reason to assume that CAP acts in precisely the same way at every promoter; on the contrary, it is clear that substantial differences exist. The center position of the CAP-binding site at various promoters varies from 40 to 200 bp upstream of the start point of transcription (+1) (58). The primary CAP binding site at the *lac* operon is centered at -61 (59), while at *gal* the primary CAP site is centered at -41 (60, 61), about 20 base pairs closer to the P1 initiation nucleotide. A second CAP binding site at *lac*, which is revealed *in vitro* in low salt

buffers, is at +15 (18). CAP binds at around -100 in the *ara* promoter (62). The stimulation of transcription by CAP apparently is achieved differently at different promoters. One CAP-cAMP complex can activate *lac* P1 transcription (59), but two CAP molecules are involved in *gal* P1 transcription (63). The second CAP at *gal* binds at precisely the same location relative to +1 as does the lone CAP at *lac* (64); perhaps it makes similar contacts with RNA polymerase. In some cases, CAP activation is assisted or antagonized by other regulatory proteins. MalT as a positive regulatory protein is required for the proper function of CAP at *mal* operons (65). CytR binds to its specific site at the *cytRP* promoter, antagonizes the CAP-induced DNA bending, and forms a repression complex (66). Furthermore, both CAP and RNA polymerase mutations which affect CAP activation of the *lac* but not the *gal* promoter have been isolated (67, 68), providing further evidence that CAP activation of polymerase transcription varies from promoter to promoter.

PROMOTERS

The sequence of DNA at which RNA polymerase binds and which specifies the start point of transcription is called the promoter. Comparison of more than 200 *E. coli* promoters has revealed two conserved sequences centered around -10 and -35 (69, 70). The -10 region has consensus sequence TATAAT,

and is separated from the -35 consensus sequence of TTGACA by 17±1 base pairs. Some promoters may have exact agreement with one or the other consensus sequences, but no naturally occurring promoter has been found with perfect homology in both regions.

The -10 and -35 regions are the two major determinants of promoter strength. Systematic base substitution in the -35 region of the CAP-independent mutant lacUV5 promoter confirmed that the -35 consensus sequence is very important for promoter function (71). RNA polymerase likely contacts the G residue in the -35 consensus sequence. Methylation of the G residue in the consensus sequence prevents open complex formation at $galP_{con}$ (in which the galP1 -35 region was replaced with the TTGACA consensus sequence). Interestingly, if the promoter has a -35 sequence that differs considerably from the consensus sequence (e.g., galP1), modifications of G residues around -35 reduce but do not prevent open complex formation (72).

An artificial promoter containing consensus sequences at both the -10 and -35 region is so strong in vivo that it is difficult to regulate (73). Moreover, one can imagine that a promoter sequence that binds RNA polymerase tightly could inhibit transcription by hindering escape of the enzyme. This is probably why nature does not create perfect promoters, since delicate regulation and balanced metabolism are vital for living systems. Promoters with poor homology

to the -35 consensus sequence are often positively controlled by other factors (74).

The spacer DNA between the two consensus sites is also important. Single base pair changes in the spacer show no mutant phenotype, thus it is likely that polymerase does not make specific contacts within the spacer. The length of the spacer DNA is important. Single base deletions or insertions in the spacer of promoters can cause severe alteration of promoter strength (75). Therefore, the spacer is thought to function by correctly positioning the -10 and -35 sequences for proper polymerase binding. When the spacer DNA is replaced by DNA with a different conformation such as Z-DNA, promoter strength can be affected (76). Collis *et al*. (77) have shown that loss of curvature in spacer DNA in the T7 promoter is associated with a nearly complete loss of promoter activity. Replacing a random DNA spacer sequence of synthetic *E. coli* promoters containing consensus sequences at both the -10 and -35 region with a properly positioned curved A:T tract enhances transcription initiation, probably by lowering transcription activation energy (78).

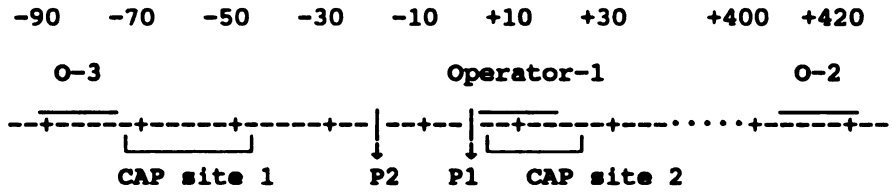
The role (if any) of sequences upstream and downstream of the consensus regions in promoter recognition by RNA polymerase is less well studied. It was reported that the DNA upstream of the -35 sequence can be deleted and replaced by heterologous DNA without affecting the strength of the

*lac*UV5 promoter *in vivo* or *in vitro* (79, 80). Upstream sequences clearly can affect promoter strength and function when regulatory proteins are involved.

However, there are somewhat ambiguous reports in the literature regarding downstream sequences and promoter strength. By systematically changing the sequence between +1 and +20 in promoters from bacteriophages T5, T7, and λ, Bujard and coworkers observed a more than 10-fold variation in promoter strength (81, 82). However, other studies show that the downstream sequences of the *lac* promoter are not essential for proper and efficient promoter utilization, both *in vivo* and *in vitro* (83, 84). Replacing natural *lac* DNA downstream of -6 with heterologous sequences using any of the four nucleotides at +1 has no effect on initiation efficiency or on the transcription start point. Thus, while RNA polymerase generally prefers to encounter a purine at +1, this is not a critical matter for the *lac* promoter. Studies of different promoters show that there is little correlation between the efficiency of promoter recognition and promoter strength *in vivo* (85), indicating that promoter strength in the cell is the result of an optimization process involving several parameters.

The lactose operon displays two overlapping and mutually exclusive promoters *in vitro*: the P1 promoter, from which CAP-dependent transcription begins at +1, and the P2 promoter, which initiates efficiently at -22 in the absence

of CAP (5, 86). The lacP2 promoter does not appear to be very efficient in initiation of transcription in vivo (84,122). The following is a simple diagram of the lactose operon control region. It depicts the relative positions of three potential operator sites, two CAP recognition sites and the start sites of the P1 and P2 promoters:



DNA BENDING AND TRANSCRIPTION ACTIVATION

A variety of models have been put forth to explain the mechanism of CAP activation. Gilbert suggested that CAP activates transcription through protein-protein contacts with RNA polymerase (87). CAP mutations that retain DNA binding activity but fail to interact cooperatively with polymerase have been isolated (68, 88). Two surface-exposed loops in the CAP DNA-binding domain have been identified as important for the cooperative interaction and enhancement of RNA polymerase. Recent genetic analyses have suggested a direct role for the C terminus of the alpha subunit as a target for several activator proteins. Reconstituted RNA polymerase carrying deletions of the C terminus of α was

competent to transcribe from many promoters but not from the lac promoter (40, 67). The C terminus of α appears essential in cases where the activator binds upstream of the -35 region (known as class I promoters), but not in cases where the activator-binding site overlaps the -35 region (known as class II promoters), such as gal (67).

Many regulatory proteins bend DNA upon binding (14), and DNA bending in the promoter region has been found to contribute to the regulation of promoters in both prokaryotes and eukaryotes. Classification of 43 promoter sequences in terms of the likelihood that they have curved structures indicates that promoters that have high transcription rates in vivo tend to have high curvature scores in their upstream regions (89). That CAP induces a sharp bend in its DNA binding site has been demonstrated by analyses of the mobility of CAP-DNA complexes in polyacrylamide gels (90, 91). The protein-induced DNA bending angle was measured by gel mobility comparison with known bent DNA standards (92); it could also be directly measured when complexes were imaged by the scanning atomic force microscopy technique (93). The CAP-induced bending angle was measured to be about 100° (92). There are other reports that CAP-induced DNA bending varies from 90° to 180°, the disagreement being apparently due to reaction conditions, the precise sequence of the CAP-binding site, the size of the DNA fragment, and the means of measurement

(57, 94, 95). Whereas the size of CAP would predict that it would interact only with ~20 base pairs of linear DNA, bending the DNA allows for interactions with a 28-bp segment that Liu-Johnson *et al*. (94) demonstrated is required for maximum CAP binding affinity.

To address the question of whether bending is a biological feature required for activation, or is only a side effect of protein binding, the CAP binding site upstream from the *lac* promoter was replaced by appropriately phased DNA bending sequences. The bent sequence modulates the relative rate of open complex formation by ten-fold. Constructs that yield optimal rates bend the promoter in the same direction as would the CAP-DNA complex, were it present (96). Similar results were obtained with modified *gal* promoter constructs (97), suggesting that bending induced by the cAMP-CAP complex upon binding to its site may be biologically relevant to the mechanism of transcriptional activation.

Recently, it was reported that appropriately phased protein-induced bending could act as a transcriptional switch factor *in vivo* (10). The CAP site in the *fur* operon was replaced with the RepA binding site, which was placed alternatively in phase with either promoter A (P_A) or promoter B (P_B). "In phase" means that the position of the protein binding site relative to the promoter is the same as in a natural promoter, both in distance and in rotation

angle along the double helix axis. P_A and P_B are two overlapping promoters in the *fur* operon. RepA (a repressor protein) bends DNA when specifically bound to DNA. Assay of transcripts using S1 nuclease showed that only the promoter phased with the RepA binding site became active. In each case, the promoter on the opposite side of the DNA helix simultaneously decreased its transcriptional activity to negligible levels. It is generally believed that CAP activation of the *lac*P1 promoter may be in part due to its excluding polymerase from binding at P2 site. In the *fur* operon, the CAP (or RepA) sites (at -70 with respect to the transcription start site) are sufficiently far from the polymerase binding sites of either promoter that protein binding should not physically exclude enzyme binding at the alternative promoters. Therefore, switching on or off of the *fur* promoters is solely dependent on the phase of RepA-induced DNA bending; the bending somehow affects polymerase binding at the promoter when the phase is not correct.

In cases where the promoter regulatory elements are located far from the transcription start site, DNA bending may lead to loop formation that enables the distantly bound regulatory protein to interact with polymerase and other proximal regulatory elements. In the *E. coli* *ara*BAD promoter, two AraC protein binding sites, *ara*I and *ara*O_2, are 211 bp apart. AraC protein binds at both sites simultaneously, causing a DNA loop to form, which is

necessary for repression of the promoter (13). The phasing of the sites is very important for loop formation. Insertion or deletion of a non-integral number of helical turns diminishes loop formation (98). The suggestion that the loop is involved in repression was subsequently strengthened by observations involving the *lac* repressor, where looping was observed directly using electron microscopy (99).

Activation of transcription can also involve loop formation. The binding site for the enhancer protein NtrC is approximately 100 bp upstream of the polymerase binding site in the *glnA* promoter, and also functions by looping back to allow NtrC to enhance RNA polymerase melting-in at the promoter at the expense of an ATP molecule (12).

It is known that RNA polymerase can stabilize a CAP molecule bound to the *lac* promoter. However, when one-half turn of DNA is inserted between the CAP and polymerase sites, the stabilization is lost. The CAP-polymerase interaction is partially restored with an 11 bp insert. Interestingly, the stimulation of transcription parallels the stabilization effect (100), so that CAP enhances both polymerase binding and the subsequent initiation process. In *gal*, the spacing between the CAP binding site and the -10 region is crucial for activation at the *galP1* promoter. A one base pair deletion (~34° phase shift) in the spacing region diminishes CAP enhancement of transcription from

galP1 (101). When an extra consensus CAP site is placed upstream of the natural CAP site at the lac promoter (at around -90), the activation of transcription is more than the cooperative DNA binding of the two CAP molecules can account for (102), suggesting synergistic activation by CAP, probably by contacting RNA polymerase at two distinct surface areas at the same time.

Although much has been learned about the mechanism of CAP activation of transcription, including possible roles of protein-protein interactions and of DNA bending and looping, many questions remain unanswered. Does CAP-induced DNA bending lead to loop formation in the lactose operon? How does CAP interact with polymerase and stimulate transcription at lacP1? What is the function, if any, of the second CAP binding site at +15, that overlaps the lacO1 sequence at around +11 (18)? In this thesis, we provide evidence that a regulatory loop likely plays a role in CAP stimulation of lacP1 transcription initiation, as the weak CAP binding site at +15, and another CAP site far upstream at -250 are apparently involved in loop formation.

CHAPTER II

TWO CAP MOLECULES INTERACT WITH THE *E. COLI Lac* OPERON AT THREE DIFFERENT SITES AND FACILITATE LOOP FORMATION

CHAPTER II

INTRODUCTION

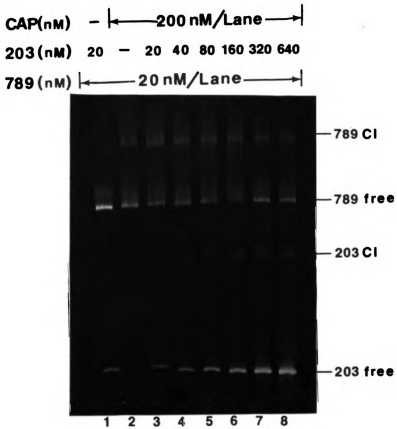
Catabolite activator protein interactions with cAMP, with the *E. coli lac* promoter and with RNA polymerase constitute an important paradigm for understanding the positive control of gene expression, because of the extensive genetic, biochemical, and structural studies which have provided significant insights into its mode of action. CAP acts at many "catabolite-sensitive" operons. While it was once expected that the mechanism of action of CAP would be the same at each of these, there is no reason to expect that every CAP-sensitive operon should respond in the same way to a change in cellular cAMP concentration. Indeed, recent studies indicate that a variety of molecular events are involved. It is of interest to compare and contrast CAP action at different promoter regions.

While much work has been done with the primary CAP binding site on the *lac* promoter centered at -61 relative to the transcription start site (+1), a second weak CAP binding site at +15 is less well studied. The CAP binding sites at different promoters involve different sequences and locations relative to the transcription start site. At the

galactose promoter, the primary CAP binding site is at -41, overlapping with the RNA polymerase binding region (60). Furthermore, when RNA polymerase is added to a solution of CAP-*gal* DNA complexes, a second CAP molecule binds immediately upstream of the first CAP to form a very stable 2(CAP):1(polymerase):1(promoter) structure that is active in transcription initiation (63). As indicated in chapter I, DNA bending and looping also appear to play a role in transcription complex function. Taken together, these observations lead us to ask whether CAP might interact with both of its potential binding sites at *lac*, whether more than one CAP molecule may be involved in CAP activation of transcription at the *lac* promoter, and whether loop formation may occur.

Several years ago we observed that about four times as much CAP is needed for titration of a relatively long (789 bp) DNA fragment of *lac* DNA (*lac* **789**) as is needed for an equivalent amount of the *lac* **203** fragment (203 bp). This result was from a gel retardation assay which measured the minimum amount of CAP which, when added to the DNA, resulted in disappearance of the band corresponding to the free DNA. Furthermore, when CAP is added to a mixture of **789** and **203** *lac* fragments, it binds preferentially to the longer fragment; even a 32-fold excess of the shorter fragment does not appreciably reduce the amount of CAP·*lac* **789** DNA complex (Figure 1). A possible explanation for these data would be

Figure 1. Gel binding assay of lac 789 and 203 DNA fragments competing for CAP binding. Lane 1, 20 nM 789 and 203 fragments each; Lanes 2 to 8, 20 nM 789 fragment per lane with increasing amounts of 203 DNA fragment as indicated and 200 nM CAP protein. CI denotes CAP-DNA complexes.


CAP(nM) – ←200 nM/Lane→
203(nM) 20 – 20 40 80 160 320 640
789(nM) ←20 nM/Lane→
789 CI
789 free
203 CI
203 free
1 2 3 4 5 6 7 8

that the additional sequences present on the longer fragment allow for cooperative binding of several CAP molecules to lac 789 DNA. Since most in vitro studies on the lac promoter have been done with the lac 203 DNA fragment, we undertook a series of experiments of the interaction of CAP with lac promoter fragments of different lengths.

MATERIALS AND METHODS

Proteins. RNA polymerase holoenzyme was isolated from E. coli strain CAG 1610 by the method of Burgess and Jendrisak (103) as modified by Peterson and Reznikoff (104). CAP was isolated from E. coli strain pHA7 by a procedure similar to that used for polymerase purification, involving Polymin P and $(NH_4)_2SO_4$ precipitation, combined with chromatography on Bio-Rex 70, Sephacryl S-200, and ssDNA-cellulose columns. Both RNA polymerase and CAP were kept in storage buffer (10 mM Tris-HCl, pH 7.9, 0.1 mM EDTA, 0.1 mM DTT, 50% glycerol, plus 0.1 M NaCl for RNA polymerase, or 0.2 M NaCl for CAP) at -20°C. The preparations were about 10% to 25% active based on assays of specific binding and stimulation of P1 transcription (105, 106). Proteins were at least 95% pure judged by SDS-polyacrylamide gel electrophoresis. Protein concentrations were determined spectrophotometrically using the extinction coefficients $\epsilon_{280}=3\times10^5$ $M^{-1}cm^{-1}$ for polymerase and $\epsilon_{280}=3.5\times10^4$ $M^{-1}cm^{-1}$ for CAP (105).

Lac Promoter Fragments. Figure 2 is a diagram of the lac DNA fragments used in these studies. The 789 bp DNA fragment, designated lac 789, extends from base pair -310 to +479 (numbered relative to the lac P1 promoter transcriptional start site at +1), has an EcoRI linker at the upstream end and a BamHI linker at the downstream end, and is cloned into pUC18 plasmid. The 652 bp DNA fragment (lac 652) extends from base pair -173 to +479. This fragment was cloned into pUC18 at the SmaI site after the overhanging ends were filled in using dNTPs and Klenow fragment. The lac 652 fragment was then isolated by digesting the plasmid with EcoRI and BamHI, which cleave at sites located on either side of the SmaI site, nearby in the polylinker region. The lac 367 (-310 to +57) sequence was obtained by partial digestion of the 789 fragment with HinfI, purify the 367 bp fragment by gel electrophoresis, and cloning into pUC18 at the SmaI site after the overhanging ends were filled in. The lac 518 (-310 to +208) fragment was generated by digesting the lac 789 fragment with BglII restriction enzyme. The lac 203 (-140 to +63) fragment has an EcoRI linker on each end and is cloned into pUC18. lac^H 487 DNA is a hybrid fragment made as follows: the plasmid containing lac 203 was cleaved with PvuII, yielding a 360 bp construct containing pUC DNA at the downstream end and the lac DNA sequence from -123 to +63 at the upstream end. This fragment was cloned into the pUC 19 at

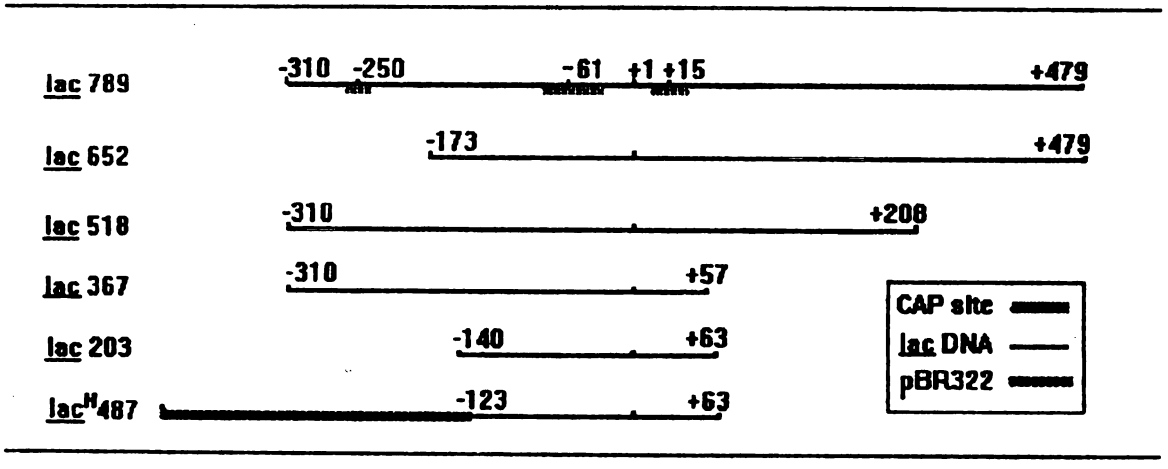


Figure 2. lac DNA fragments used in this work. The lengths are not strictly to scale.

the HindII site. The recombinant pUC19 was digested with EcoRI and HindIII, generating a 200 bp lac fragment (-123 to +63 plus pUC sequence from HindII to HindIII) that was saved for further cloning. Simultaneously, a 280 bp DNA fragment was isolated from pBR322, extending from pBR bp 376 to 656. The pBR DNA fragment ends were filled in with dNTPs and Klenow fragment, then a BamHI linker was added to one end and a HindIII linker to the other end, making it 287 bp long and ready for further cloning. The 287 bp pBR322 fragment was ligated to the upstream end of the 200 bp lac fragment previously prepared, and was cloned in pUC18.

Restriction fragments were purified by separation in polyacrylamide gels followed by electroelution. DNA fragments used in binding gels, DNase I footprinting, and exonuclease III protection experiments were 5′-end labelled with [γ-^{32}P] ATP and T4 polynucleotide kinase following treatment with calf intestinal phosphatase (107). Alternatively, they were 3′-end labelled with α-^{32}P-dATP, using Klenow fragment to fill in the recessed ends (107). Concentrations of DNA were determined spectrophotometrically using ϵ_{260}=13,000 $M^{-1}cm^{-1}$ (per mole of base pairs) and were expressed as molar in promoter fragments (105).

Binding Reactions. Binding reactions were performed in a binding buffer containing 20 mM Tris-HCl (pH 8), 10 mM $MgCl_2$, 0.1 mM EDTA, 0.1 mM DTT, and 100 mM NaCl or KCl, unless otherwise indicated. Reactions generally contained

10 to 50 nM DNA fragments, 10 to 1000 nM CAP, 20 μM cAMP and, where applicable, 100 to 750 nM RNA polymerase. Protein concentrations reflect numbers of molecules, some of which may not be fully active. The level of protein in a reaction is often presented as a ratio of moles protein per mole of DNA fragment present. CAP was incubated with DNA fragments for 15 min at 37°C; if polymerase was used, it was added after CAP and incubated for an additional 30 min at 37°C.

Binding Assays by Gel Retardation. Protein-DNA complexes were formed as described above using either unlabeled or radioactively-labelled lac DNA fragments. If RNA polymerase was a reactant, then after incubation for the appropriate length of time heparin (100 μg/ml) or poly d(A-T) (50 μM in bp) was added to eliminate nonspecifically bound polymerase (83, 100). Then 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 15 cm vertical 4% polyacrylamide (30:1 acrylamide:bis-acrylamide), 0.2% agarose mixed gel in TBE buffer (90 mM Tris, 90 mM H_3BO_3, 2.5 mM EDTA). 20 μM cAMP was added to the TBE running buffer and the gel was pre-run for about 5 min at 20 mA before samples were loaded. Electrophoresis was performed at 20 mA until the bromphenol blue was at the bottom of the gel (the xylene cyanol was about halfway down the gel at this point). A cooling fan was used to prevent

the gel from overheating during electrophoresis. Gels were either stained with ethidium bromide or were dried and autoradiographed.

DNase I Footprinting. Proteins were mixed with 5′- or 3′- end labelled fragments under binding conditions as described above. For footprinting experiments in solution, following the incubation period 1 μl of DNase I solution (1 μg/ml DNase I in 20 mM Tris-HCl, pH 8, 20 mM $MgCl_2$, 20 mM $CaCl_2$) was added. After incubation for 30 sec at 37°C, the reaction was halted by addition of an equal volume of stop buffer (0.3 M NaOAC, 0.02 M EDTA, 100 μg/ml tRNA), followed by phenol-chloroform extraction and ethanol precipitation to concentrate the DNA. The digested DNA was dissolved in 10 to 15 μl denaturing loading buffer (90% formamide in TBE, 0.1% bromphenol blue, 0.1% xylene cyanol) so that the sample contained approximately 3000 cpm per μl. Samples were then heated for 3 min at 97°C and loaded onto a 5% polyacrylamide sequencing gel (20:1 acrylamide:bis-acrylamide, with 7 M urea in TBE buffer).

Improved resolution in footprinting experiments can be obtained if the complexes are gel purified before DNase I treatment, as described by Crothers et al. (100), with slight modifications. In this case, the DNA-protein solution is first electrophoresed on a non-denaturing gel. The gel is then exposed to X-ray film for 30 min at room temperature, so that the location of the desired band of

complexes can be identified. The gel slice containing the desired complexes is cut out and 4 μl DNase I solution (1 μg/ml DNase I in 10 mM Tris, pH 8, 2 mM DTT, 5% glycerol, 0.5 mg/ml BSA) is spread on it. After 5 min at room temperature, 4 μl of salt solution (50 mM $MgCl_2$, 50 mM NaCl) is spread on gel and incubation is continued for another 1.5 min at 37°C, after which the reaction is quenched with 15 μl of 100 mM EDTA. The DNA, which has now undergone a limited digestion by the nuclease, is eluted from the gel, ethanol precipitated, washed with 70% ethanol and dried. The DNA sample is then dissolved in denaturing loading buffer without further purification, and electrophoresed on a sequencing gel.

Direct Measurement of CAP Stoichiometry in Complexes. Binding reactions and gel retardation experiments were performed as described above. The highly radioactive DNA fragments allowed desired complex bands to be visualized by a 30 min exposure of the X-ray film. Bands of complex were cut out and inserted into the well of a slightly thicker non-denaturing gel and electrophoresed to further purify complexes from any contaminating proteins. The second binding gel was dried and autoradiographed, and the desired complex bands were excised.

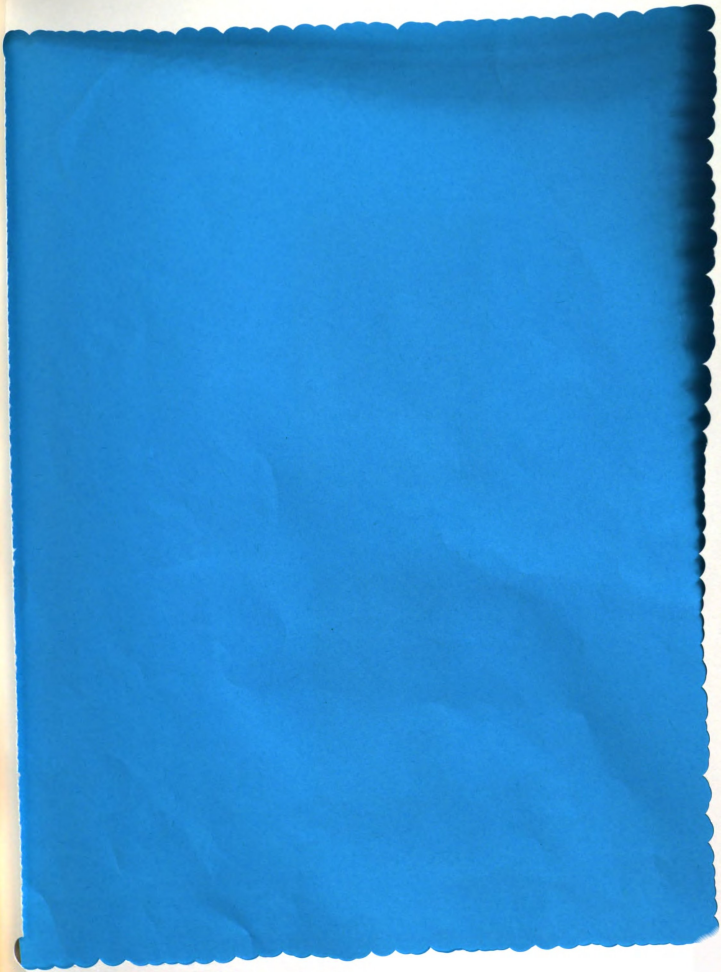
The amount of DNA in each complex band was determined by comparison with standards, as follows. A known quantity of labeled DNA was run in an additional lane in the second

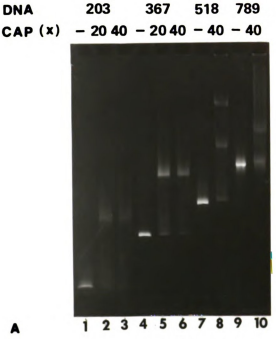
binding gel. A dried gel slice containing the control DNA, and dried gel slices containing the complex bands of interest were placed in individual scintillation vials and counted with an open window (no scintillation fluid added). To determine the level of protein, the gel slice was then soaked with denaturing buffer (5% β-mercaptoethanol, 1% SDS) (28) for 15 min at room temperature, keeping the volume as small as possible, so there was no extra solution left before insertion into the well of a Laemmli SDS-polyacrylamide gel (4% stacking/10% separating gel) (108). The gel was electrophoresed at 200 v until the bromphenol blue marker dye reached the bottom, then was silver stained (109). The amount of CAP in each complex band was determined by densitometry of the protein bands compared with bands run in control lanes containing known amounts of CAP.

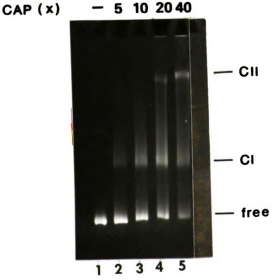
RESULTS

CAP Binding Affinity as a Function of Length of Lac Promoter Fragments. Figure 3 shows the results of titration of lac 203, 367, 518, and 789 DNA fragments with CAP. The data show that the amount of CAP needed to shift all free DNA to complex (i.e., to eliminate the "free DNA band") is larger for longer fragments. Thus 20X CAP is enough to shift all the 203 DNA to the complex band (Figure 3, Panel A, lane 2), but not for longer fragments. 40X CAP shifts

Figure 3. Gel binding assay of CAP complexes formed with different lac promoter fragments. Panel A, lanes 1, 4, 7, and 9, lac 203, 367, 518, and 789 DNA fragments alone; lanes 2 and 3, lac 203 plus CAP; lanes 5 and 6, lac 367 plus CAP; lane 8, lac 518 plus CAP; lane 10, lac 789 plus CAP. "(X)" indicates the ratio of CAP molecules to DNA fragments. Panel B, lane 1, lac 789 DNA fragment alone; lanes 2 through 5, lac 789 fragment plus indicated amount of CAP. CI denotes faster migrating CAP-DNA complexes, CII denotes slower migrating CAP-DNA complexes.




DNA 203 367 518 789
CAP (x) – 20 40 – 20 40 – 40 – 40
A
1 2 3 4 5 6 7 8 9 10


CAP (x) – 5 10 20 40
CII
CI
free
B
1 2 3 4 5

relatively less 789 DNA than it does 518 and 367 DNA fragments from the free DNA band.

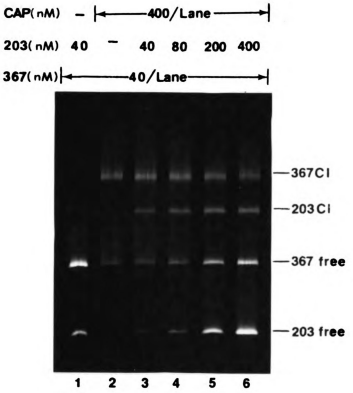
These results immediately raise the possibility that there may be cooperative binding of more than one CAP molecule at the longer fragments. As noted above, the lac promoter does have at least two CAP binding sites, although one of them is weak. However, since both sites are present in all fragments used here, it is necessary to postulate some special effects from the additional DNA sequences in the longer DNA molecules. The situation is further complicated by the discovery of a second, slower migrating, CAP complex, designated as CII, which forms only in the presence of excess CAP on lac 518 and 789 (Figure 3, Panel A, lanes 8, 10 and Panel B, lanes 4, 5). Thus it appears that the longer fragments are capable of binding multiple CAP molecules, in more than one kind of complex.

It is unlikely that the apparent higher affinity of CAP for longer fragments is related to the fact that there are more nonspecific binding sites on the longer DNAs. Theoretically, each base pair can be considered as the beginning of a nonspecific binding site, hence a 789 bp fragment would have about 4 times as many potential sites as a 203 bp DNA, though each fragment would have a single specific binding region. However, at the ionic conditions in these reactions, little nonspecific CAP binding is expected (123). Furthermore, the stoichiometry data along

with other control experiments involving a construct with a non-lac DNA sequence (see below) imply that the differences in CAP affinity for long and short lac DNA fragments resides in interactions that are unique to the lac sequence.

The relative affinities for CAP binding to different length lac DNA fragments were assessed by gel retardation assays of solutions containing pairs of fragments. Different quantities of any two lac DNA fragments, such as 203 plus 367 or 789, or 367 plus 789, were mixed and incubated with limiting quantities of CAP, followed by gel electrophoresis. Figure 4 provides a sampling of results from these competition experiments (also see Figure 1). We find that the longer fragment always competes better for CAP binding than the shorter one does. In this instance, the 367-CAP complex remains at an almost constant level even at 10-fold excess of 203 DNA (compare lanes 2 and 6). In the reaction run in lane 6, the input CAP:203 ratio is 1:1. Thus, if the 203 fragments were as capable in binding CAP as the 367 fragment, and even if all CAP molecules were fully active, 90% of the CAP molecules should be bound to the 203 DNA, and little 367-CAP complex would be detected. That this is not found indicates that the longer 367 fragment binds CAP better than does the shorter 203 DNA. The same pattern is seen for other lac DNA fragments as well (data not shown). The longer fragments consistently compete better for the protein.

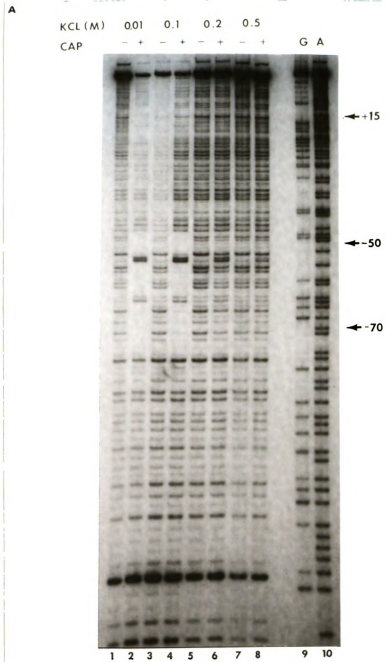
Figure 4. Gel binding assay of lac 367 and 203 fragments competing for CAP binding. Lane 1, 40 nM of 203 and 367 DNA fragments each; lanes 2 to 6, 40 nM 367 DNA per lane, with increasing amount of 203 DNA as indicated, plus 400 nM CAP.


CAP(nM) – ←400/Lane→
203(nM) 40 – 40 80 200 400
367(nM) ←40/Lane→
367CI
203CI
367 free
203 free
1 2 3 4 5 6

The gel shift experiments imply that CAP binding affinity is a function of the length of the lac DNA fragments used, and that multiple copies of CAP may bind at longer lac DNA fragments. The complexes were therefore characterized further by DNase I footprinting and by stoichiometry measurements. The lac CII complex is particularly interesting. It forms only when the CAP concentration is high and the lac DNA fragment is sufficiently long. Thus, it is possible that the longer lac fragments may contain additional unidentified CAP binding sites that are missing on the shorter fragments, which may affect CAP·lac interaction and CII formation.

DNase I Footprinting of Various CAP Complexes. Protection by bound proteins of DNA from digestion by nucleases is a very sensitive technique for detecting sites for proteins on a given DNA fragment. To ascertain whether more than one CAP molecule is bound to the longer lac DNA fragments (and where the binding sites are), DNase I protection experiments were performed with each lac fragment after incubation with CAP. The results are as follows. At 100 mM salt (NaCl or KCl) concentration, only one region protected by CAP is seen for the 203 DNA fragment, which is the usual -61 (-50 to -70) site (Figure 5). Although DNA-protein interactions are generally weakened at elevated salt concentrations, CAP binding at the -61 site of lac P1

Figure 5. DNase I footprinting assay of salt effect on CAP binding to the lac 203 DNA fragment. Panel A, lac 203 is labelled at the 5′-end of the upper strand. Panel B, lac 203 is labelled at the 5′-end of the lower strand. "+" and "-" symbols denote whether CAP is present or absent. Salt concentrations are as indicated. Lanes 9 and 10, panel A, show the "G" and "G+A" Maxam-Gilbert sequencing reactions for the lac 203 fragment upper strand, which allow identification of the specific base to which each band corresponds.


A
KCL (M) 0.01 0.1 0.2 0.5
CAP − + − + − + − +
G A
+15
-50
-70
1 2 3 4 5 6 7 8 9 10
B
− + − + − + − + CAP
0.02 0.1 0.2 0.5 NaCl (M)
-77
-44
-16
+23
1 2 3 4 5 6 7 8

promoter is very stable, as even 500 mM KCl does not completely eliminate the CAP-DNA interaction at this site (Figure 5, Panel A, lane 8). There are clearly two CAP protection regions on the lac 367, 518, and 789 DNA fragments, one at -61 and the other located around +15 (Figures 6, 7, and 8). The +15 site on the lac 203 fragment can be protected by CAP but only at a much lower salt concentration (10 to 20 mM KCl) (Figure 5, Panel B, lane 2). These experiments were performed at CAP:DNA ratios (20-50X) at the higher of which the longer lac DNA fragments 518 and 789 may form CII complexes when reactions are analyzed by gel retardation assays. However, at 100 mM salt, their DNase I footprinting patterns are unaffected by the amount of CAP present.

The lac 789 fragment has the same 5'-end as the 367 fragment, thus extends the 367 DNA more than 400 bp downstream. Figure 3 (and other data not shown) reveals that while the CII band is seen for longer fragments, it does not appear when the 367 fragment is used. This leads to the question of whether there are additional CAP binding sites located in this downstream region of the 789 fragment. Careful inspection of footprinting data reveals no CAP protection in the downstream sequence of lac 789 DNA. In spite of this finding, we pursued the question of whether the length of the DNA fragment might be crucial for CII complex formation, even though no specific new CAP binding

Figure 6. DNase I footprinting assay of CAP-lac 367 complexes. Lane 1, 367 DNA alone; lane 2, 367 DNA plus CAP (20X); lane 3, 367 DNA plus RNA polymerase; lane 4, 367 DNA plus CAP, then polymerase. Heparin was added to reactions containing RNA polymerase before samples were exposed to DNase I.

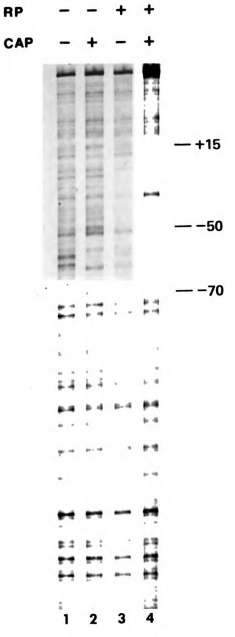
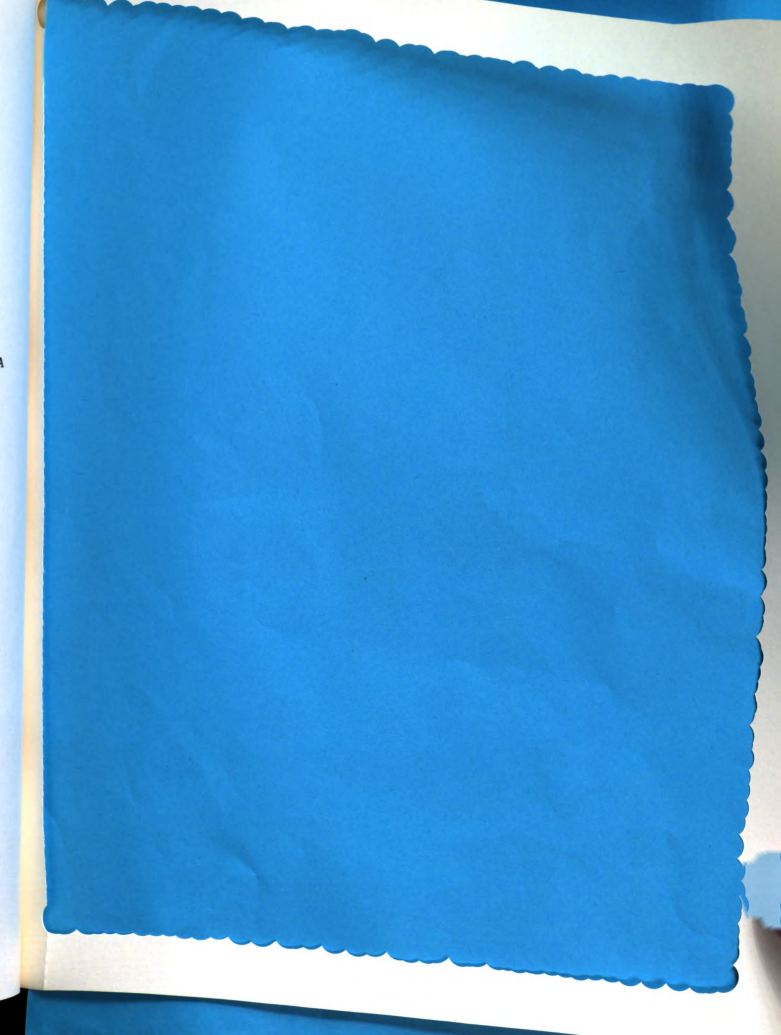
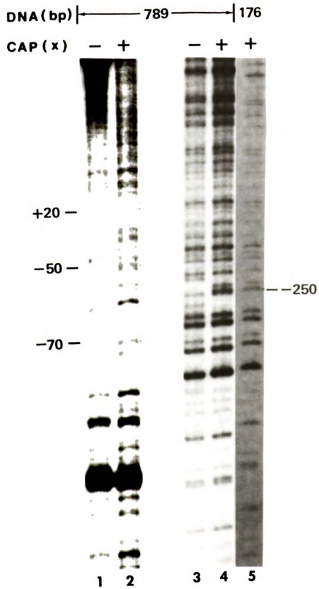

RP - - + +
CAP - + - +
+15
−50
−70
1 2 3 4

Figure 7. DNase I footprinting assay of CAP-*lac* 518 complexes. Lane 1, 518 DNA alone; lane 2, 518 DNA plus CAP (40X); lane 3, 518 DNA plus RNA polymerase; lane 4, 518 DNA plus CAP, then polymerase. Heparin was added to reactions containing RNA polymerase before samples were exposed to DNase I.


Heparin – – + +
RP – – + +
CAP – + – +
+20
–70
–250
1 2 3 4

Figure 8. DNase I footprinting of CAP protection at lac DNA fragments. "-" and "+" symbols denote absence or presence of CAP. DNA fragments are labelled at the 5'-end of the upper strand. Lanes 1 and 3, DNase I digest of free 789 fragments; these are derived from one reaction, but because of the large size of the fragment, are electrophoresed for different lengths of time; lanes 2 and 4, same as lanes 1 and 3 but with CAP (50X) present during DNase I digestion; lane 5, DNase I digestion in the presence of CAP of a 176 bp lac fragment (extending from base pair -310 to -134), which was isolated by digesting the labelled 789 fragment with HinfI, followed by gel purification.




DNA (bp)
789
176
CAP (x)
–
+
–
+
+
+20
–50
–70
–250
1
2
3
4
5

sites appear in the footprints. For these experiments, we constructed a 631 bp fragment in which 264 bp of lacZ DNA were added to the downstream end of the 367 fragment. Gel retardation assays showed that the longer hybrid fragment (631 bp) is capable of forming the CII complex (Figure 9, lane 8). The hybrid fragment (LP 631) was digested with DdeI at -242 to produce a 560 bp fragment (LP 560); this fragment is also capable of CII complex formation with CAP (Figure 9, lane 6). As controls, purified random pBR322 DNA fragments longer then 518 bp were also tested for their ability to bind CAP. Figure 10 shows the results of CAP interaction with random non-lac DNA sequences. It is clear from this figure that random DNA molecules longer than 500 bp do not form CII-like complexes when incubated with CAP (nor do they form CI-type complexes with CAP either); thus CII complex formation does not solely depend on the length of DNA fragment, but also requires the presence of specific lac promoter sequences. The CII complex forms only when the DNA fragment is long enough (roughly 500 bp or longer) and contains lac promoter sequences as well. An exception to this conclusion is seen in Figure 10, lane 5, where the 527 bp pBR DNA fragment is seen to form both CI and CII-like complexes. This may be due to a CAP binding site in pBR322 at the P4 promoter, upstream of the RNA polymerase binding region, though this CAP site is not very strong compared to the primary CAP site at the lactose promoter (124).

Figure 9. Gel binding assay of the effect of lac DNA fragment length on the formation of CAP complexes. L, wild type lac DNA fragments. LP 560, hybrid construct containing lac DNA sequences -242 to +54 plus 264 bp lacZ DNA from pUC18. LP 631, construct containing lac DNA sequences -310 to +54 plus 264 bp lacZ DNA. CI, CAP complex I, faster migrating complex band; CII, CAP complex II, slower migrating complex band.

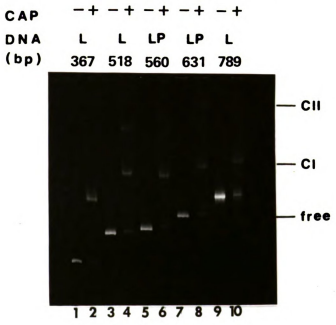
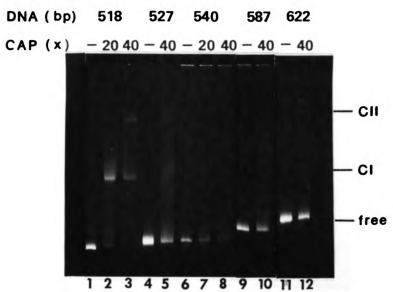

CAP
− + − + − + − + − +
DNA
L L LP LP L
(bp)
367 518 560 631 789
CII
CI
free
1 2 3 4 5 6 7 8 9 10

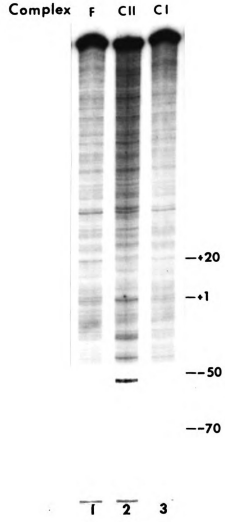
Figure 10. Gel binding assay of length effect on CII formation with or without CAP primary -61 site. Lanes 1 to 3, lac 518; lanes 4 through 12, purified pBR322 DNA fragments (most of which containing no known CAP binding sequences).


DNA (bp) 518 527 540 587 622
CAP (x) – 20 40 – 40 – 20 40 – 40 – 40
CII
CI
free
1 2 3 4 5 6 7 8 9 10 11 12

To probe further the CI and CII complexes, 789 CI and CII were purified by electrophoresis through a nondenaturing gel, then were footprinted in the gel slice using DNase I. The results are shown in Figure 11. The 789 CI complex CAP protection pattern is the same as found in footprinting of a solution of 789 DNA plus CAP; DNA is protected at both the -61 and the +15 sites (lane 3). However, the 789 CII complex shows a single protection region extending from -70 to +20 (lane 2). This surprising result implies that at high CAP concentration, CAP interacts with lac DNA in a cooperative fashion, provided that the DNA is sufficiently long. Assuming that each CAP can protect 20 nucleotides, there should be at least four CAP molecules bound in the 789 CII complex. These CAP molecules are bound through the region in which RNA polymerase binds. Can the CII-type protein-DNA interaction occur under physiological conditions, does it have an impact on RNA polymerase open complex formation at the lac promoter, or is it an artifact of the in vitro experimental conditions? The following experiments are aimed at further elucidating DNA-protein interactions in the 789 CII complex.

CAP-lac DNA Interaction as a Function of the Concentrations of Salt and cAMP. Specific binding of CAP to promoter DNA sites requires that the protein be in complex with cAMP. Binding of the cyclic nucleotide causes an

Figure 11. DNase I footprinting assay of gel-purified lac 789 DNA complexes with CAP. CAP was incubated with 789 DNA before loading samples on the binding gel. Two CAP complexes with the lac 789 fragment were resolved (cf. Figure 3, panel B). Lane 1, free 789 DNA control; lane 2, the slower migrating 789-CAP complex, CII; lane 3, the faster migrating 789-CAP complex, CI.


Complex
F
CII
CI
+20
+1
-50
-70
1
2
3

allosteric conformational change that activates CAP for specific interaction with its binding site. The ionic strength of the system is a second critical factor influencing DNA-protein interactions. The specificity and binding constant of a protein for any DNA sequence usually is a strong function of the ionic strength and composition, especially the cation concentration. This reflects the fact that positive charges on the protein are involved in a competition with buffer cations for sites on the highly negative DNA anion. Electrostatic attraction generally is a larger component of nonspecific DNA-protein binding affinities; specific interactions involve other forces as well. Thus nonspecific binding that can be detected at low salt may not be evident at higher salt. In gel shift and footprinting assays, only specific CAP-*lac* promoter DNA complexes are seen at 100 mM salt. When the salt concentration is lower (10 to 20 mM), nonspecific binding may become apparent.

The effects of cAMP and salt concentration on *lac* CII and CI complex formation were investigated. The cAMP dependence of CAP complex formation was tested with the 789 DNA fragment. Figure 12 shows the results of the mobility shift assay when cAMP is not present in either the gel itself or the running buffer. With no cAMP in the gel system, CII complexes are not observed at either low or high salt concentration, even with 50 μM cAMP present in the

Figure 12. Gel binding assay of cAMP and salt effects on lac 789 CII complex formation. There is no cAMP present in the gel or running buffer; cAMP concentrations shown are for the binding reactions. "+" and "-" symbols denote whether a component is present or absent. Panel A, 0.1 M KCl in each binding reaction; left two lanes, no cAMP in binding reaction; lanes 3 to 7, increasing cAMP concentration. Panel B, 20 μM cAMP in each reaction; lanes 1 and 2, no KCl in binding reactions; lanes 3 to 8, increasing KCl concentration.

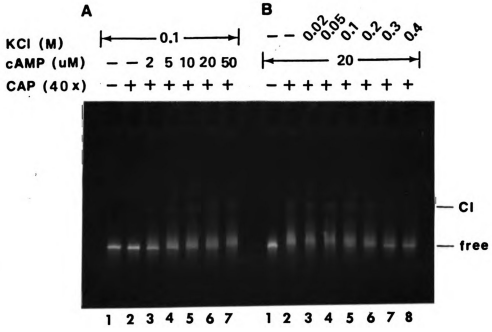
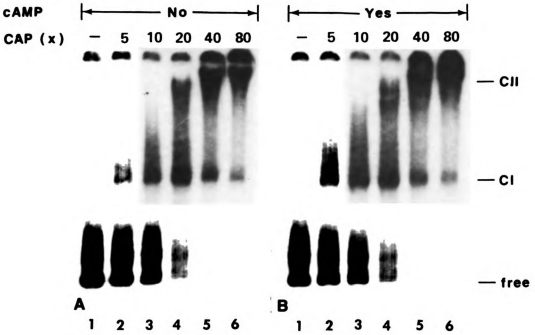

A
B
KCl (M)
0.1
0.02
0.05
0.1
0.2
0.3
0.4
cAMP (uM)
2 5 10 20 50
20
CAP (40x)
CI
free
1 2 3 4 5 6 7
1 2 3 4 5 6 7 8

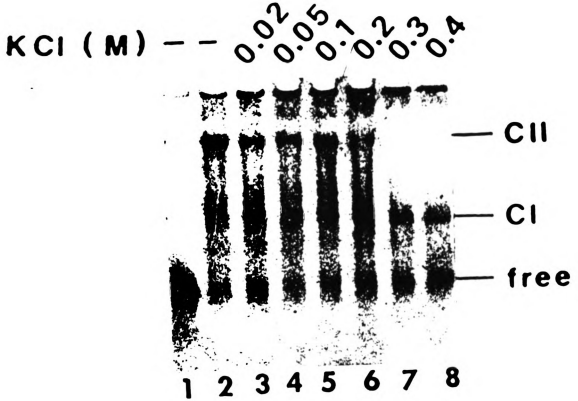
Figure 13. Gel binding assay of CAP-lac 789 DNA CII complex formation as a function of cAMP concentration. There is 20 μM cAMP present in the gel and gel running buffer. Panel A, no cAMP present in binding reaction. Panel B, 20 μM cAMP present in binding reaction. The CAP/DNA ratio is as indicated.


cAMP
No
Yes
CAP (x)
–
5
10
20
40
80
CII
CI
free
A
B
1
2
3
4
5
6

reaction. However, as little as 2 μM cAMP in the binding reaction is enough for stable 789 CI complexes to be seen in the gel (Figure 12, Panel A, lane 3). In the absence of cAMP, of course, no CAP complex is detected (Figure 12, Panel A, lane 2). In contrast, Figure 13 shows that 789 CII complexes can form at 20X CAP as long as there is cAMP in the gel system (at 20 μM cAMP) even when cAMP is absent in the binding reaction (Figure 13, Panel A), though the extent of CII complex formation is greater when cAMP is present in the binding reaction (Figure 13, Panel B). When cAMP is present in the gel system, the CII complex probably forms after the sample is loaded onto the gel (before it runs into the gel); when cAMP is absent from the gel and electrophoresis buffer, it appears that electrophoresis of cAMP away from the DNA-protein complexes markedly reduces the stability of the CII complexes.

Figure 14 shows the salt effects on CAP-789 DNA complex formation. It is of interest that, while the gel assay indicates that the 789 CI complex is stable at [KCl] as high as 0.4 M, the amount of CII complex diminishes at elevated salt concentrations. This result is relevant to the question of the relationship between CI and CII complexes. Does CII form when extra CAP molecules bind to a CI complex? If this were true, one might expect that CII would convert to CI at high salt concentration. Figure 14 shows, however, that as [KCl] increases, the number of CII complexes

Figure 14. Gel binding assay of lac 789 CI and CII complex formation as a function of reaction buffer salt concentration. Lane 1, lac 789 DNA alone; lanes 2 though 8, 789 DNA plus 40X CAP (i. e., moles CAP/moles DNA fragment = 40). KCl concentrations are as indicated.


KCl (M) – – 0.02 0.05 0.1 0.2 0.3 0.4
CII
CI
free
1 2 3 4 5 6 7 8

diminishes (and [free 789 DNA] increases), while the level of CI complexes is relatively unaffected. Thus the CI-CII relationship may not be a simple one. Finally, it should be kept in mind that the TBE buffer generally used in gel shift assays is a low ionic strength buffer, hence may support CII complex formation.

Stoichiometry of CAP in Complexes at lac DNA Fragments of Different Length. Footprinting results of CAP-lac complexes presented above indicate that there may be multiple CAP molecules bound at longer lac DNA molecules (the 367, 518, and 789 fragments), rather than the single CAP that apparently binds to the 203 fragment. To determine precisely how many CAP molecules are bound to each lac fragment, complexes were twice purified by gel electrophoresis and DNA was quantified as described in Materials and Methods, then were analyzed by SDS-polyacrylamide gel electrophoresis to quantify CAP. Figures 27-29 (in Chapter III) are examples of typical silver stained SDS gels. CAP stoichiometries in various CI and CII complexes, with different length lac fragments, are summarized in Table I. There are two CAP molecules on each 789 DNA fragment in the CI complex band, two CAPs per promoter in the 367 CI-equivalent band, but only one CAP molecule per 203-CAP complex (CI-equivalent band). There are approximately 15 CAP molecules per DNA fragment in the

Table I

CAP Stoichiometry in Lac Promoter-Protein Complexes

CAP only complexes	CAP / promoter	open ("M_o") complexes	CAP / promoter
203-CAP	1	203 ternary	1
367-CAP	2	367 ternary	2
789-CAP (CI)	2	789 ternary	2
789-CAP (CII)	15		
652-CAP (CI)	1		

Notes: All complexes were formed in reaction buffers at 100 mM salt concentration.

The data for open "M_o" complexes are discussed in Chapter III.

789 CII complex. The lone CAP in complex with 203 DNA binds at the -61 site (see Chapter I), while in the 789 CI and 367 CI-equivalent complexes, the footprinting results show that one CAP binds at -61 and the other at +15. As for the 789 CII complex, footprinting reveals protection mainly in the -70 to +20 region, implying extensive protein-protein interaction among the 15 CAP molecules in addition to DNA contacts by a limited number of CAP molecules (there simply is not enough space within this 90 bp region for each of 15 CAP molecules to contact DNA).

DISCUSSION: The Possibility of CAP-induced Loop Formation at the lac Promoter.

The mechanism of action of regulatory proteins may involve their binding at two sequences that are linearly separated on DNA, and bringing those two sites together to form a DNA loop. The prototype for this behavior is the AraC protein interaction at the ara BAD operon (13). The AraC-induced loop structure represses transcription from this promoter. At the E. coli lac promoter, DNA looping apparently plays a role in repression by the lac repressor protein; three repressor binding sites have been identified (99). In my study of lac DNA fragment interaction with CAP protein, several CAP binding sites have been detected. For the 203 DNA molecule, there is the primary CAP binding site centered at -61, and a weak CAP binding site around +15

which is detectable in CAP-DNA solutions only at low salt concentration (Figure 5, lane 2). [This has also been reported by Schmitz (18)]. When the DNA fragment used includes additional lac sequences, such as the 367 DNA fragment which has the upstream sequence from -140 to -310, or the 789 DNA fragment which also has the downstream +63 to +479 sequence, the weak CAP binding site around +15 is strengthened and can be detected at a higher salt concentration (100 mM KCl) generally used in binding studies (lane 2 of Figures 6 and 8). Consistent with this observation from DNase I footprinting is a change in CAP stoichiometry in CI complexes from 1 (203 DNA) to 2 (367 and 789 DNA) (Table I). This is a quite unexpected result. How can the presence of additional DNA sequences alter the binding affinity of CAP for a site (at +15) that is not adjacent to the CAP site at -61? One possibility involves DNA looping, which would involve sequences in addition to those at -61 and +15.

To study this question further, the lac 789 DNA fragment interaction with CAP was studied by DNase I footprinting. Careful inspection of the data indicates that a third region of the DNA, around -250, is also affected when CAP is bound (Figure 8, lanes 2, 4). The -250 site protection is limited, but is seen independently of the presence of the other two sites. For example, using a 176 bp (-310 to -144) DNA fragment isolated following HinfI digestion of the 789

fragment, we see that CAP protects the -250 region from DNase attack just as in the 789 fragment (Figure 8, lane 5).

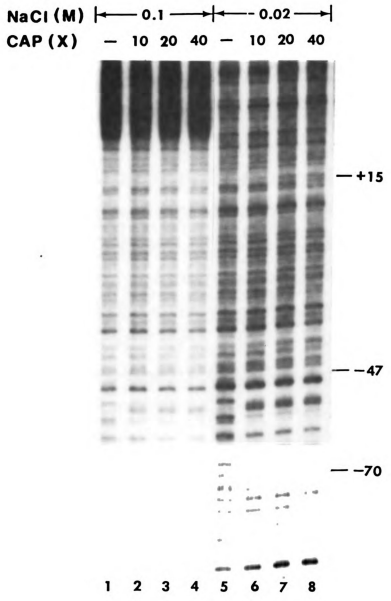
Comparing CAP interactions with the lac 203 and 367 fragments reveals that CAP protection at the +15 site is influenced by the sequences upstream of -140. The upstream sequences may affect CAP interaction via a DNA loop that involves CAP binding to a region in the upstream sequence (presumably at -250) and to the +15 region simultaneously. If CAP interaction with both the -250 and the +15 sites indeed facilitates the formation of a loop structure at the lac promoter, an important question is its possible role in stimulating transcription initiation by RNA polymerase at lac P1. This matter will be further considered in Chapter III.

It should be noted that the DNase I protection pattern at the -250 region of lac is a weak one, and is not always convincingly seen in every experiment. It is not clear why this is the case. Nevertheless, the upstream sequences present in the 367 fragment markedly affect CAP binding as compared with the 203 fragment, and it seems likely that the -250 region is the sequence upstream of the -140 that is responsible for enhancing the binding of the second CAP molecule to the +15 site.

Further evidence for the hypothesis that the second CAP molecule interacts with both the -250 and the +15 regions simultaneously at normal salt concentration (100 mM) to

facilitate the formation of a loop structure, was derived from studies with a lac DNA fragment (lac 652) from which the -250 region was removed. This fragment was made by digesting the lac 789 fragment with NarI (which cuts at -173), to generate a 652 bp lac fragment (-173 to +479). DNase I footprinting and stoichiometry experiments were used to study lac 652 interaction with CAP. If the -250 region is required for the second CAP molecule binding at the +15 site, the lac 652-CAP CI complex should behave in DNase I footprinting and stoichiometry assays just as the 203 fragment does; that is in each 652 CI complex, there would be one CAP molecule at -61 site, with the +15 site protected by CAP only at low salt, but not at 100 mM salt. Figure 15 shows the results of DNase I footprinting of the lac 652 fragment at both 100 mM and 20 mM NaCI. As predicted, the primary CAP site at -61 is protected at both high and low salt, but protection at the +15 region is only seen at low salt concentration, and when the CAP:DNA molar ratio is higher than 10 (see dotted site in Figure 15, lanes 7 and 8). At 100 mM salt, the +15 region is not protected even at a much higher CAP:DNA molar ratio (40x) (Figure 15, lane 4). Thus the 652 fragment is indeed a "203-type" fragment, even though it has additional lac sequences. The CAP content of the lac 652-CAP CI complex was determined as described above. The results show that there is a single CAP molecule in each lac 652 CI complex at 100 mM salt (Table I), while

Figure 15. DNase I footprinting assay of CAP interaction with *lac* 652 DNA as a function of salt concentration. Lanes 1 and 5, 652 DNA alone; lanes 2, 3, 4, 6, 7, and 8, 652 DNA plus indicated amount of CAP. Salt concentrations are as indicated.

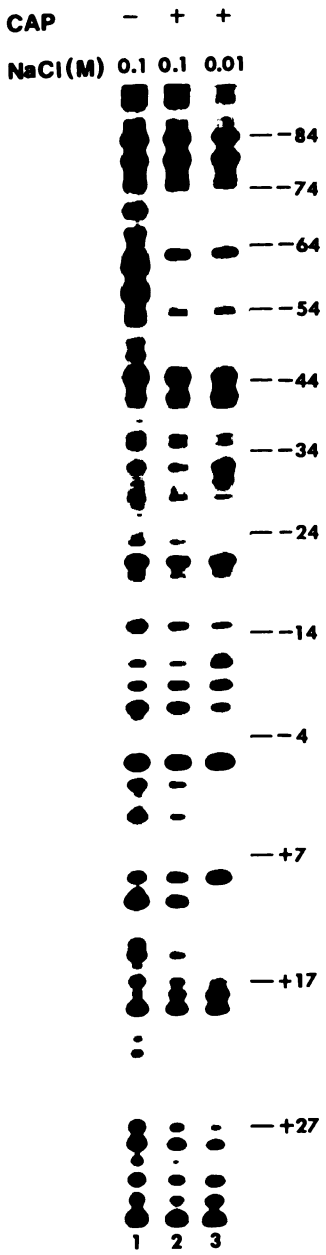

NaCl (M)
0.1
0.02
CAP (X)
— 10 20 40 — 10 20 40
+15
−47
−70
1 2 3 4 5 6 7 8

the CI complex of 652 DNA formed at low salt (20 mM) has about two CAP molecules for each DNA fragment (data not shown). Thus the stoichiometry data substantiate the footprinting results, and are as expected for a 203-type fragment that does not contain the -250 sequences. The -250 region appears to be indispensable for stable binding of a second CAP molecule at +15 at 100 mM salt, which most probably involves a loop structure.

To investigate if any DNA at -250 will do, or whether a specific -250 sequence is required, studies were done using a hybrid DNA construct, $lac^{H}487$. This fragment contains a 287 bp pBR322 DNA sequence attached to the upstream end of lac 203 DNA, so that it extends more than 300 bp upstream of +1. Figure 16 shows the DNase I footprinting of complexes of $lac^{H}487$ DNA with CAP. The data show that $lac^{H}487$ DNA is a 203-type lac DNA fragment. At 100 mM NaCl, the +15 site is not very well protected by CAP (lane 2); however, when the salt concentration is lowered to 10 mM, sequences around +15 (+2 to +22) are protected by CAP (lane 3). Thus, in order for a second CAP to stably bind at the +15 site of the lac promoter at 100 mM salt, the "far-upstream" DNA sequences (around -250) are not only necessary, but also should be specific lac sequences.

The CAP consensus sequence is composed of two symmetrical half sites, with 5′-TGTGA-3′ being the most important specificity determinant for CAP recognition in

Figure 16. DNase I footprinting assay of CAP interaction with lac^{H} 487 DNA as a function of salt concentration. Lane 1, lac^{H} 487 DNA alone; lanes 2 and 3, lac^{H} 487 DNA plus CAP (40X). Salt concentrations are as indicated.


CAP
NaCl(M)
0.1
0.1
0.01
-84
-74
-64
-54
-44
-34
-24
-14
-4
+7
+17
+27
1
2
3

each half site. Comparing sequences around the +15 and the -250 regions of the wild type lac promoter reveals two TGTGA motifs around +15, one on each strand; however, these are 13 bp apart instead of the usual 6 bp apart for a consensus CAP binding sequence. CAP protection from DNase I digestion around +15 covers only about 20 bp, which is 8 bp fewer than the 28 bp that CAP protects at the -61 site (94). This may indicate that CAP binding at +15 does not bend DNA as much as it does at the -61 site, assuming that CAP would protect fewer base pairs on a linear DNA fragment, more base pairs on a bent fragment.

At the -250 region, there exists a TGAGA sequence, on the lower strand from -242 to -246. This sequence is found in the CAP site of the cat promoter (53). There is no sequence resembling TGTGA on the upper strand around -250. Thus the proposed third CAP site at -250 is perhaps only a half site. This corresponds well with the footprinting data showing that CAP protection at -250 region covers at most about 6 nucleotides (Figure 8). Perhaps the -250 site only interacts with a monomer of the CAP dimer molecule.

Since the two TGTGA motifs in each half site of the lac +15 CAP site are separated by 13 base pairs, to bind to this site, the CAP molecule would need to distort its structure considerably, which would likely reduce the affinity of CAP for this site. Lowering the ionic strength would make the binding more favorable; indeed, CAP protection from DNase I

at +15 is seen at low salt. There appears to be additional stabilization obtained from interactions of CAP with -250 sequences as well via a DNA loop. Perhaps the function of the DNA looping is to strengthen binding of the second CAP at +15 under physiological salt conditions (~100 mM monovalent ion).

Is the CAP binding at +15 dependent on binding of the first CAP molecule around -61? Data in this chapter do not directly answer this question. In all cases, CAP binding to the -61 region is observed, but this could simply result from a strong independent interaction there. However, in Chapter III we see that in the presence of RNA polymerase, the affinity of CAP for the +15 is related to binding of CAP at -61.

While prokaryotes are much more complicated than biochemists might prefer, they are also very economical in their metabolic activities. What purpose is served by the existence of two extra CAP binding sites and DNA looping at the lac promoter? A single CAP at the -61 site can activate RNA polymerase initiation of transcription from the P1 promoter on a 203 fragment. Does the proposed loop structure have any observable biological role in regulating RNA polymerase function at the lactose operon? This question will be the topic of further investigation in the following chapter.

CHAPTER III

THE LOOP-STABILIZED SECOND CAP AT THE *lac* PROMOTER PARTICIPATES IN ACTIVATION OF TRANSCRIPTION INITIATION FROM +1

CHAPTER III

INTRODUCTION

The molecular mechanisms involved in the stimulation of transcription by CAP/cAMP at catabolite sensitive promoters remain the object of much study. Various models have been proposed, including: (1) conformational changes in RNA polymerase induced by polymerase-CAP contacts that make the enzyme more efficient; (2) structural changes in the DNA that accompany CAP binding, which may be transmitted to the polymerase site in the promoter and facilitate open complex formation (this generally involves the notion of "action-at-a-distance", and of possible melting of the DNA by CAP); (3) the exclusion by CAP of RNA polymerase binding to a non-productive site thereby making more enzyme available for transcription from the active promoter region; and (4) CAP-induced bending of DNA which might alter polymerase-DNA contacts and enhance initiation of RNA synthesis, or which may lead to formation of a DNA loop that somehow affects transcription initiation. The situation is made even more complicated by the fact that CAP binds at different positions relative to RNA polymerase at different promoters, and that there may be more than one CAP molecule involved in some cases. The action-at-a-distance model is not supported

by the finding that CAP is not a helix-destabilizing protein (110, 111), though it remains an interesting possibility in other systems. At both the lac and gal promoters it is known that CAP prevents RNA polymerase from binding at the P2 site, thereby shifting it to initiation from P1, at nucleotide +1, although the exclusion effect does not fully account for the observed level of stimulation. There is strong evidence for direct CAP-RNA polymerase contacts from studies using CAP and polymerase mutants (Chapter I). And we have presented evidence above favoring the involvement of CAP in loop formation at the lac promoter, provided that the appropriate upstream DNA sequences are present. In this chapter we investigate the potential effects of the DNA loop on transcription initiation, and report the discovery of previously unknown lac transcription complexes that require CAP for their formation but which can initiate transcription in the absence of the activator protein.

MATERIALS AND METHODS

Materials. Proteins and lac DNA fragments were prepared as described in Chapter II. 5′-Ribonucleoside triphosphates were from Pharmacia Biochemical. The dinucleotide ApA and UpU were from Sigma. The ^{32}P-nucleotides (3000 Ci/mmole) were from NEN.

Binding Assays by Gel Retardation. All binding reaction and gel retardation assays were performed as described in

Chapter II using ^{32}P-labelled lac DNA fragments. Gel electrophoresis was done in mixed polyacrylamide-agarose gels, except those involving the 203 lac DNA fragment. 203 DNA open complexes were resolved using a 0.1 x 20 cm vertical 5% polyacrylamide gel (acrylamide:bisacrylamide = 40:1) in TBE buffer in the presence of cAMP. To isolate open complexes, all gels were electrophoresed at 20 mA until the xylene cyanol marker dye ran off the bottom of the gel. Desired complex bands were located and excised from the gel as previously described.

The rate of open complex formation was estimated by means of a gel shift assay. Radioactive labelled DNA fragment was incubated with CAP for 15 min, RNA polymerase was added to the mixture and incubation continued at 37°C. At various times after polymerase addition, aliquots were withdrawn, quenched with heparin and Ficoll loading buffer, and loaded onto the running retardation gel. Following electrophoresis, the gel was dried and autoradiographed, the free DNA and complex bands were located and excised, and their radioactivity determined by scintillation counting.

Exonuclease III Protection Assays. DNA fragments were labelled at the 5′-end as described in Chapter II, then digested with an appropriate restriction enzyme at a site near the upstream end of the fragment. DNA at a concentration of 20 nM was used in each binding reaction. Following the incubation period, reactions containing only

DNA or DNA plus CAP were directly treated with exonuclease III. To reactions containing RNA polymerase were added either 50 μM poly d(A-T) (followed by incubation for 5 min at 37°C) or 100 μg/ml heparin (followed by incubation of less than 1 min at 37°C). Then exonuclease III digestion was performed, using 1 unit/μl Exo III enzyme for poly d(A-T)-competed reactions, or 5 units/μl Exo III for heparin-competed reactions. (The enzyme was diluted from stock solution into double distilled water.) Digestion times were from 1 to 10 min at 37°C. The reaction was stopped by mixing with an equal volume of stop buffer (10 M urea, 50 mM EDTA, 0.1% SDS, and 50 mM NaOH), followed by heating for 3 min at 97°C and denaturing gel electrophoresis on a sequencing gel. Localization of the protein-protected sites on the DNA fragment was determined by comparison with Maxam-Gilbert sequencing reactions for that fragment.

In Vitro Transcription Assays.

a) *Runoff transcription in solution.* A runoff transcription assay can be used to estimate the strength of a promoter of interest. The presence of heparin or other competitor in the reaction guarantees that each active promoter complex produces only one transcript, since polymerase cannot then reassociate with the promoter after it catalyzes the synthesis of a single transcript and dissociates from the DNA. Thus, comparison of the number of transcripts produced in each reaction is a direct measure of

the number of functional open complexes present in the solution, thereby indicating how active a promoter fragment is in binding transcription factors and RNA polymerase. The promoter strengths of various lac fragments were tested by the following procedure: 10 nM lac DNA fragment was mixed with CAP and RNA polymerase as usual in binding buffer, in a 10 μl reaction volume. At the end of the incubation period, 2.5 μl of nucleotide mixture was added (working concentration: 200 μM GTP, ATP, and CTP; 25 μM UTP; 0.5 μl (α-^{32}p)UTP (5 μCi); 100 μg/ml heparin in binding buffer) to the reaction, and mixed gently. After 10 min at 37°C, the reaction was quenched with an equal volume of RNA stop buffer (10 M urea, 50 mM EDTA, 0.1% SDS, 0.1% bromphenol blue, and 0.1% xylene cyanol), and heated for 3 min at 97°C before loading onto a denaturing gel. RNA transcripts were resolved on a 5% sequencing gel and visualized via autoradiography. Gel slices containing transcript bands were excised and quantified by liquid scintillation counting.

b) *Transcription assay of gel-purified open complexes.* Gel-purified open complexes were obtained as follows. Labelled DNA fragments (30 nM) were incubated with CAP and RNA polymerase as usual, then heparin and Ficoll loading buffer were added and the solutions were electrophoresed through a nondenaturing gel. Complex bands were visualized by autoradiography at room temperature for 30-60 min. Gel

slices containing desired complex bands were cut out and placed in separate 1.5 ml tubes; the amount of DNA in each gel slice was quantified in a scintillation counter with an open window (no scintillation fluid added).

The transcription assay was performed with only three of the four NTPs as described in (28) with some modification. The nucleotide solutions were as follows:

NTP mixture for P1 transcript detection: 30 mM $MgCl_2$, 500 μM ApA, 120 μM ATP and GTP, 60 μM UTP, and 5 μCi (α-^{32}P) UTP in binding buffer.

NTP mixture for P2 transcript detection: 30 mM $MgCl_2$, 500 μM UpU, 120 μM CTP and GTP, 60 μM UTP, and 5 μCi (α-^{32}P) UTP in binding buffer.

A 10 μl aliquot of NTP mixture was spread on the gel slice in each tube, followed by incubation at 37°C for the designated length of time. The reaction was stopped with 20 μl RNA stop buffer (see above). Reaction tubes were kept at room temperature overnight to allow the transcripts to diffuse out of the gel slices (in fact, there is an equilibrium reached in which the (small) transcripts are uniformly distributed in solution within and outside the gel). To detect transcripts produced by complexes in the various gel slices, equal volumes of solution from each tube were loaded onto a 0.04 x 40 cm 20% sequencing gel (acrylamide:bisacrylamide = 20:3), which was run at 40 watts until the bromphenol blue was about 3/4 of way down the gel.

The wet gel was autoradiographed at -70°C with an intensifying screen; keeping the gel frozen reduces band spreading from diffusion of the small transcript molecules. Gel slices containing transcript bands were excised, and the numbers of transcripts in each reaction were quantified by scintillation counting. Useful comparisons of the data require that the original gel slices be of equal size. This was achieved by careful cutting of gel slices of the transcript bands and the background control slices.

RESULTS

Different Length lac Fragments Can Participate in Several Kinds of Open Complexes. The strength of a promoter can be characterized by its ability to participate in formation of heparin-resistant open complexes, or by its efficiency in initiating transcription. The gel mobility shift assay was used to measure the rate of open complex formation at lac promoters on different DNA fragments. The results are shown in Figures 17, 18, and 19. The data illustrate that both the 367 and 789 lac fragments form at least two complexes, designated M_o and L_o (Figures 18 and 19) plus a minor species (U_o), while the 203 fragment appears from this experiment to participate in only one kind of complex (Figure 17). Figure 20 shows that, when the CAP input is low enough, the 203 lac fragment can also form two complexes, which are equivalent to the M_o and L_o complexes

Figure 17. Gel retardation assay of the rate of open complex formation with lac 203 DNA. Lane 1, 203 DNA only; lane 2, 203 DNA plus CAP; lanes 3 through 12, 203 DNA plus CAP, then RNA polymerase, with reactions quenched at various times. The "stair-step" pattern occurs because samples were loaded onto the gel at different times, immediately after being quenched.

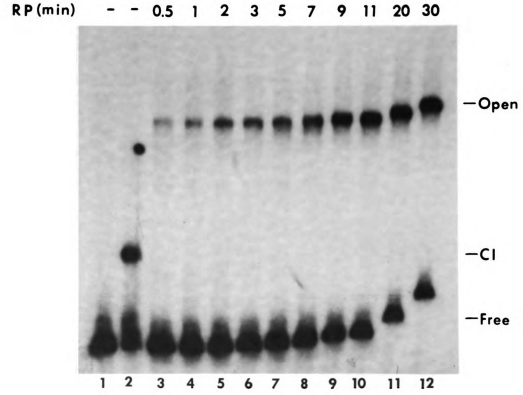

RP(min) - - 0.5 1 2 3 5 7 9 11 20 30
—Open
—CI
—Free
1 2 3 4 5 6 7 8 9 10 11 12

Figure 18. Gel retardation assay of the rate of open complex formation with lac 367 DNA. Lane 1, 367 DNA only; lane 2, 367 DNA plus CAP; lanes 3 through 12, 367 DNA plus CAP, then RNA polymerase, with reactions quenched at various times. U_o, M_o, L_o denote the heparin-resistant polymerase-containing complexes.

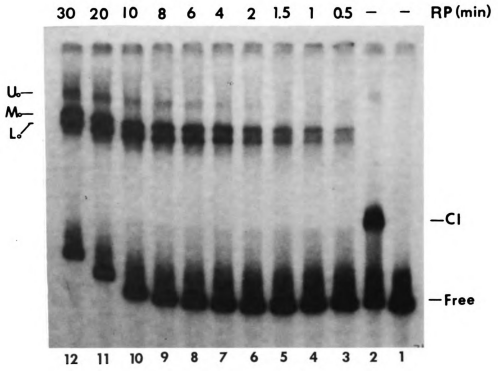

30 20 10 8 6 4 2 1.5 1 0.5 – – RP (min)
Uo
Mo
Lo
—CI
—Free
12 11 10 9 8 7 6 5 4 3 2 1

Figure 19. Gel retardation assay of the rate of open complex formation with lac 789 DNA. Lane 1, 789 DNA only; lane 2, 789 DNA plus CAP; lanes 3 through 12, 789 DNA plus CAP, then RNA polymerase, with reactions quenched at various times. U_o, M_o, L_o denote the heparin-resistant polymerase-containing complexes.

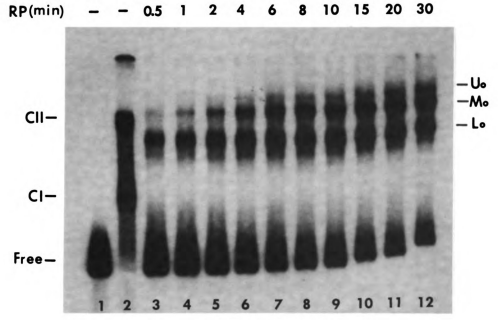
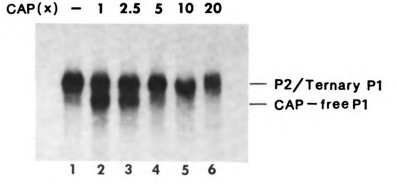

RP(min) – – 0.5 1 2 4 6 8 10 15 20 30
CII–
CI–
Free–
–Uo
–Mo
–Lo
1 2 3 4 5 6 7 8 9 10 11 12

Figure 20. Gel retardation assay showing two different heparin-resistant complexes of 203 lac DNA with RNA polymerase and CAP. Lane 1, 203 DNA plus polymerase; lanes 2 through 6, 203 DNA plus different amounts of CAP as indicated, then plus polymerase. Complexes were resolved using a 20 cm gel of 5% acrylamide:bis-acrylamide (40:1) in TBE buffer. (Note: The gels used in Figures 17-19 were mixed polyacrylamide-agarose.) As described later in the text, the lower band seen in lanes 2 and 3 was found to contain RNA polymerase that initiates transcription from +1 (P1), but no CAP. The upper band seen in these lanes contains either RNA polymerase-203 DNA complexes that initiate from P2, or CAP-polymerase-lac promoter P1 complexes.


CAP(x) – 1 2.5 5 10 20
P2/Ternary P1
CAP–free P1
1 2 3 4 5 6

of the longer fragments. Considering, for example, lane 3 of Figure 20, the slower moving complex band apparently corresponds to M_o while the faster moving complex band corresponds to L_o. There seems to be no U_o-type complex detected for the 203 DNA fragment.

Plotting the relative radioactivity of each complex band versus time of incubation with polymerase (from Figures 17-19) generally shows no significant differences in the rates of open complex formation between the three fragments (not shown). An exception is the "lower" open complex (L_o) of the 789 fragment; a substantial fraction of the L_o complex forms within the first 30 seconds of polymerase incubation (Figure 19, lane 3).

The runoff transcription assay was used to measure the transcription efficiency of complexes formed at each lac fragment. Transcript length varies depending on the fragment and on whether initiation occurs from lac P1 or P2. It was therefore necessary to correct the observed level of radioactivity to account for the different total number of radioactive U residues in the various transcripts. This involved normalizing the scintillation counting data to a baseline according to the number of UTPs incorporated in transcripts from each fragment. An example of a typical runoff transcription gel is as shown in Figure 35.

Table II shows runoff transcription data for the 203, 367, and 518 fragments. The 518 DNA was used because there

Table II

Transcription Efficiency of Different *Lac* Fragments

transcripts from reactions of	relative amount of transcripts from		
	203	367	518
15X polymerase (P2)	1	1.07	0.45
10X CAP 15X polymerase (P1)	1	1.10	0.72
20X CAP 15X polymerase (P1)	1	0.94	1.05

is some premature termination when the long 789 fragment is used; moreover it is difficult to quantify the number of full length transcripts from 789 DNA. The data reveal that the 518 fragment produces fewer P2 transcripts than do the 203 and 367 fragments. When CAP input is in 10-fold excess over promoter fragment concentration, the 518 DNA fragment produces somewhat fewer P1 transcripts than the other two fragments do. However, with doubled CAP input (20X), more P1 transcripts were produced from each fragment, and the lac 203, 367 and 518 DNA fragments yield about equal numbers of P1 transcripts.

The results presented in Table II imply that lac 203, 367, and 518 DNA fragments are potentially equally active in initiating CAP-activated transcription from +1. However, this interpretation is complicated by the fact that at least two different heparin-resistant complexes are involved (M_o, L_o). Each fragment can form more than one complex in the presence of CAP and RNA polymerase, and it is not clear whether each complex is active in transcription from P1 (though no P2 transcripts are evident in the gels at the CAP concentrations used in Table II). If both M_o and L_o complexes are active in initiation from P1, what is the relative efficiency of each? This question is considered below.

It is also worth noting that the longer fragment (518) produces relatively fewer P2 transcripts (in the absence of

CAP) than do the shorter DNAs, while not much difference was found in CAP-activated P1 transcription (Table II). This suggests that as the lac fragment gets longer, the DNA fragment conformation may change in a way that makes the lac P2 promoter less accessible for RNA polymerase binding. This might be the reason why RNA polymerase does not initiate transcription from P2 efficiently in vivo when the lac operon is in an intact supercoiled plasmid (122).

Lac Promoter DNA Participates in a CAP-Free and in a CAP-Containing Open Complex in the Presence of RNA Polymerase. Formation of transcriptionally-competent complexes at the lac P1 promoter requires pre-incubation of lac DNA fragments with CAP prior to the addition of RNA polymerase. In this situation the only transcripts obtained are from P1, while P2 is repressed. Furthermore, there is evidence from footprinting experiments that CAP is present in heparin-resistant open complexes; that is, there is stabilization of CAP binding by RNA polymerase, and protection from nuclease digestion is seen at -61 (100).

The data in Chapter II reveal that more than one kind of CAP-lac DNA complex can be formed in vitro, depending on the length of the DNA fragment and the concentration of CAP used. And, as we have just seen, it appears that two or more types of open complexes (at P1) are found when RNA polymerase is present. What are the differences between

these transcription complexes? Is there a relationship between the CI and CII CAP-DNA complexes and the several P1 open complexes that have been described above? To address these questions, additional gel shift experiments were done (at different CAP concentrations), along with transcription studies and stoichiometry determinations on gel-purified complexes.

The relative amounts of each complex formed when RNA polymerase is added to solutions of lac DNA and CAP are a function of the CAP concentration in the reaction (Figures 20, 21 & 22). For each lac DNA fragment, at low CAP input a faster moving complex ("L_o") in Figure 21 predominates, while with increasing CAP, the amount of higher mobility complexes decreases and the level of slower migrating complexes ("M_o") increases. A third complex with an even lower mobility (U_o) is formed in small amounts, but is not considered further here. As will be seen shortly, the M_o band contains CAP-polymerase-DNA P1 complexes, the L_o band contains polymerase-DNA P1 complexes but no CAP. The mobilities of these polymerase-DNA complexes relative to that of the corresponding P2 open complex differ from fragment to fragment. For 203 DNA, the P2 open complex has the same mobility as the slower migrating complex formed at high CAP input (Figure 20, lane 1); for the 789 fragment, the P2 complex migrates at the same rate as the faster moving P1 complex (Figure 21, lane 6); the P2 complex of 367

Figure 21. Gel retardation assay of the formation of lac 789-polymerase complexes as a function of CAP concentration. The "+" and "-" symbols denote the presence or absence of the indicated components. Lane 1, 789 DNA alone; lanes 2 through 5, 789 DNA plus increasing amounts of CAP as indicated (no RNA polymerase); lane 6, 789 DNA plus polymerase; lanes 7 through 10, 789 plus increasing amounts of CAP as indicated, then plus polymerase. Binding reactions involving polymerase were treated with heparin before loading on the binding gel, to eliminate nonspecific binding of RNA polymerase. See text for evidence regarding the composition of each RNA polymerase-containing band.

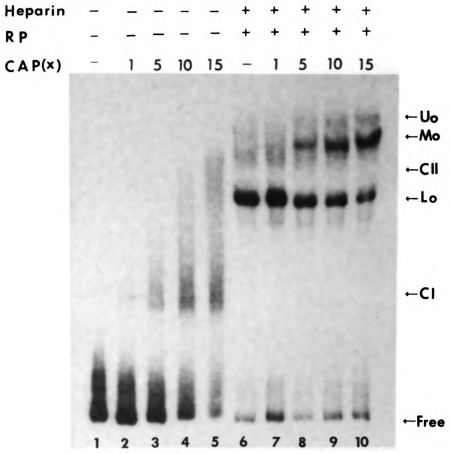
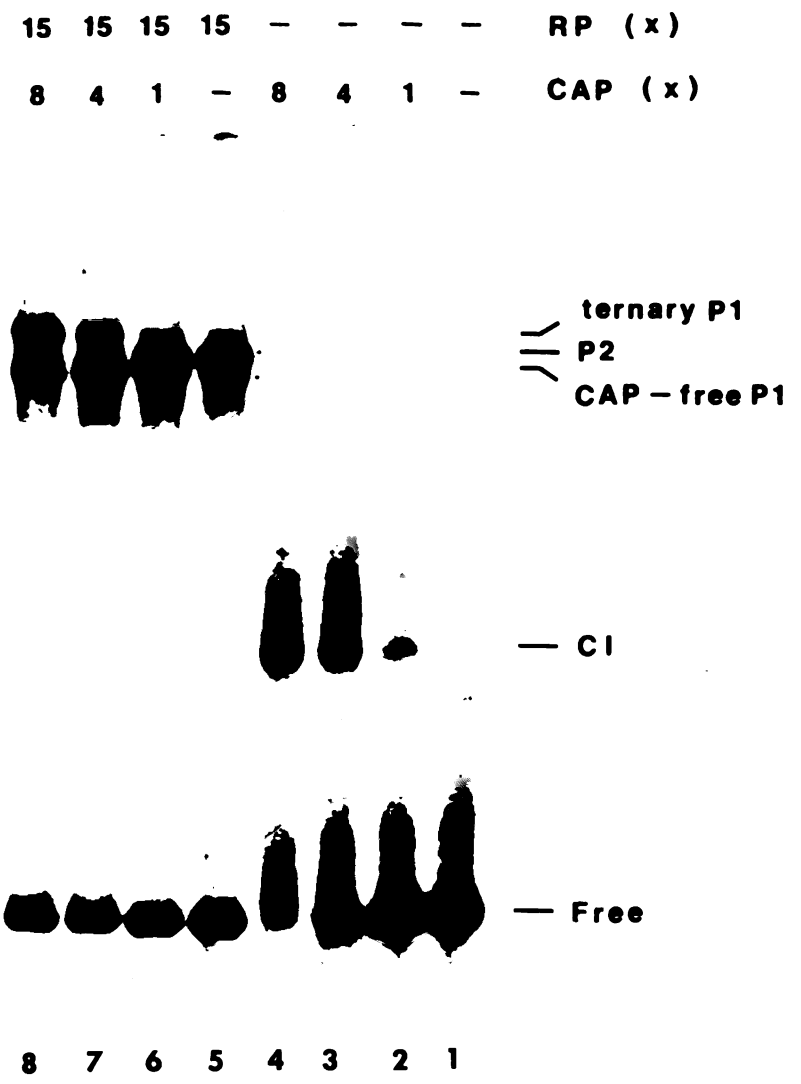

Heparin – – – – – + + + + +
RP – – – – – + + + + +
CAP(x) – 1 5 10 15 – 1 5 10 15
←Uo
←Mo
←CII
←Lo
←CI
←Free
1 2 3 4 5 6 7 8 9 10

Figure 22. Gel retardation assay of the formation of lac 367-polymerase complexes as a function of CAP concentration. Lane 1, 367 DNA alone; lanes 2 through 4, 367 DNA plus increasing amounts of CAP as indicated (no RNA polymerase); lane 5, 367 DNA plus polymerase; lanes 6 through 8, 367 plus increasing amounts of CAP as indicated, then plus polymerase. Binding reactions involving polymerase were treated with heparin before loading on the gel, to eliminate nonspecific binding of RNA polymerase. See text for evidence regarding the composition of each RNA polymerase-containing band.


15 15 15 15 – – – – RP (x)
8 4 1 – 8 4 1 – CAP (x)
ternary P1
P2
CAP – free P1
– CI
– Free
8 7 6 5 4 3 2 1

DNA migrates between the two complexes seen when CAP is present (Figure 22, lane 5).

An immediate inference from these band shift patterns is that the slower migrating complex contains more CAP than the faster moving complex, hence has a reduced mobility. We thus turned to characterizing the complexes in terms of stoichiometry and transcription properties.

Transcription assays were performed on gel purified complexes to identify which are active and how active they are. Using the dinucleotide primers ApA and UpU, and only three of the four NTPs in the reactions, permits us to determine whether complexes initiate at P1 or P2. In the presence of ApA but absence of CTP, initiation occurs from +1, and a 9-nucleotide Primary transcript is made. The lack of CTP prevents the formation of any transcripts longer than a dimer that would originate at P2 (see lac promoter sequence below). Likewise, when UpU is present and ATP is omitted from the reaction, initiation is from -22 (P2), and an 11-nucleotide transcript is produced. The absence of ATP means that no transcripts at all are made from the P1 promoter.

```
-35          UpU                    ApA          +15
.                                                .
TTACACTTTA TGCTTCCGGC TCGTATGTTG TGTGGAATTG TGAGCGGATA
AATGTGAAAT ACGAAGGCCG AGCATACAAC ACACCTTAAC ACTCGCCTAT
```

Sequence of the lac promoter from -35 to +15.

The results presented in Figure 23 (lanes 2 & 3), Figure 24 (lanes 3, 4, & 5), and Figure 25 (lanes 5, 6, & 7) show

Figure 23. Transcription assays of gel-purified lac 203 heparin-resistant open complexes. Lanes 1 and 4, P2 complex formed by incubating the 203 DNA plus polymerase in the absence of CAP; lanes 2 and 5, the faster migrating (lower "L_o" band) complex observed in solutions of 203 DNA plus CAP plus polymerase; lanes 3 and 6, the slower migrating complex (middle "M_o" band) seen in solutions of 203 DNA plus CAP plus polymerase. ApA and UpU are used in detecting P1 and P2 complexes as indicated. The subscript "$_o$" denotes open complex.

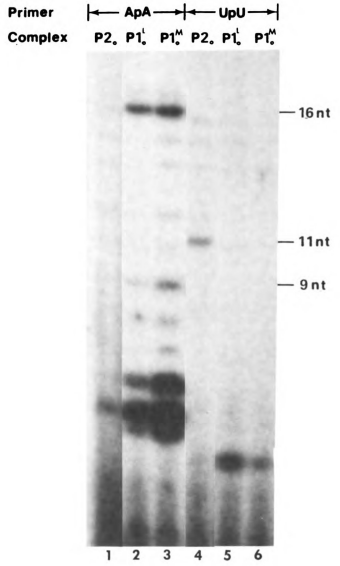

Primer
ApA
UpU
Complex
P2∘
P1∘I
P1∘M
P2∘
P1∘I
P1∘M
16nt
11nt
9nt
1
2
3
4
5
6

Figure 24. Transcription assays of gel-purified lac 367 heparin-resistant open complexes. Lanes 1 and 6, P2 complex formed by incubating 367 DNA plus polymerase in the absence of CAP; lanes 2 and 7, P2 complex observed in solution of 367 DNA plus CAP(1X) plus polymerase (this is the middle complex band in Figure 22); lanes 3, 4, 8, and 9, faster migrating complexes (lower "L_o" band) seen in solutions of 367 DNA plus CAP(1X or 8x) plus polymerase; lanes 5 and 10, slower migrating complex (middle "M_o" band) formed in solutions of 367 DNA plus CAP(8X) plus polymerase. The $P2_o$ complexes in lanes 1 and 2 are contaminated by the L_o complex because they do not separate well on a gel as shown in Figure 22.

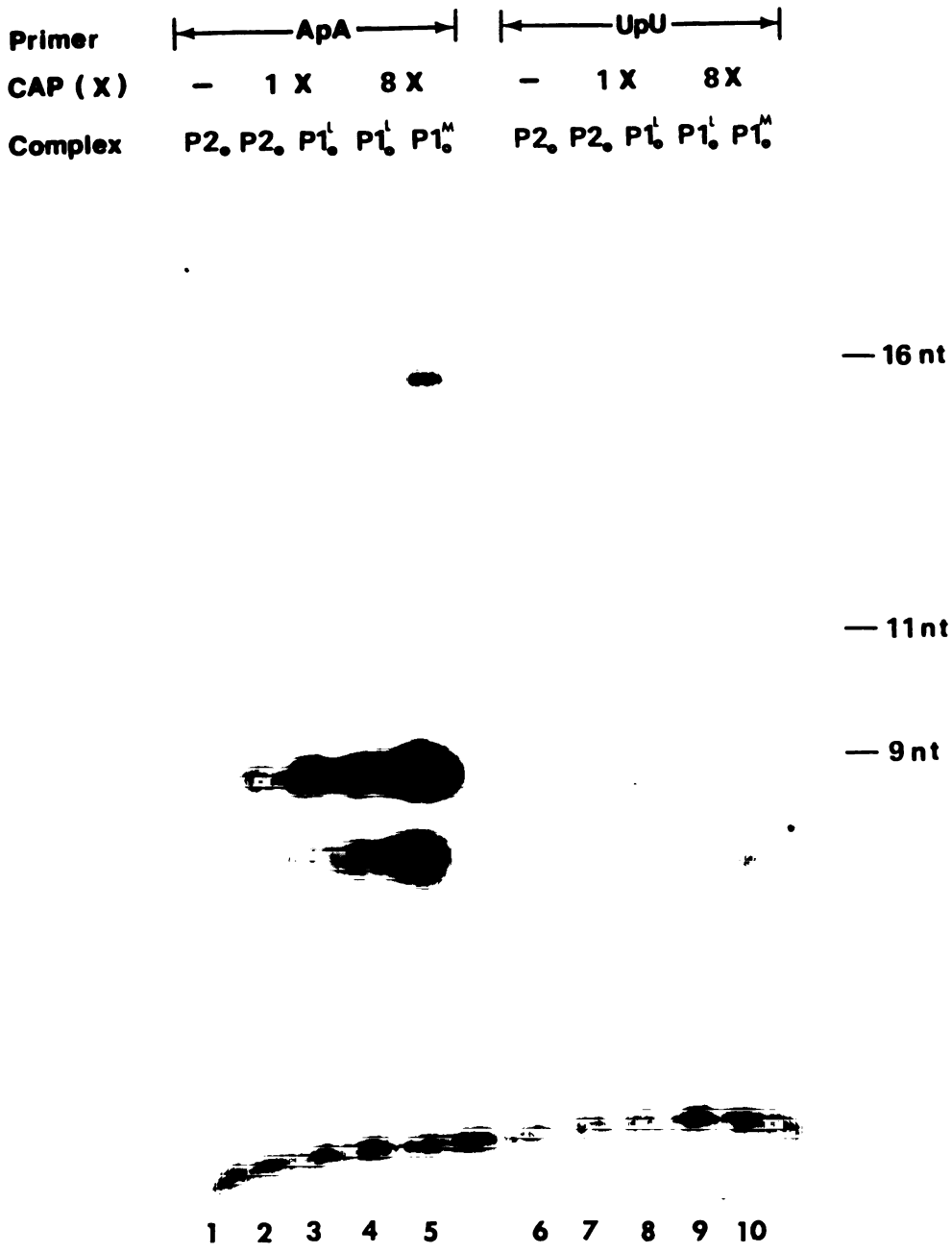
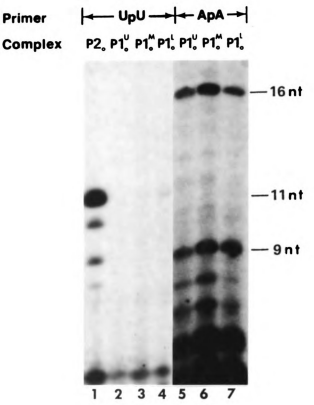

Primer
ApA
UpU
CAP (X)
– 1 X 8 X – 1 X 8 X
Complex
P2o P2o P1Lo P1Lo P1Mo P2o P2o P1Lo P1Lo P1Mo
16 nt
11 nt
9 nt
1 2 3 4 5 6 7 8 9 10

Figure 25. Transcription assays of gel-purified lac 789 heparin-resistant open complexes. Lane 1, P2 complex formed by incubating 789 DNA plus polymerase; lanes 2 and 5, the slowest migrating complex seen in solutions of 789 DNA plus CAP plus polymerase ("U_o" band); lanes 3 and 6, the middle complex band observed in solutions of 789 DNA plus CAP plus polymerase ("M_o" band); lanes 4 and 7, the fastest migrating complex formed in solutions of 789 DNA plus CAP plus polymerase ("L_o" band).


Primer
UpU
ApA
Complex
P2∘ P1∘U P1∘M P1∘I P1∘U P1∘M P1∘I
16 nt
11 nt
9 nt
1 2 3 4 5 6 7

that for all three lac DNA fragments, both the faster and slower migrating complexes ("L_o" and "M_o") formed in the presence of CAP are active P1 complexes. Initiation from +1 is seen when ApA is used as primer, but there are no transcripts when UpU is the primer. [The presence of transcripts that are longer than predicted (16-nt long) is likely due either to impurities in the purchased ribonucleotides, or to occasional incorporation of a wrong nucleotide by RNA polymerase; the shorter transcripts result from abortive initiation (cycling) by the enzyme.]

The Faster Migrating P1 Complexes are CAP-free. Mixing RNA polymerase with lac DNA fragments in vitro leads to formation of stable complexes that initiate from the P2 promoter; very few P1 transcripts are seen. If CAP is then added, some P1 transcripts are made; this is thought to involve CAP-stimulated melting in of preformed closed complexes at the P1 site (79). However, if CAP and lac promoter DNA are mixed prior to addition of RNA polymerase, then the enzyme forms only P1 complexes (P2 is excluded by the bound CAP). A wealth of evidence shows that CAP and RNA polymerase bind to the lac promoter DNA cooperatively. For example, a CAP molecule bound at the -61 site is sensitive to heparin, but upon polymerase binding at P1, CAP binding at -61 is stabilized by polymerase and becomes heparin-resistant (100, 112, 113). In the current studies, we find that the ratio of the two most abundant P1 complexes (M_o and

L_o bands) is directly affected by the CAP concentration in the reaction. Thus the CAP stoichiometry in these complexes is of much interest.

The structure of these P1 open complexes was probed using DNase I footprinting. As seen in Figure 26, the faster migrating heparin-resistant lac 789 P1 complex (L_o band) shows DNA protection from about -45 to +20 (lane 3); the usual CAP site (-50 to -70) is not protected in this complex. The slower moving P1 complexes (M_o and U_o bands) of 789 DNA have the usual ternary complex DNase I protection pattern (lanes 1 and 2); their DNA sequences are protected by CAP and polymerase from -70 to +20. The M_o and U_o complexes may somehow differ in conformation, perhaps in a way which is not crucial for complex function, because no differences are found in footprinting assays, and both are active P1 complexes. The stoichiometry of the U_o complex was not studied because there was insufficient material; this complex is ignored due to its low abundance and similarity to the M_o complex.

The above results strongly suggest that, in the presence of heparin, the faster migrating (L_o) P1 complex has no CAP bound at the -61 site. To investigate this hypothesis, gel purified P1 complexes were loaded onto an SDS gel and silver stained following the quantification procedures described in Chapter II. The SDS gel electrophoresis results are shown

Figure 26. DNase I footprinting assay of gel-purified lac 789 DNA complexes with CAP and RNA polymerase. CAP was incubated with 789 DNA, then RNA polymerase was added to reactions. Heparin was added before loading samples on the binding gel. Three polymerase-containing complexes with the lac 789 fragment were resolved (cf. Figure 19): lane 1, the upper complex ($P1_U$) band is the slowest migrating CAP-polymerase-DNA complex; lane 2, the middle complex ($P1_M$) band; lane 3, the fastest migrating polymerase-containing complex band ($P1_L$); lane 4, free 789 DNA control; lane 5, DNA length markers.

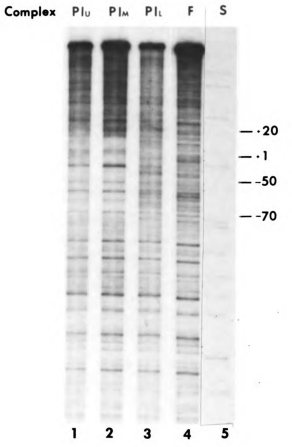

Complex
PIU
PIM
PIL
F
S
-20
-1
-50
-70
1
2
3
4
5

Figure 27. SDS-polyacrylamide gel electrophoresis assay to identify CAP content in lac 203 P1 open complexes. Proteins were visualized by silver staining. Lane 1, faster migrating P1 complex (L_o band) from 203 DNA plus CAP plus RNA polymerase reaction; lane 2, slower migrating P1 complex (M_o band) from the reaction as in lane 1; lane 3, P2 complex control. All complexes were purified in the presence of heparin.

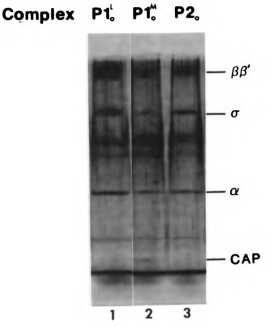

Complex P1°. P1°° P2°.
ββ'
σ
α
CAP
1 2 3

Figure 28. SDS-polyacrylamide gel electrophoresis assay to identify CAP content in lac 367 P1 open complexes. Proteins were visualized by silver staining. Lane 1, CI complex formed with 367 DNA plus CAP (no RNA polymerase); lane 2, P2 complex formed in solutions of 367 DNA plus polymerase; lane 3, slower migrating P1 complex (M_o) from 367 DNA plus CAP plus polymerase reaction; lane 4, faster migrating P1 complex (L_o) from reaction as in lane 3; lane 5, CAP standard; lane 6, RNA polymerase standard. All open complexes were purified in the presence of heparin.

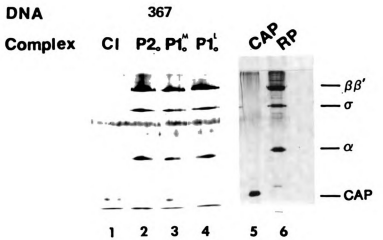
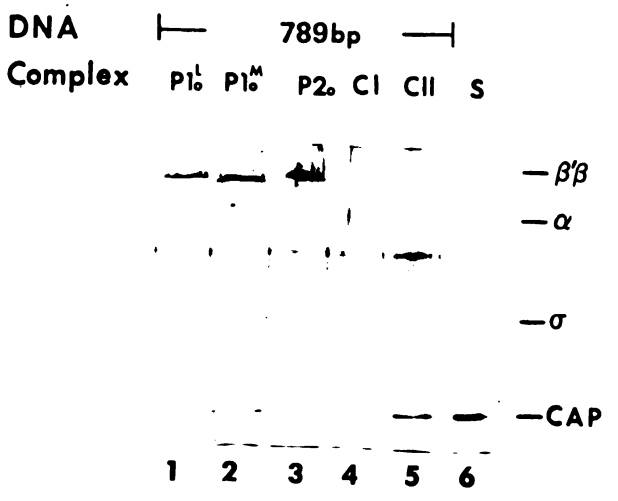

DNA
367
Complex
CI
P2
P1
P1
CAP
RP
ββ'
σ
α
CAP
1
2
3
4
5
6

Figure 29. SDS-polyacrylamide gel electrophoresis assay to identify CAP content in lac 789 open complexes. Lane 1, fastest migrating P1 complex band (L_o) formed in solutions of 789 DNA plus CAP plus RNA polymerase; lane 2, the middle P1 complex band (M_o) from the reaction as in lane 1; lane 3, P2 complex from 789 DNA plus polymerase reaction (no CAP); lane 4, CI CAP complex formed in solution of 789 DNA plus CAP (no polymerase); lane 5, the slower migrating CII CAP complex from the 789 DNA plus CAP reaction; lane 6, CAP only control. All open complexes were purified in the presence of heparin.


DNA
789bp
Complex
P1L
P1M
P2
CI
CII
S
β'β
α
σ
CAP
1
2
3
4
5
6

in Figures 27 to 29. The L_o complexes of 203 DNA (Figure 27, lane 1), of 367 DNA (Figure 28, lane 4), and of 789 DNA (Figure 29, lane 1), as predicted, contain no CAP.

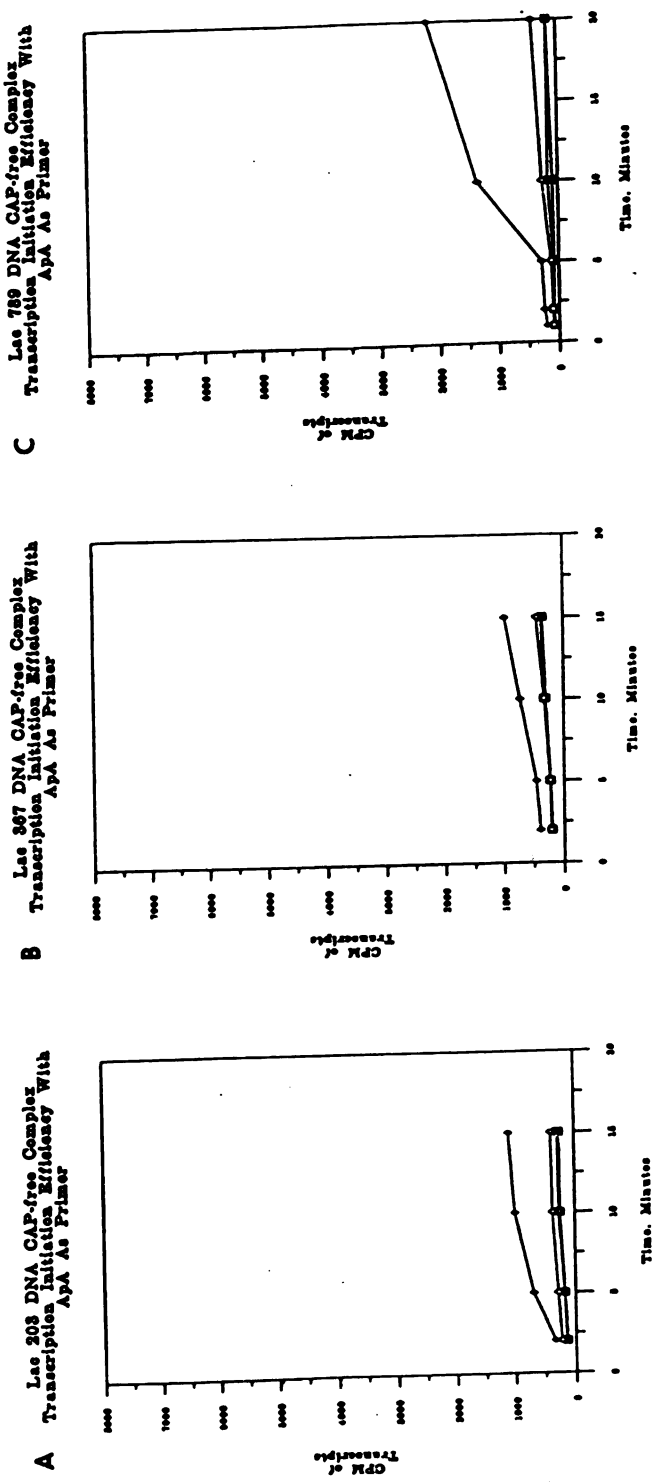
The stoichiometry of CAP in CAP-containing complexes formed with the three lac DNA fragments was determined by quantitative comparison of the amount of CAP in each gel slice with standards (as described in Chapter II, Materials and Methods). Results are presented in Table I (Chapter II). The 203 DNA open complex (M_o) contains one CAP per promoter fragment, while the 367 DNA and 789 DNA M_o complexes have two CAP molecules per fragment. These are therefore ternary complexes in that they contain lac DNA, RNA polymerase, and CAP. The CAP content of each lac ternary complex coordinates well with the corresponding CI complex CAP stoichiometry described in the previous chapter (Table I). Thus polymerase may form the ternary complex by binding to the CI complex that has been formed during the incubation of CAP with DNA prior to addition of the enzyme.

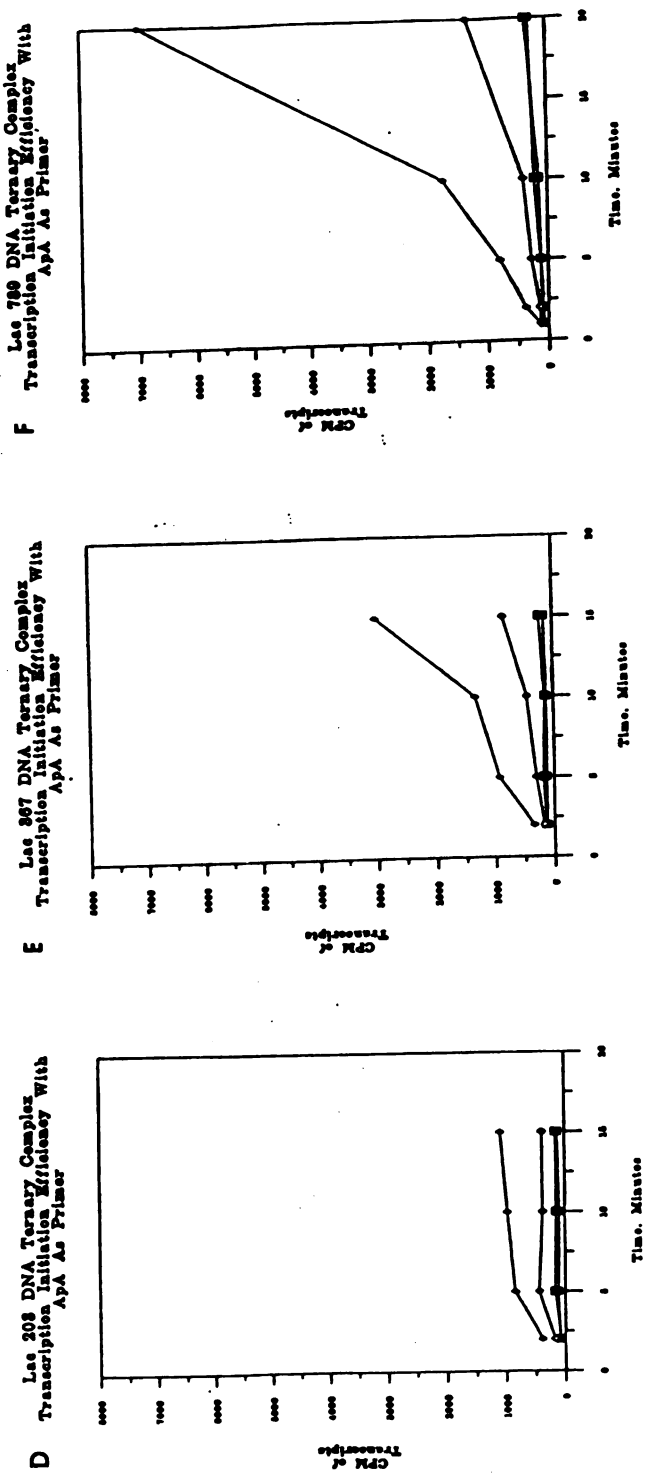
Ternary Complexes Containing Two CAP Molecules Show Elevated Transcription Initiation Efficiency. As demonstrated in Chapter II, the length of lac promoter fragments affects CAP binding affinity---the longer the DNA the more stable its CAP complex. Does the length of the lac DNA also affect transcription initiation efficiency of either P1 complex? To answer this question, labelled DNA fragments were incubated with CAP and RNA polymerase, the

CAP-free and ternary P1 complexes were isolated by gel electrophoresis, and the DNA in each band was quantified by scintillation counting. In-gel transcription assays primed with ApA were performed with each complex; the transcription reaction was stopped at different times with RNA stop buffer as described in Materials and Methods. The transcripts produced by each complex band were resolved on a 20% denaturing gel and analyzed quantitatively by scintillation counting. Although 9-nucleotide transcripts are expected to be synthesized, for unknown reasons, RNA polymerase tends to abort transcription more frequently in these experiments, which leads to the presence of abundant shorter 5- and 6-nucleotide transcripts and relatively fewer 9 nt transcripts. Since polymerase complex is at abortive initiation stage no matter transcripts are 9 nt long or shorter, it is reasonable to evaluate complex transcription initiation efficiency based on data of transcripts which are 5- or 6-nucleotide long. Data were normalized to account for the different amounts of DNA in each complex. The results presented in Figure 30 are based on 0.1 picomole of lac promoter fragment.

The transcription initiation efficiencies of P1 open complexes formed on the 203, 367, and 789 lac promoter fragments are shown in Figure 30. These data indicate that, in general the ternary complex at a longer fragment has a higher transcription initiation efficiency than that at a

Figure 30. Summary of gel-purified lac DNA ternary and CAP-free P1 open complex transcription initiation efficiency. Time "zero" was the point at which nucleotides were added to complexes. Reactions were stopped at indicated later times, and transcripts resolved on a denaturing gel, then quantified by scintillation counting. Only short transcripts are obtained, since CTP is omitted from the reaction. The level of each short RNA was measured. Panel A, 203 CAP-free P1 complex; Panel B, 367 CAP-free P1 complex; Panel C, 789 CAP-free P1 complex; Panel D, 203 ternary P1 complex; Panel E, 367 ternary P1 complex; Panel F, 789 ternary P1 complex. "■" denotes 16 nt transcripts; "□" denotes 9 nt transcripts; "◆" denotes 6 nt transctipts; "◇" denotes 5 nt transcripts.


A
Lac 203 DNA CAP-free Complex
Transcription Initiation Efficiency With
ApA As Primer
B
Lac 367 DNA CAP-free Complex
Transcription Initiation Efficiency With
ApA As Primer
C
Lac 789 DNA CAP-free Complex
Transcription Initiation Efficiency With
ApA As Primer
CPM of Transcripts
Time, Minutes


D
Lac 208 DNA Ternary Complex
Transcription Initiation Efficiency With
ApA As Primer
CPM of Transcripts
Time, Minutes
E
Lac 267 DNA Ternary Complex
Transcription Initiation Efficiency With
ApA As Primer
CPM of Transcripts
Time, Minutes
F
Lac 789 DNA Ternary Complex
Transcription Initiation Efficiency With
ApA As Primer
CPM of Transcripts
Time, Minutes

shorter fragment (Panels D, E and F). Furthermore, the ternary P1 complexes are more efficient than are the corresponding CAP-free complexes for both the 367 and 789 DNA fragments; comparing Panels B with E, and C with F in Figure 30 shows that ternary complexes generate 2.5 - 3 fold more transcripts than the CAP-free complexes. In separate experiments, using NTPs and ApA at higher concentration, the 789 DNA ternary complex was found to have about a ten-fold increase in initiation efficiency relative to the CAP-free 789 P1 complex (data not shown). On the contrary, there is no significant difference in initiation efficiency between the 203 DNA CAP-free and ternary complexes (Figure 30, Panels A and D). This may reflect the fact the 203 DNA ternary complex has only one CAP molecule bound instead of two.

Exonuclease III protection assay. Exo III protection assays can provide information about the structures of the various transcription complexes. As discussed in Chapter I, heparin is a strong polyanion which is widely used as a competitor in RNA polymerase studies. Heparin can disrupt sequence-specific binding of CAP at the lac promoter if CAP is not stabilized by polymerase. An alternative to heparin is poly d(A-T), which binds RNA polymerase strongly, but which interacts with CAP only in a nonspecific manner, hence does not appreciably affect specific binding of CAP at the lactose promoter. By alternatively using these two

competitors in Exo III studies of polymerase binding and open complex formation in the presence of CAP, we can better characterize these processes.

Exonuclease III digestions were performed by using lac DNA fragments ^{32}P-labelled at the 5'-end of the lower strand, hence provide information as to how far upstream the complexes extend. Data were obtained in the presence of different amounts of CAP, using either poly d(A-T) or heparin as competitor for nonspecific binding. This allowed study of the degree of formation of different open complexes as a function of CAP input, and of the protein-protein interactions in open complexes as well.

As expected from the DNase I footprinting data of Figure 11 (Chapter II) and Figure 26, exonuclease III assays show that the action of this processive nuclease can be halted by bound protein molecules at three different sites on lac DNA fragments, representing four different complexes under various reaction conditions. The nuclease may be stopped upstream of -70 (at -78), which is diagnostic for the CAP CI complex (in solutions containing no polymerase) or the ternary P1 complex. An RNA polymerase bound at P2 would stop Exo III at around -55, while the CAP-free complex with polymerase at P1 stops digestion farther downstream (at about -47).

These results are shown in Figures 31 and 32, which contain data from Exo III protection assays of lac 203 and

Figure 31. Exonuclease III protection assays of lac 203 DNA interacting with CAP and/or RNA polymerase in the presence of poly d(A-T). Lane 1, 203 DNA alone; lane 2, 203 DNA plus polymerase; lane 3, 203 DNA plus CAP; lanes 4 through 7, 203 DNA with increasing amounts of CAP as indicated, then plus RNA polymerase; lanes 8 and 9, G and G+A Maxam-Gilbert sequencing reactions for 203 DNA. All reactions involving RNA polymerase were competed with poly d(A-T) before Exo III digestion.

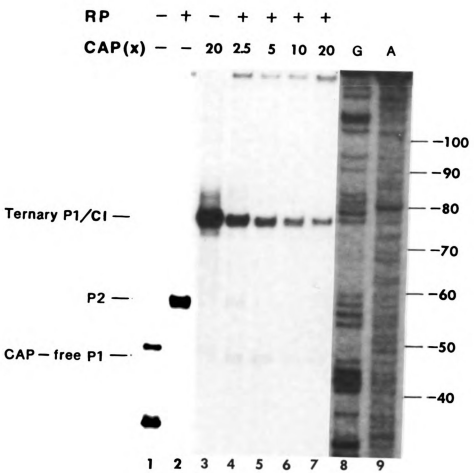
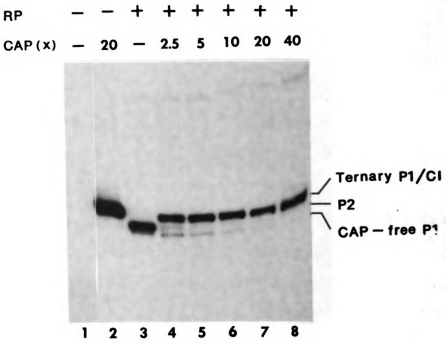

RP − + − + + + +
CAP(x) − − 20 2.5 5 10 20 G A
Ternary P1/CI
P2
CAP − free P1
−100
−90
−80
−70
−60
−50
−40
1 2 3 4 5 6 7 8 9

Figure 32. Exonuclease III protection assays of lac 789 DNA interacting with CAP and/or RNA polymerase in the presence of poly d(A-T). Lane 1, 789 DNA alone; lane 2, 789 DNA plus CAP; lane 3, 789 DNA plus polymerase; lanes 4 through 8, 789 DNA with increasing amounts of CAP as indicated, then plus RNA polymerase. All reactions involving RNA polymerase were competed with poly d(A-T) before Exo III digestion.


RP – – + + + + + +
CAP (x) – 20 – 2.5 5 10 20 40
Ternary P1/CI
P2
CAP – free P1
1 2 3 4 5 6 7 8

789 complexes. When DNA alone is digested, the Exo III digests each strand of the fragment from the 3'-end, meeting in the center region of the fragment and preventing further digestion. Consequently, DNA bands about half the size of the lac fragment are detected on the sequencing gel as shown in Figure 31, lane 1. [Note: occasionally, when excess enzyme is added or the incubation time is long, the DNA fragment appears to be completely digested by the enzyme.] When lac DNA fragments are incubated with CAP, protection by CAP stops the Exo III at around -78 (Figure 31, lane 3; Figure 32, lane 2). When 203 DNA or 789 DNA is incubated with RNA polymerase only, the Exo III is stopped primarily at -55 (P2 complex), with a minor protection at -47, probably from polymerase bound at P1 in the absence of CAP (Figure 31, lane 2; Figure 32, lane 3). These results are consistent with in vitro transcription data showing that lac DNA fragments plus polymerase only produce a large majority of transcripts from P2, only a few from P1 (Figure 35, lanes 1 and 5).

The Exo III protection pattern for complexes containing both CAP and RNA polymerase depends on which competitor was used and how much CAP was present. When poly d(A-T) was used as competitor, at low level of CAP input (2.5X), partial protection is seen at both -55 (the P2 site), and at -47 (the P1 site without CAP), while the major protection is at -78 (the ternary P1 complex with CAP bound)

(lane 4 of Figure 31 and of Figure 32). The protection at both the P2 (-55) and the CAP-free P1 (-47) sites disappears with increased CAP. When sufficient CAP is present (20X), the only protection detected is at -78, representing a ternary P1 complex with CAP bound (lane 7 of Figure 31 and of Figure 32). Figure 33 shows typical data obtained from exonuclease III digestions with heparin as competitor, in this case using lac 203 DNA. At 2.5X CAP input (lane 3), few ternary P1 complexes are detected; however, the amount of CAP-free P1 complex is increased significantly relative to that found in the polymerase-only reaction (lane 2) and in the poly d(A-T)-competed reactions (cf. Figure 31, lane 4). With increased CAP input, the P2 complex eventually disappears, the amount of CAP-free P1 complex diminishes but does not disappear, and the level of the ternary P1 complex increases (lanes 4-6). Thus the CAP-free P1 complex is detected only when heparin is used as competitor.

Similar experiments comparing the competitor effect on P1 complex formation were conducted with the gel retardation assay, as shown in Figure 34. In this case, lac 789 DNA was incubated with CAP, then polymerase was added at time "zero". At various later times, two aliquots of sample were withdrawn from the reaction. Poly d(A-T) (50 μM) was added to one aliquot, heparin (100 μg/ml) to the other. After a 5 min incubation, the solutions were subjected to gel mobility shift electrophoresis in a nondenaturing gel. When poly

Figure 33. Exonuclease III protection assays of lac 203 DNA interacting with CAP and/or RNA polymerase in the presence of heparin. Lane 1, 203 DNA alone in the absence of heparin; lane 2, 203 DNA plus polymerase; lanes 3 through 6, 203 DNA incubated with increasing amounts of CAP as indicated, then plus RNA polymerase; lane 7, Maxam-Gilbert G reaction of the 203 DNA fragment. All reactions containing RNA polymerase were competed with heparin before Exo III digestion.

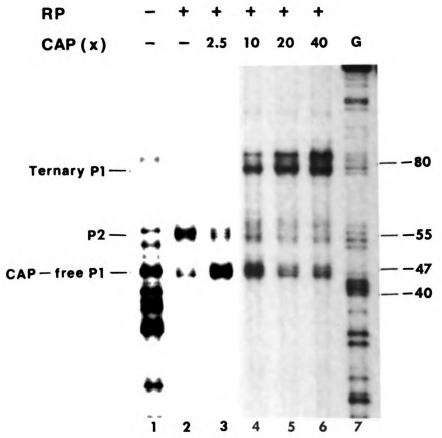
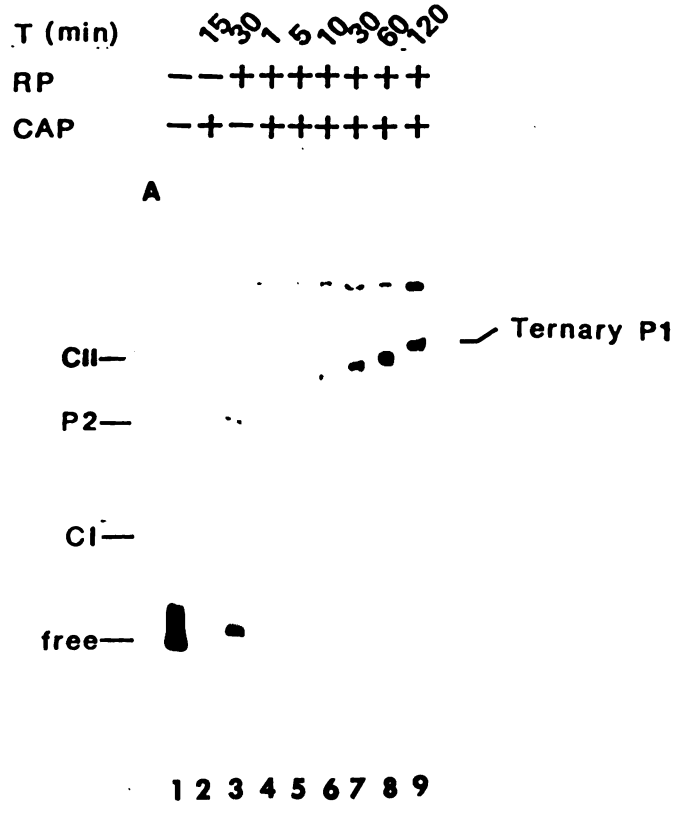
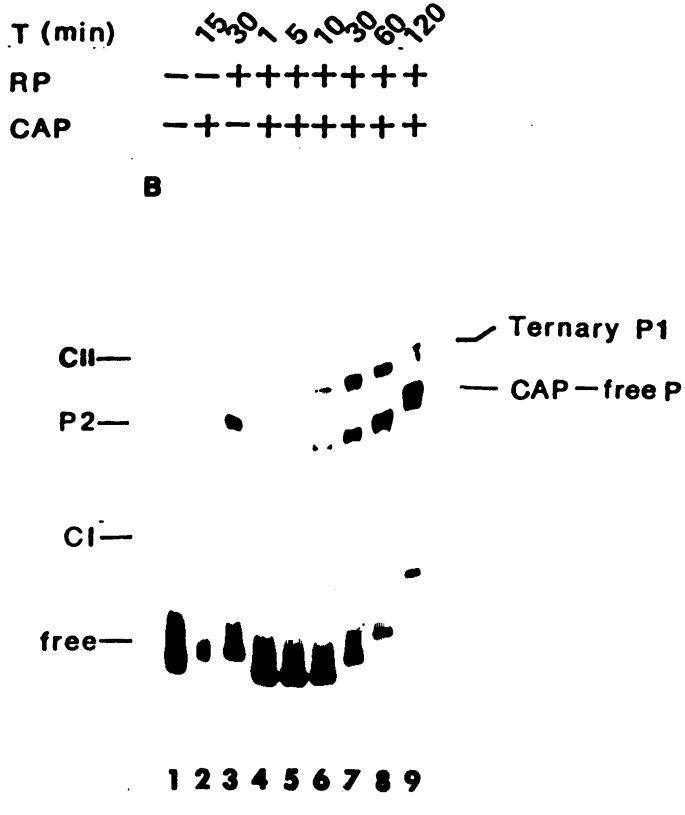

RP
CAP (x)
– + + + + +
– – 2.5 10 20 40 G
Ternary P1
P2
CAP – free P1
80
55
47
40
1 2 3 4 5 6 7

Figure 34. Gel retardation assay of the effect of competitor on lac 789 P1 open complex formation. Panel A, poly d(A-T) was used as competitor; Panel B, heparin was used as competitor. Lane 1, 789 DNA alone; lane 2, 789 DNA plus CAP, incubate for 15 min; lane 3, 789 DNA plus RNA polymerase, incubate for 30 min; lanes 4 through 9, 789 DNA plus CAP for 15 min, then add polymerase and start timing; at indicated times mix aliquots of sample with appropriate competitor and begin electrophoresis immediately. The apparent difference in mobility of a complex is due to the different times of loading on the gel.


T (min) 15 30 1 5 10 30 60 120
RP − − + + + + + +
CAP − + − + + + + +
A
Ternary P1
CII—
P2—
CI—
free—
1 2 3 4 5 6 7 8 9


T (min) 15 30 1 5 10 30 60 120
RP − − + + + + + +
CAP − + − + + + + +
B
Ternary P1
CAP — free P1
CII—
P2—
CI—
free—
1 2 3 4 5 6 7 8 9

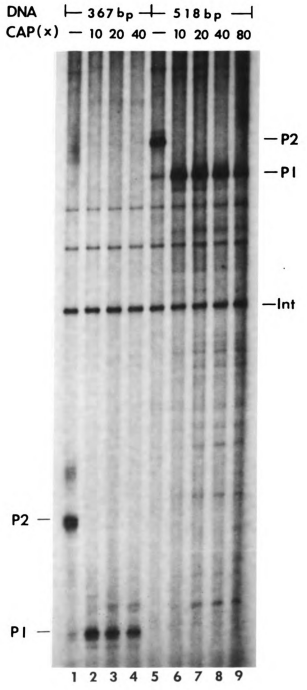
d(A-T) was used as competitor, almost no CAP-free P1 complex is detected at any time after RNA polymerase addition (Figure 34, Panel A). However, when the binding reaction is competed with heparin, both CAP-free and ternary P1 complexes are found (Figure 34, Panel B).

DISCUSSION

The catabolite activator protein stimulates initiation of transcription when bound at its primary binding site (-61) at the *E. coli lac* promoter. The lactose operon has overlapping RNA polymerase binding sites (P1 and P2) which are mutually exclusive (8, 84, 86). The effect of CAP on transcription from *lac* P1 has two aspects: first, CAP enhances the binding of RNA polymerase at P1 and represses its binding at P2 (the *lac* P2 promoter initiates *in vitro* 22 bp upstream of P1, thus placing its -35 region in the middle of the CAP-binding domain); second, CAP facilitates isomerization of closed complexes to open complexes (without CAP, RNA polymerase binds at *lac* P1 to form closed complexes *in vitro*, but melting-in is very slow and few P1 transcripts can be detected (79, 100; and see Figure 35, lanes 1 and 5).

In this thesis are presented results from a variety of approaches to study the protein-DNA and protein-protein interactions involved in open complex formation and transcription initiation at the *lac* promoter. To elucidate the interaction of CAP with RNA polymerase and with *lac* DNA

Figure 35. Runoff transcription of lac 367 and 518 DNA fragments as a function of CAP input. Lanes 1 and 5, transcripts from DNA plus polymerase interaction (no CAP); lanes 2 through 4 and 6 through 9, transcripts from complexes made by incubating DNA with increasing amounts of CAP as indicated, then adding polymerase. The "Int" band represents a radioactive DNA fragment added to the reaction as an internal control to monitor equal loading of transcription samples onto the gel. P1 and P2 indicate transcripts from P1 and P2 complexes at each fragment.


DNA
367 bp
518 bp
CAP(x)
10 20 40
10 20 40 80
P2
P1
Int
P2
P1
1 2 3 4 5 6 7 8 9

and to try to understand the function of the proposed loop structure in lac transcription initiation, gel retardation, in vitro transcription, DNase I footprinting, and exonuclease III protection assays were performed using complexes formed at lac promoter fragments.

Gel mobility shift assays of lac fragments in the presence of CAP and RNA polymerase, when challenged with heparin, reveal the existence of two major types of heparin-resistant complexes. Both of these have been identified as active P1 complexes; one (the ternary complex) contains promoter DNA plus CAP and polymerase, while the other contains only DNA and RNA polymerase (the CAP-free complex). What is the origin of these different P1 complexes?

Comparing the Exo III protection pattern using different competitors, it was found that a fraction of the ternary complexes that are stable in the presence of poly d(A-T) lose their CAP molecules and are converted to CAP-free P1 complexes in the presence of the stronger competitor heparin. This was further conformed by gel retardation assays (shown in Figure 34). Thus some lac complexes contain CAP molecules that are specifically bound at the -61 site, but are somehow not stabilized by RNA polymerase, since they are susceptible to heparin. In solutions containing lac DNA fragments, CAP, and RNA polymerase, there is a mixture of complexes, which differ in CAP-polymerase interactions. The existence of the CAP-free P1 open complex

leads to two conclusions about lac open complex formation and transcription initiation. First, it suggests that RNA polymerase binding and melting-in at the lac P1 promoter may not require extensive CAP-polymerase interaction; the mere presence of CAP at its binding site (-61) may be sufficient to help polymerase to bind and melt-in at lac P1, perhaps aided by the CAP-induced DNA bending. Second, it indicates that once RNA polymerase forms an open complex, it can initiate transcription on its own---CAP bound at the -61 site is not essential for initiation.

The formation of CAP-free P1 complexes appears to be related to heterogeneity in the population of CAP molecules. This conclusion is based on the properties of solutions containing different amounts of CAP. When more CAP is present, the CAP-free P1 complex band disappears and additional ternary P1 complexes are formed (Figures 20, 21, 22). Moreover, if this hypothesis is true, then incubating isolated CAP-free P1 complexes with more CAP protein might transform them to ternary P1 complexes. This was proved to be the case by the following experiment (data not shown). Gel electrophoresis was used to isolate lac 789 CAP-free P1 complexes. The relevant bands were incubated with CAP (spread onto the surface of the excised gel slices) at 37°C for 90 min. Then either poly d(A-T) or heparin were spread on the gel. After incubation for 5 min. the gel slices were loaded onto a second nondenaturing gel and electrophoresed

until the xylene cyanol marker dye ran off the gel. This experiment showed that when the binding was competed with poly d(A-T), all the CAP-free complexes were converted to CAP-containing ternary complexes. Moreover, even when heparin was used as competitor about half the original CAP-free complexes shifted to ternary complexes, while the rest remained CAP-free. We conclude that the CAP protein stock used in these experiments probably is not homogeneously active in establishing molecular contact with RNA polymerase. We know that some CAP molecules are completely inactive--they do not bind to the promoter CAP site. It is clear from the above data that some "active" CAP molecules are capable of binding to the CAP site at -61 and of enhancing RNA polymerase binding at P1, but are in some way deficient in contacts with RNA polymerase. Other CAP molecules can bind to the promoter and interact with polymerase to form heparin-resistant ternary P1 complexes.

Assuming that there are several forms of CAP present in a normal preparation, we can only speculate as to the molecular differences between these forms. As described in Chapter I, there are two potential surface-exposed activator loops in CAP. CAP molecules with mutations in these loops can repress P2 expression when bound at the lac CAP site, but are incapable of positive activation. It was proposed that these two loops might be the sites that interact with RNA polymerase in alternative situations, to activate

transcription initiation (68, 88). Residues in these two loops that affect CAP activation include His159, Gly162, Thr158 and Lys52.

Cell metabolism is regulated in a variety of ways. The rates of chemical reactions are controlled by regulation of enzyme activity, which can be achieved at the level of enzyme concentration (protein synthesis and degradation), or of enzyme activity (via effector molecules or by post-translational modification of the relevant proteins). Phosphorylation of proteins is used to modulate regulatory protein activity in both prokaryotes and eukaryotes (114, 115). In eukaryotic cells, cAMP is involved in phosphorylation events via cAMP-dependent protein kinases (116). So far, there is no published report about post-translational modification of CAP. It is interesting to speculate that such modifications do occur, that they affect CAP-polymerase interactions, and that at a given time the cell contains a mixture of modified and nonmodified molecules that is reflected in the purified protein used in these studies.

While our CAP preparations are no less active (10 to 40% active) than CAP prepared in other labs (117, 118) [except when CAP is purified by an affinity chromatography technique (reported to be 70% active) (100)], there remain uncertainties about the "inactive" molecules. Activity seems to be a function of the technique used to assay it;

for example, CAP that is 10% active in specific binding in a gel shift experiment seems to be fully active when studied by fluorescence (M. Sinton, personal communication).

A very interesting finding in these studies is the apparent role of a second CAP molecule (bound around +15) in enhancing initiation of transcription at the lactose P1 promoter, provided that sufficient lac DNA sequences are present for formation of a specific DNA loop. A single CAP molecule binding at the -61 site in the lac 203 ternary P1 complex represses lac P2 expression and may stimulate transcription initiation to some extent, but only in the presence of CAP molecules at both the +15 and -61 sites (which occurs with longer DNA fragments) do the ternary complexes show maximum initiation efficiency relative to their CAP-free counterparts (Figure 30).

The CAP stoichiometry study shows that the CAP:DNA ratio is 1.0 for CAP/ternary complexes containing the lac 203 fragment, and 2.0 for CAP/ternary complexes at lac 367 and 789 DNA. If the two CAP molecules at the lac 367 and 789 DNA promoters bind independently, we should find complexes that have CAP at only the -61 or the +15 site, but, so far, no such complexes have been detected. Thus, it is reasonable to assume that the binding of these two CAP molecules to a longer lac fragment (e.g. the 789) is somewhat cooperative. Since a single CAP can stably bind at the -61 site on the 203 DNA, it is likely that the formation

of ternary complexes at longer fragments involves one CAP molecule binding at the -61 site first, then the second CAP interacting at the +15 site in the presence of RNA polymerase and the upstream -250 site sequence. This is probably accomplished via loop formation bridged by the second CAP molecule between the +15 and -250 sites. Perhaps stable binding of the first CAP molecule at -61 affects the longer lac fragment DNA structure in such a way (bending?) that it is energetically more favorable for the second CAP molecule to bind at +15, assisted by the interaction with -250 site and subsequent loop formation. The loop can form either in the presence or absence of RNA polymerase.

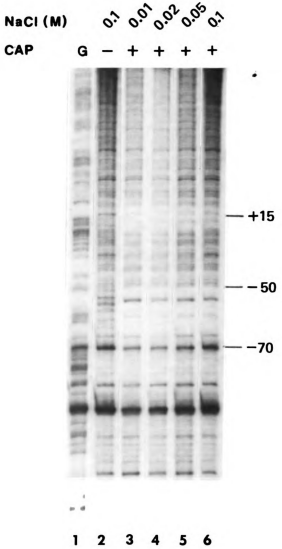
At an early stage in studying CAP interactions with the lac 789 DNA fragment, it was observed that at high levels of CAP a large fraction of DNA was found in slower moving CII complexes. Furthermore, only under conditions where complex CII could be detected by electrophoresis assays are there significant amounts of ternary complex formed when RNA polymerase is added. It thus seemed reasonable to assume that the CII and CI complexes of CAP plus lac 789 would be precursors of ternary and CAP-free P1 complexes, respectively. However, there were some inconsistences found when different lac DNA fragments were compared, as well as when different assays were used, indicating that above assumption is not correct.

One such inconsistency appeared in the following

experiment. An aliquot of a reaction mixture containing CAP plus lac 789 DNA was electrophoresed to separate CII and CI complexes, which were then probed, in the gel, by DNase I footprinting. A second aliquot was footprinted with DNase I in solution without gel purification. Different results were obtained. The gel-purified CII complex shows uniform protection from -70 to +20 (Figure 11, lane 2). Since there are as many as 15 CAP molecules per DNA fragment in each 789 CII complex, it is reasonable that the polymerase binding site is occupied by the bound CAP, although the precise structure of the CII complexes is not known. The in-solution footprinting experiment at 100 mM salt (Figure 8, and Figure 36, lane 6) shows the same protection pattern as that of the purified CI complex (Figure 11, lane 3), at -61 and around +15, even with 80X excess CAP present. These data suggest that the CII complex only forms after the reaction sample is loaded onto the retardation gel, in low ionic strength buffer (1X TBE), and most probably before the sample migrates into the gel.

Combine this with the results of Chapter II, it seems evident that the CII complex formation is determined by three factors, which include the length of the lac DNA fragment, cAMP and CAP concentrations and the ionic strength of the reaction buffer. Since CII can only be detected by gel shift assay at low salt but not by DNase I footprinting at 100 mM salt, it is possible that the CII complexes are

Figure 36. DNase I footprinting assay of lac 789 DNA interaction with CAP as a function of salt concentration. Lane 1, Maxam-Gilbert G sequencing reaction for 789 DNA; lane 2, 789 DNA alone; lanes 3 through 6, 789 DNA plus CAP (80X). Salt concentrations are as indicated for each reaction.


NaCl (M)
0.1
0.01
0.02
0.05
0.1
CAP
G
–
+
+
+
+
+15
–50
–70
1
2
3
4
5
6

nonspecific complexes of gel artifact. Chang et al. (125) detected the existence of a nonspecific complex of CAP with double-stranded DNA with a structure of regularly striated fibres under the electron microscope observation, the complex are four times shorter than the free DNA and has high protein content. Salemme (126) proposed that nonspecific interaction of CAP with DNA containing two CAP binding sites may leads to the formation of localized solenoidal coil, which in turn may affect RNA polymerase binding to the DNA. It is possible that the CII complex we saw may be this type of a complex but is not definitive. CII complexes are apparently not related to the formation of P1 open complexes, this is further proved by the finding that both 203 and 367 lac DNA fragments can form two different P1 complexes, yet neither shows a CAP-DNA complex that is equivalent to the lac 789 CII complex.

When different lac DNA fragments are transcribed in vitro at relatively low CAP input, transcription efficiency increases with CAP concentration to a certain extent; however, further addition of CAP actually reduces the amount of transcripts produced (Figure 35). This is in agreement with the data of Maquat (119). Correlating these results with the observation that excess CAP is required for lac 789 DNA-CII complex formation, it is reasonable to predict that the 789 DNA-CII complex is a dead complex. The 789 DNA-CII complex does not contribute to the formation of active open

complexes, nor does it convert back to a CI complex when extra CAP molecules are dissociated at higher salt (where the CI complex is stable) (Figure 14, lane 8).

If CII complex formation is responsible for repression of transcription at high CAP levels from the lac 789 DNA fragment, what about the shorter lac DNA fragments 203 and 367? Their in vitro transcription efficiency is also decreased by excess CAP input, but they do not form any CAP-DNA complexes equivalent to the 789 DNA-CII complex. The results in Chapter II imply that CII complex formation depends on both the length of the DNA fragment and the presence of a specific CAP recognition site. Therefore, it is possible that with the shorter lac DNA fragments (203 and 367), similar CAP-DNA and CAP-CAP interactions occur in the lac promoter region. This might involve clustering of CAP molecules at the several 5′-TGTGA-3′-like sequences that lie between the two CAP sites at -61 and +15. The cluster of CAP molecules might inhibit transcription initiation by blocking polymerase binding just as is presumed to occur at the lac 789 promoter region, except that these two fragments are not long enough to form the stable, looped CII complex, at least not as shown by retardation gel electrophoresis.

The notion is that CAP not only can bind to the lac promoter at the -61 CAP site, but it can also interact with lac promoter DNA in a more nonspecific fashion at additional sites immediately downstream. These interactions may lead

to formation of a "super complex", that prevents polymerase binding and therefore reduces transcription initiation from this promoter fragment. This conclusion is somewhat different from Reznikoff's suggestion that the inhibition of transcription by excess CAP is due to CAP binding at the +15 CAP site (8). Recall that we find that the second CAP, at +15, in a 789 ternary complex causes enhanced transcription relative to the CAP-free complex.

A gel retardation experiment using 789 DNA with CAP and RNA polymerase with different competitors (Figure 34) confirmed that both 789 P1 complexes are derived from the faster moving CAP complex, CI. Thus we conclude that in order to have RNA polymerase primarily bind and melt in at the lac P1 promoter, a CAP bound at -61 is necessary and sufficient, although its ultimate presence at this site is not necessary for polymerase to initiate transcription from +1. RNA polymerase melted in at the P1 site is in close proximity to CAP bound at -61, thus the proteins may interact with each other with consequent stabilization of the CAP molecule. As discussed above, for unknown reasons, some CAP molecules in our preparations have an altered ability to interact with RNA polymerase at the lac promoter, which leads to the formation of the CAP-free P1 complex in the presence of heparin.

The fact that the CAP-free P1 complex formed on 203 lac DNA initiates transcription about as well as the 367

fragment CAP-free P1 complex indicates that the sequences upstream of -140, specifically the -250 site, have no direct effect on RNA polymerase, but are important in recruiting the second CAP molecule to stably bind to the lac DNA fragment at the +15 site. This second CAP molecule binding provides enhanced stimulation of transcription initiation from open complexes. A CAP bound at -61 helps polymerase bind to the P1 site by preventing it from binding at the P2 promoter; this CAP may further help the polymerase to melt in, perhaps by physically bending the promoter DNA. Without the upstream -250 site, the second CAP molecule does not stably bind to the lac DNA fragments at +15 (at 100 mM salt), and the increased CAP stimulation of transcription initiation by RNA polymerase at lac P1 is not seen.

Our data imply that the sequences downstream of +60 make the complexes formed at 789 DNA better at transcription than those formed on 367 DNA fragments, but that some enhanced CAP activation of transcription occurs even in their absence (see Figure 30). Thus these sequences seem to have no fundamental function in CAP activation, but do help when they are present, perhaps by stabilizing open complexes via an unknown complex structure, or by somehow keeping the complexes in a more active conformational state.

Sequences downstream of the -10 region, including the CAP site at +15, allow the second CAP to be involved in activation of transcription initiation at the lacP1

promoter. Lorimer et al. (83) and Xiong et al. (84) reported that these downstream sequences are not essential for lac promoter function, but they used the lac 203 fragment, and not the lac 367 or 789 fragments (containing the -250 site) for assays. For the same reason, Hudson and Fried (118) could not demonstrate any cooperativity of CAP binding at the two CAP sites when using the lac 203 promoter fragment. Because the lac 203 promoter fragment does not bind the second CAP at +15 under the usual reaction conditions (100 mM salt), due to the absence of the -250 site sequence, deletion of sequences downstream of -6 has no impact on CAP activation of transcription.

It has been shown that a mutation which makes the +15 site a better CAP binding site represses lac expression (120) suggesting that if the second CAP is to function in activating transcription initiation it should bind "loosely" at this site.

The CAP molecules bound at -61 and +15 obviously interact with RNA polymerase differently, contacting polymerase at different surface domains. The CAP molecule at the +15 site may help polymerase escape from the promoter, while the CAP at -61 may stimulate polymerase binding and melting in at P1. It is not clear whether the activation of transcription initiation in the presence of the second CAP at +15 is caused by its interaction with polymerase and promoter DNA, or whether the -61 site bound

CAP is also involved. We do know that without CAP at the -61 site the second CAP does not stay on at the +15 site.

Measurement of association rates for complex formation at the lac 789 fragment with CAP and RNA polymerase shows that formation of the CAP-free P1 open complex is very fast when DNA is preincubated with CAP. The ternary P1 complex forms relatively slowly. Figure 19 shows that a substantial amount of CAP-free P1 complex forms only 30 sec after polymerase is added (lane 3), while up to 9 min is needed to form an equivalent amount of ternary complexes (lane 8). On the other hand, formation of CAP-free complexes at the lac 367 fragment takes minutes rather than seconds (Figure 18). For all three lac fragments, the establishment of interactions between CAP and polymerase appears to be a slow process (Figures 17, 18, and 19). These findings lead us to infer the following: (1) the interaction between CAP and RNA polymerase is not essential for polymerase initiating transcription from lacP1 promoter, as long as there is a CAP bound at -61 site to exclude binding at P2 and enhance binding and melting in at P1; (2) the establishment of the interaction between CAP and RNA polymerase is a rate limiting step for lac ternary complex formation, which implies that significant readjustment of conformation for both proteins is occurring; (3) the sequences downstream of +60 probably play a role in stimulating open complex formation (e.g., at lac 789 P1), which accounts for the

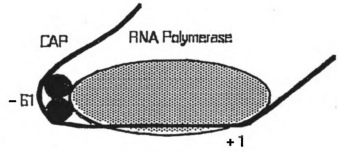
difference in the rate of CAP-free P1 complex formation at 789 and 367 DNA fragments (see Figures 18 and 19).

CAP effects on polymerase binding and melting in at promoter regions have been reported by other researchers. Kinetic experiments with gal DNA fragments show that CAP binding at the promoter region can affect either the formation of closed complexes or the isomerization to open complexes, depending on the nature of the sequence upstream of the -35 region (121). Lorimer (79) suggested that CAP bound at -61 stimulates the isomerization step at the lac P1 promoter, at least on lac 203 DNA. Moreover, data presented in this chapter suggest that, in addition to facilitating polymerase binding and melting in at lac P1, CAP also directly stimulates initiation of transcription from +1 when a second CAP molecule is bound at +15 (with the help of the -250 CAP site-facilitated loop structure).

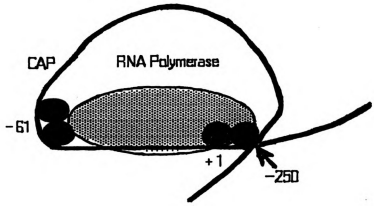
In conclusion, CAP activation of transcription initiation at the lac P1 promoter has two important features: (1) CAP binding at -61 prevents polymerase binding at P2, and ensures that polymerase exclusively binds at P1; (2) the natural lac sequence at -250 facilitates binding of a second CAP at +15, provided that a CAP molecule is bound at -61---this additional CAP has a direct role in stimulating transcription initiation from P1 open complexes. Figure 37 depicts the hypothetical structures of ternary P1 complexes at different lac DNA fragments. There is no

Figure 37. Proposed structures of lac ternary P1 complexes. A. Ternary complex of lac promoter DNA with one CAP dimer and RNA polymerase; this lac DNA fragment does not contain the -250 CAP recognition sequence. B. Ternary complex of lac promoter DNA containing RNA polymerase plus two CAP molecules; the CAP molecule at +15 interacts with the -250 sequence, and facilitates loop formation.

A.



B.



second CAP bound at +15 without the upstream -250 sequence, as in the case of *lac* 203 DNA (Figure 37, A). In the presence of the -250 region, the *lac* fragment binds two CAP molecules plus RNA polymerase; the CAP molecule at the +15 site also interacts with the upstream -250 sequence which leads to the formation of a loop structure (Figure 37, B).

CHAPTER IV

SUMMARY

SUMMARY

A variety of techniques have been used to study the interaction of catabolite activator protein (CAP) and RNA polymerase with lactose promoter DNA fragments. During the course of this work, a number of interesting observations were made:

(1) Longer *lac* DNA fragments containing the promoter region have higher affinity for CAP binding then shorter ones.

(2) Three sites of CAP interaction are found on *lac* DNA fragments that extend upstream to -310; they are centered at -250, -61, and +15. *lac* DNA fragments that do not contain the -250 region do not bind CAP at +15 at physiological salt concentration.

(3) *lac* DNA fragments containing the -61 and +15 sequences, with or without the -250 region, can form a high-CAP-content complex (CII) provided that they are about 500 bp or longer. The CII complex, however, does not contribute to open complex formation; on the contrary, it seems to prevent RNA polymerase binding at the promoter. CII-type interaction of CAP with *lac* promoter DNA leads to repression by CAP of *in vitro* transcription.

(4) *lac* DNA fragments containing all three CAP sites,

when mixed with CAP, bind two CAP molecules in a structure that likely involves a DNA loop. Open P1 transcription complexes formed at such fragments also contain two CAP molecules, in addition to RNA polymerase. However, lac 203 DNA, which is lacking the -250 site, stably binds only one CAP, at -61, even in the presence of RNA polymerase.

(5) CAP binding at -61 and +15, in the presence of the upstream -250 site, is strongly cooperative. lac DNA fragments with all three CAP sites bind either two CAP molecules or no CAP at all.

(6) Extensive CAP-RNA polymerase interaction at the lac promoter may not be required for polymerase melting in at P1. Once RNA polymerase melts in at the P1 promoter, the upstream (-61)-bound CAP is not essential for polymerase to initiate transcription from +1.

(7) Heparin-resistant ternary P1 open complexes containing two CAP molecules show elevated transcription initiation efficiency compared to the corresponding CAP-free P1 open complexes. But a heparin-resistant ternary P1 open complex containing only one CAP molecule (at lac 203 DNA) is no better at initiation than the related CAP-free open complex.

These results indicate the complexity of CAP and RNA polymerase interactions with promoter DNA sequences. The data in Chapter II indicate that there are three sequences that may interact with CAP. CAP binding at -61 is

independent of the other two sites, -61 being a very strong CAP site. CAP binding at the +15 site is dependent on the presence of -250 site, under our usual buffer conditions (100 mM salt); in agreement with other reports, CAP can bind at +15 at low salt in the absence of the -250 site. The necessity of the -250 sequence for CAP binding at +15 under more physiological conditions, and the fact that DNA fragments containing these three sites have two CAPs bound, strongly suggests the possibility of a loop formation between the -250 and +15 sites, bridged by a CAP dimer. CAP can also interact with the lac promoter in yet another mode, involving partially nonspecific CAP-DNA binding and extensive protein-protein interactions between CAP molecules in the lac promoter region, which in turn, may block the entry of RNA polymerase at the lactose promoter. This conclusion is derived from the observation that excess CAP represses in vitro transcription of the lac P1 promoter.

Chapter III is aimed at understanding the role of the CAP molecule at the +15 site in transcription initiation and in the mechanism of CAP activation of transcription. A CAP-free P1 open complex was discovered when the binding reaction was competed with heparin, which probably results from the failure of CAP to establish interaction with RNA polymerase bound at the P1 site. This suggests that all CAP molecules in our solution are not the same. Of those that can successfully bind to the lac CAP site at -61, some are

capable of interacting with RNA polymerase while others are not. Those that cannot establish the proper interaction with polymerase are removed from the promoter by heparin, while at the same time, the second CAP (if there is one) at the +15 site stabilized by looping has a reduced ability to bind the promoter, hence leading to the CAP-free P1 complex. Binding of a second CAP, at +15, depends on the presence of the first CAP molecule at -61. The purified CAP-free P1 complex is biologically active in initiating transcription, indicating that presence of CAP is not essential for RNA polymerase to begin mRNA synthesis.

The presence of a heparin-resistant CAP molecule at the +15 site (which occurs if RNA polymerase and the natural -250 sequence are present), leads to elevated transcription initiation efficiency relative to the CAP-free P1 complex at the same *lac* fragment. However, a heparin-resistant ternary complex containing only one CAP molecule at the -61 site (*lac* 203 DNA) does not show enhanced transcription initiation efficiency relative to the CAP-free P1 complex of the same DNA. This suggests that, in addition to helping polymerase to bind at and melt into the *lac* P1 promoter, CAP binds at +15, directly interacts with RNA polymerase, and stimulates transcription initiation from +1. Further studies are needed to elucidate exactly how the second CAP interacts with polymerase and stimulates the transcription initiation process.

APPENDIX

APPENDIX

The nucleotide sequence of the 789 lac DNA fragment used in this dissertation.

```
-310
 .
 AACCACCATC AAACAGGATT TTCGCCTGCT GGGGCAAACC AGCGTGGACC
 TTGGTGGTAG TTTGTCCTAA AAGCGGACGA CCCCGTTTGG TCGCACCTGG

-260
 .      [ -250  ]
 GCTTGCTGCA ACTCTCTCAG GGCCAGGCGG TGAAGGGCAA TCAGCTGTTG
 CGAACGACGT TGAGAGAGTC CCGGTCCGCC ACTTCCCGTT AGTCGACAAC

-210
 .
 CCCGTCTCAC TGGTGAAAAG AAAAACCACC CTGGCGCCCA ATACGCAAAC
 GGGCAGAGTG ACCACTTTTC TTTTTGGTGG GACCGCGGGT TATGCGTTTG

-160
 .
 CGCCTCTCCC CGCGCGTTGG CCGATTCATT AATGCAGCTG GCACGACAGG
 GCGGAGAGGG GCGCGCAACC GGCTAAGTAA TTACGTCGAC CGTGCTGTCC

-110
 .                                              [   CAP
 TTTCCCGACT GGAAAGCGGG CAGTGAGCGC AACGCAATTA ATGTGAGTTA
 AAAGGGCTGA CCTTTCGCCC GTCACTCGCG TTGCGTTAAT TACACTCAAT

-60
 .
  site 1]                  [ -35  ]       P1          [
 GCTCACTCAT TAGGCACCCC AGGCTTTACA CTTTATGCTT CCGGCTCGTA
 CGAGTGAGTA ATCCGTGGGG TCCGAAATGT GAAATACGAA GGCCGAGCAT
 [ -35  ]        P2          [  -10  ]

-10
 .
  -10]        [        CAP site 2        ]
 TGTTGTGTGG AATTGTGAGC GGATAACAAT TTCACACAGG AAACAGCTAT
 ACAACACACC TTAACACTCG CCTATTGTTA AAGTGTGTCC TTTGTCGATA
```

```
+41
 .
 GACCATGATT ACGGATTCAC TGGCCGTCGT TTTACAACGT CGTGACTGGG
 CTGGTACTAA TGCCTAAGTG ACCGGCAGCA AAATGTTGCA GCACTGACCC

+91
 .
 AAAACCCTGG CGTTACCCAA CTTAATCGCC TTGCAGCACA TCCCCCTTTC
 TTTTGGGACC GCAATGGGTT GAATTAGCGG AACGTCGTGT AGGGGGAAAG

+141
 .
 GCCAGCTGGC GTAATAGCGA AGAGGCCCGC ACCGATCGCC CTTCCCAACA
 CGGTCGACCG CATTATCGCT TCTCCGGGCG TGGCTAGCGG GAAGGGTTGT

+191
 .
 GTTGCGCAGC CTGAATGGCG AATGGCGCTT TGCCTGGTTT CCGGCACCAG
 CAACGCGTCG GACTTACCGC TTACCGCGAA ACGGACCAAA GGCCGTGGTC

+241
 .
 AAGCGGTGCC GGAAAGCTGG CTGGAGTGCG ATCTTCCTGA GGCCGATACT
 TTCGCCACGG CCTTTCGACC GACCTCACGC TAGAAGGACT CCGGCTATGA

+291
 .
 GTCGTCGTCC CCTCAAACTG GCAGATGCAC GGTTACGATG CGCCCATCTA
 CAGCAGCAGG GGAGTTTGAC CGTCTACGTG CCAATGCTAC GCGGGTAGAT

+341
 .
 CACCAACGTA ACCTATCCCA TTACGGTCAA TCCGCCGTTT GTTCCCACGG
 GTGGTTGCAT TGGATAGGGT AATGCCAGTT AGGCGGCAAA CAAGGGTGCC

+391
 .
 AGAATCCGAC GGGTTGTTAC TCGCTCACAT TTAATGTTGA TGAAAGCTGG
 TCTTAGGCTG CCCAACAATG AGCGAGTGTA AATTACAACT ACTTTCGACC

+441
 .
 CTACAGGAAG GCCAGACGCG AATTATTTTT GATGGCGTT
 GATGTCCTTC CGGTCTGCGC TTAATAAAAA CTACCGCAA
                                          .
                                        +479
```

LIST OF REFERENCES

REFERENCES

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

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

3. Zubay, G., Schwartz, D. and Beckwith, J. (1970) Proc. Natl. Acad. Sci. USA 66, 104-110.

4. Magasanik, B. (1970) in The Lactose Operon, Beckwith, J. R. and Zipser, D. eds., pp. 189-219. Cold Spring Harbor, New York. Cold Spring Harbor Labs.

5. 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.

6. Botsford, G. L. and Harman, J. G. (1992) Microbiol. Rev. 56, 100-122.

7. Scholte, B. J. and Postma, P. W. (1980) J. Bacteriol. 141, 751-757.

8. Reznikoff, W. S. (1992) Molecul. Microbiol. 6, 2419-2422.

9. Musso, R. E., Di Lauro, R., Adhya, S. and de Crombrugghe, B. (1977) Cell 12, 847-854.

10. Perez-Martin, J. and Espinosa, M. (1993) Science 260, 805-807.

11. Shanblatt, S. H. and Revzin, A. (1986) Biochemistry 25, 5539-5546.

12. Su, W., Porter, S., Kustu, S. and Echols, H. (1990) Proc. Natl. Acad. Sci. USA 87, 5504-5508.

13. Lee, D.-H. and Schleif, R. F. (1989) Proc. Natl. Acad. Sci. USA 86, 476-480.

14. Matthews, K. S. (1992) Microbiol. Rev. 56, 123-136.

15. Irani, M. H., Orosz, L., and Adhya, S. (1983) Cell 32, 783-788.

16. Majumdar, A. and Adhya, S. (1984) Proc. Natl. Acad. Sci. USA 81, 6100-6104.

17. Haber, R. and Adhya, S. (1988) Proc. Natl. Acad. Sci. USA 85, 9683-9687.

18. Schmitz, A. (1981) Nucleic Acids Res. 9, 277-292.

19. Burgess, R. R., Travers, A. A., Dunn, J. J., Bartz, E. K. F. (1969) Nature 221, 43-46.

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

21. Burgess, R. R., Erickson, B., Gentry, D., Gribskov, M., Hager, D., Lesley, S., Strickland, M., 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. R. eds., pp. 3-15, Elsevier, New York.

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

23. McClure, W. R. (1985) Ann. Rev. Biochem. 54, 171-204.

24. Schickor, P., Metzger, W., Werel, W., Lederer, H. and Heumann, H. (1990) EMBO J. 9, 2215-2220.

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

26. Carpousis, A. J. and Gralla, J. D. (1985) J. Mol. Biol. 183, 165-177.

27. Munson, L. M. and Reznikoff, W. S. (1981) Biochem. 20, 2081-2085.

28. Straney, D. C. and Crothers, D. M. (1985) Cell 43, 449-459.

29. Bowser, C. A. and Hanna, M. M. (1991) J. Mol. Biol. 220, 227-239.

30. Lee, D. N. and Landick, R. (1992) J. Mol. Biol. 228, 759-777.

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

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

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

34. Kornberg, A. (1974) DNA Synthesis, pp. 329-47, Freeman, W. H. and Co., San Francisco, CA.

35. Lee, J., Kashlev, M., Borukhov, S. and Goldfarb, A. (1991) Proc. Natl. Acad. Sci. USA 88, 6018-6022.

36. Martin, E., Sagitov, V., Burova, E., Nikiforov, V. and Goldfarb, A. (1992) J. Biol. Chem. 267, 20175-20180.

37. Zillig, W., Palm, P. and Heil, A. (1976) in RNA Polymerase, Losick, R. and Chamberlin, M. eds., pp. 101-125, Cold Spring Harbor Labs, Cold Spring Harbor, N.Y.

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

39. Venezia, N. D. and Krakow, J. S. (1990) J. Biol. Chem. 265, 8122-8126.

40. Igarashi, K. and Ishihama, A. (1991) Cell 65, 1015-1022.

41. Gross, C. A., Lonetto, M. and Losick, R. (1992) Bacterial Sigma Factors, in Transcriptional Regulation, Yamamoto, K. and McKnight, S. eds., Cold Spring Harbor, New York: Cold Spring Harbor Labs.

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

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

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

45. Lesley, S. A. and Burgess, R. R. (1989) Biochemistry 28, 7728-7734.

46. Dombroski, A. J., Walter, W. A., Record, M. T., Jr., Siegele, D. A. and Gross, C. A. (1992) Cell 70, 501-512.

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

48. McKay, D., Weber, I. and Steitz, T. (1982) J. Biol. Chem. 257, 9518-9524.

49. Weber, I. T. and Steitz, T. A. (1987) J. Mol. Biol. 198, 311-326.

50. Heyduk, T. and Lee, J. C. (1990) Proc. Natl. Acad. Sci. USA 87, 1744-1748.

51. Kolb, A., Busby, S., Buc, H., Garges, S. and Adhya, S. (1993) Ann. Rev. Biochem. 62, in press.

52. Brown, A. M. and Crothers, D. M. (1989) Proc. Natl. Acad. Sci. USA 86, 7387-7391.

53. Ebright, R. H., Ccssart, P., Gicquel-Sanzey, B., Beckwith, J. (1984) Nature 311, 232-235.

54. Gunasekera, A., Ebright, Y. W. and Ebright, R. H. (1992) J. Biol. Chem. 267, 14713-14720.

55. Zhang, X.-P. and Ebright, R. H. (1990) J. Biol. Chem. 265, 12400-12403.

56. Ebright, R. H., Ebright, Y. W. and Gunasekera, A. (1989) Nucleic Acids Res. 17, 10295-10305.

57. Schultz, S. C., Shields, G. C. and Steitz, T. A. (1991) Science 253, 1001-1007.

58. Collado-Vides, J., Magasanik, B. and Gralla, J. D. (1991) Microbiol. Rev. 55, 371-394.

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

60. Busby, S., Aiba, H. and de Crombrugghe, B. (1982) J. Mol. Biol. 154, 211-227.

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

62. Lee, N. L., Gielow, W. O. and Wallace, R. G. (1981) Proc. Natl. Acad. Sci. USA 78, 752-756.

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

64. Shanblatt, S. H. and Revzin, A. (1987) J. Biol. Chem. 262, 11422-11427.

65. Vidal-Ingigliardi, D. and Raibaud, O. (1991) Proc. Natl. Acad. Sci. USA 88, 229-233.

66. Pedersen, H., Sφgaard-Andersen, L., Holst, B., Gerlach, P., Bremer, E. and Valentin-Hansen, P. (1992) J. Mol. Biol. 227, 396-406.

67. Igarashi, K., Hanamura, A., Makino, K., Aiba, H., Mizuno, T., Nakata, A., and Ishihama, A. (1991) Proc. Natl. Acad. Sci. USA 88, 8958-8962.

68. Bell, A., Gaston, K., Williams, R., Chapman, K., Kolb, A., Buc, H., Minchin, S., Williams, J. and Busby, S. (1990) Nucleic Acids Res. 18, 7243-7250.

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

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

71. Kobayashi, M., Nagata, K. and Ishihama, A. (1990) Nucleic Acids Res. 18, 7367-7372.

72. Minchin, S. and Busby, S. (1993) Biochem. J. 289, 771-775.

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

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

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

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

77. Collis, C. M., Molloy, P. L., Both, G. W. and Drew, H. R. (1989) Nucleic Acids Res. 17, 9447-9468.

78. Lozinski, T., Adrych-Rozek, K., Markiewicz, W. T. and Wierzchowski, K. L. (1991) Nucleic Acids Res. 19, 2947-2953.

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

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

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

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

83. Lorimer, D. D., Cao, J. and Revzin, A. (1990) J. Mol. Biol. 216, 275-287.

84. Xiong, X., Cruz, N. and Reznikoff, W. S. (1991) J. Bacteriol. 173, 4570-4577.

85. Brunner, M. and Bujard, H. (1987) EMBO J. 6, 3139-3144.

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

87. Gilbert, W. (1976) in RNA Polymerase, PP. 192-205, Losick, R. and Chamberlin, M. J. eds., Cold Spring Harbor, New York: Cold Spring Harbor Labs.

88. Eschenlauer, A. C. and Reznikoff, W. S. (1991) J. Bacteriol. 173, 5024-5029.

89. Plaskon, R. R. and Wartell, R. M. (1987) Nucleic Acids Res. 15, 785-796.

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

91. Kolb, A., Spassky, A., Chapon, C., Blazy, B. and Buc, H. (1983) Nucleic Acids Res. 11, 7833-7852.

92. Zinkel, S. S. and Crothers, D. M. (1990) Biopolymers 29, 29-38.

93. Rees, W. A., Keller, R. W., Vesenka, J. P., Yang, G. and Bustamante, C. (1993) Science 260, 1646-1649.

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

95. Kim, J., Zwieb, C., Wu, C. and Adhya, S. (1989) Gene 85, 15-23.

96. Gartenberg, M. R. and Crothers, D. M. (1991) J. Mol. Biol. 219, 217-230.

97. Bracco, L., Kotlarz, D., Kolb, A., Diekmann, S. and Buc, H. (1989) EMBO J. 8, 4289-4296.

98. Dunn, T. M., Hahn, S., Ogden, S. and Schleif, R. F. (1984) Proc. Natl. Acad. Sci. USA 81, 5017-5020.

99. Ptashne, M. (1986) Nature 322, 697-701.

100. Straney, D. C., Straney, S. B. and Crothers, D. M. (1989) J. Mol. Biol. 206, 41-57.

101. Lavigne, M., Kolb, A. and Buc, H. (1992) Biochemistry 31, 9647-9656.

102. Joung, J. K., Le, L. U. and Hochschild, A. (1993) Proc. Natl. Acad. Sci. USA 90, 3083-3087.

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

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

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

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

107. Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular cloning: A laboratory manual. Second edition, Cold Spring harbor labs, Cold Spring harbor, New York.

108. Laemmli, U. K. (1970) Nature 227, 680-685.

109. Giulian, G. G., Moss, R. L. and Greaser, M. (1983) Analyt. Biochem. 129, 277-287.

110. Kolb, A. and Buc, H. (1982) Nucleic Acids Res. 10, 473-485.

111. Douc-Rasy, S., Kolb, A. and Prunell, A. (1989) Nucleic Acids Res. 17, 5173-5189.

112. Li, X.-M. and Krakow, J. S. (1985) J. Biol. Chem. 260, 4378-4383.

113. Ren, Y. L., Garges, S., Adhya, S. and Krakow, J. S. (1988) Proc. Natl. Acad. Sci. USA 85, 4138-4142.

114. Hunter, T. and Karin, M. (1992) Cell 70, 375-387.

115. Ninfa, A. J. and Magasanik, B. (1986) Proc. Natl. Acad. Sci. USA 83, 5909-5913.

116. Taylor, S. S., Buechler, J. A. and Yonemoto, W. (1990) Ann. Rev. Biochem. 59, 971-1005.

117. Hudson, J. M. and Fried, M. G. (1990) J. Mol. Biol. 214, 381-396.

118. Hudson, J. M. and Fried, M. G. (1991) J. Bacteriol. 173, 59-66.

119. Maquat, L. E. (1979) The effect of base-pair alterations on E. coli lactose promoter efficiency and nucleotide sequence analyses of lactose promoter mutants. Ph.D Thesis, University of Wisconsin, USA.

120. Irwin, N. and Ptashne, M. (1987) Proc. Natl. Acad. Sci. USA 84, 8315-8319.

121. Lavigne, M., Herbert, M., Kolb, A. and Buc, H. (1992) J. Mol. Biol. 224, 293-306.

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

123. Saxe, S. and Revzin, A. (1979) Biochemistry 18, 255-263.

124. Brierley, I. and Hoggett, J. G. (1992) Biochem. J. 285, 91-97.

125. Chang, J. J., Dubochet, J., Baudras, A., Blazy, B. and Takahashi, M. (1981) J. Mol. Biol. 150, 435-439.

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