P22 Arc Repressor:
Cooperative DNA Binding
and Transcriptional Regulation

by

Tracy Lynn Smith

B.A., Biology
Carleton College
June, 1990

Submitted to the Department of Biology
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

at the
Massachusetts Institute of Technology
January, 1996

Signature of Auth‹   Department of Biology

Certified by   Robert T. Sauer, Thesis Supervisor

Accepted by   Frank Solomon, Chairman, Biology Graduate Committee

P22 Arc Repressor: Cooperative DNA Binding and Transcriptional Regulation

by

Tracy Lynn Smith

Submitted to the Department of Biology on January 23, 1996
in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Abstract

The work presented here investigates Arc, a protein which binds to adjacent subsites in its operator and represses transcription from two divergent promoters, P_{ant} and P_{mnt}, in the *immunity I* operon of bacteriophage P22. Arc dimers bound to each subsite interact cooperatively to stabilize binding further. Mutagenesis studies of the protein, the operator, and the promoters regulated by Arc are presented here to address specific questions about cooperative DNA binding and transcriptional regulation by this protein. As an introduction, **Chapter 1** describes prokaryotic transcriptional initiation and focuses on the mechanisms of regulation by three well-characterized proteins. **Chapter 2** investigates the flexibility and the importance of Arc cooperative DNA binding. Mutation of Ser35 (a residue in the dimer-dimer interface) to Arg or Leu disrupts cooperative binding. The mutant proteins have near wild-type stabilities, give operator footprints like wild type, and prevent open-complex formation by RNA polymerase at the P_{ant} promoter *in vitro* but are largely inactive *in vivo*. Thus, although cooperativity is not structurally required for repression, it appears that the additional DNA binding energy from cooperativity is required for normal biological function. An analysis of Arc binding to operators in which the spacing between the DNA half-sites has been altered indicates that the cooperative contacts are inflexible. These experiments were published as "P22 Arc: Role of Cooperativity in Repression and Binding to Operators with Altered Half-Site Spacing" (Smith, T. L., & Sauer, R. T. (1995), *J. Mol. Biol.* 249, 729-742). **Chapter 3** analyzes regulation by Arc at two steps in transcription initiation, showing that an Arc dimer can slow open-complex formation and accelerate promoter clearance when bound to a single subsite. These dual activities of Arc allow it to act as either a repressor or an activator, depending on which step is rate-limiting in the presence of Arc. This work has been submitted for publication as "Dual Regulation of Distinct Steps in Transcription Explains a Novel Repressor to Activator Switch" (Smith, T. L., & Sauer, R. T. (1996), *submitted*). **Chapter 4** further investigates negative regulation of P_{ant} or P_{mnt} variants that contain only a single *arc* subsite. Occupancy of the subsite proximal to the P_{ant} -35 region results in more efficient repression than occupancy of the P_{ant} -10 proximal subsite. In P_{mnt}, Arc bound to the -10 proximal subsite is more effective than Arc bound to the -35 proximal subsite. Because of the divergent orientations of the two promoters, the -35 proximal site in P_{ant} is the same as the -10 proximal site in P_{mnt}. Thus occupancy of the same subsite results in the strongest repression of both promoters, suggesting that an Arc dimer at one position is primarily responsible for repression by the DNA-bound Arc tetramer. **Chapter 5** summarizes the results of Chapters 2-4 and suggests possible future experiments on the Arc system.

Thesis Supervisor: Robert T. Sauer, Whitehead Professor of Biochemistry

for my family,
especially Mom and Dad

.

Acknowledgements

I would like to thank the following people who participated in my graduate education (scientifically or otherwise!):

Bob Sauer, for being an excellent scientific advisor and for allowing me to make my own mistakes and learn from them.

All the members of the Sauer Lab, past and present, for advice and assistance over the years. Special thanks to Anne Paulin for her patient willingness to help in any situation.

Tania Baker, Alan Grossman, Peter Kim, and Ann Hochschild for serving on my thesis committee and for their interest and advice. Special thanks to Tania Baker, whose continual encouragement and daily good cheer have been much appreciated.

Don Rio, for letting me get started in his lab, Paul Kaufman for showing me how to deal with proteins, and the rest of the 1991 Rio Lab for teaching me the basics.

Susan Singer and Will McClure for allowing me to try out doing research in their respective labs and for encouraging me to go to graduate school.

My classmates. I could not have entered graduate school with a better group of people.

Eric Schmidt, for his friendship, love, and support, and for helping me realize that I really do like Brussels sprouts.

And most of all, my family, for supporting my education and for their constant encouragement of everything that I do.

Table of Contents

List of Tables

List of Figures

Chapter 1
Regulation of Prokaryotic Transcription Initiation

Regulation of Transcription

The regulation of gene expression at the level of transcription is a common feature of both prokaryotes and eukaryotes. At one level, much is known about transcriptional control: the outcome of many biological processes, such as a response to a particular nutrient, the development of an organism or the signaling between the differentiated cells of a tissue, is the activation or repression of specific genes. The regulators, the genes regulated, and the types of regulation (positive or negative) are known in many systems. What is often lacking is a detailed understanding of the mechanisms of activation or repression of particular genes.

The biochemical steps involved in both the eukaryotic and the prokaryotic transcription initiation cycles are presumably similar. However, significant differences do exist between transcription in prokaryotes and eukaryotes which may reflect the different complexities of the two systems. The different levels of complexity are illustrated by the differences in the RNA polymerase enzymes themselves. There are three eukaryotic RNA polymerases, Pol I which transcribes the rRNA genes, Pol II which transcribes messenger RNA, and Pol III which transcribes 5S RNA and tRNA. Each is a large multiprotein complex and requires several other accessory proteins to recognize promoter sequences and begin appropriate productive transcription (Geiduschek & Tocchini-Valentini, 1988; Young, 1991). In contrast, the single core RNA polymerase of prokaryotes is composed of only three types of subunits and requires in most cases only one additional subunit for promoter recognition and transcription initiation (Ishihama, 1988; von Hippel *et al.*, 1984).

In spite of these and many other differences, similar general mechanisms of transcriptonal regulation seem to apply to both systems. In both prokaryotes and eukaryotes, regulatory proteins bound at varying distances from the enzyme bound at the promoter can affect the basal transcription machinery through protein-protein contacts to increase or decrease the rate of transcription (Busby & Ebright, 1994; Gralla, 1989; Gralla, 1991; Guarente, 1988; North *et al.*, 1993; Tjian & Maniatis, 1994). In addition, alterations in the DNA structure around the promoter can affect transcription (Grosschedl *et al.*, 1994; Laurenson & Rine, 1992; Matthews, 1992; Perez-Martin *et al.*, 1994; Wolffe, 1994). However, in only a few cases, and most of those are prokaryotic, do we have even a partial understanding of the detailed mechanism by which a regulator affects the transcription apparatus (Adhya & Garges, 1990; Busby & Ebright, 1994). Therefore, an in-depth investigation of any specific system will add to our knowledge about how transcription is

regulated. Moreover, detailed models of regulation in prokaryotes may lead, by analogy or by direct comparison, to a better understanding of more complex regulatory circuits in eukaryotes.

In this introduction, I will review prokaryotic transcription initiation and regulation by three prokaryotic proteins. The remainder of the thesis will focus on understanding transcriptional control by the Arc protein of bacteriophage P22.

Transcription initiation in prokaryotes

The *E. coli* RNA polymerase holoenzyme consists of four types of subunits in the stoichiometry $\alpha_2\beta\beta'\sigma$, and each of these subunits has defined roles in transcription initiation at promoter sequences (Ishihama, 1988; von Hippel *et al.*, 1984). The most common promoters in *E. coli* contain two hexameric sequences termed the -35 and -10 regions for their distances from of the +1 transcriptional start site (Harley & Reynolds, 1987; McClure, 1985). Different types of promoters with varying consensus sequences exist. The σ subunit of the RNA polymerase holoenzyme confers promoter recognition to the core polymerase ($\alpha_2\beta\beta'$) (Chamberlin, 1974; Helmann & Chamberlin, 1988; Siegele *et al.*, 1988; Siegele *et al.*, 1989; Waldburger *et al.*, 1990), and there are many different σ factors to allow discrimination between the various promoters (Helmann & Chamberlin, 1988). In addition, the σ subunit is the target of certain transcriptional activators (Busby & Ebright, 1994; Hu & Gross, 1985; Hu & Gross, 1988; Kumar *et al.*, 1994; Popham *et al.*, 1989; Sasse-Swight & Gralla, 1990). The β and β' subunits form the catalytic center of the complex as shown by mutagenesis (Jin & Gross, 1991; Kashlev *et al.*, 1990) and substrate crosslinking studies (Borukhov *et al.*, 1991). The α subunit has multiple functions: it has an important role in assembling the core polymerase complex (Igarashi *et al.*, 1991; Ishihama, 1988; Kimura & Ishihama, 1995), it can bind specifically to A-T-rich sequences (UP elements) located upstream of certain promoters (Blatter *et al.*, 1994; Ross *et al.*, 1993), and it contacts many transcriptional regulators (Busby & Ebright, 1994; Russo & Silhavy, 1992).

Figure 1 shows a basic model for the interaction of RNA polymerase with a promoter. Upon recognizing the promoter, RNA polymerase forms the competitor-sensitive closed complex. The closed complex can dissociate or isomerize to form the more stable open complex, in which approximately 10 to 12 bp around the +1 start site of transcription are melted. The open complex is competent to initiate transcription upon

addition of ribonucleoside triphosphates (NTPs), forming an initiating complex. This stable ternary complex can produce short, abortive transcripts by recycling on the promoter. Loss of the σ subunit at the promoter clearance step commits the enzyme to leaving the promoter, and the transcript is elongated by core polymerase. Although this general model is well accepted (Chamberlin, 1974; McClure, 1985; von Hippel et al., 1984), specific intermediates have been postulated to occur before open complex formation at various promoters based on footprinting and kinetic evidence (Buc & McClure, 1985; Cowing et al., 1989; Roe et al., 1984; Schickor et al., 1990; Spassky et al., 1985; Straney & Crothers, 1987; Suh et al., 1992). In the most simple terms, initiation can be viewed as three steps: formation of the closed complex, isomerization to the open complex, and clearance from the promoter after initiation. Regulators of transcription initiation can act at any of the steps.

$$R + P \underset{}{\overset{K_B}{\rightleftharpoons}} RP_C \xrightarrow{k_f} RP_O \xrightarrow{NTPs} RP_{init} \xrightarrow{\sigma\ subunit} RP_{el}$$

Closed-Complex   Open-Complex   Promoter
Formation   Formation   Clearance
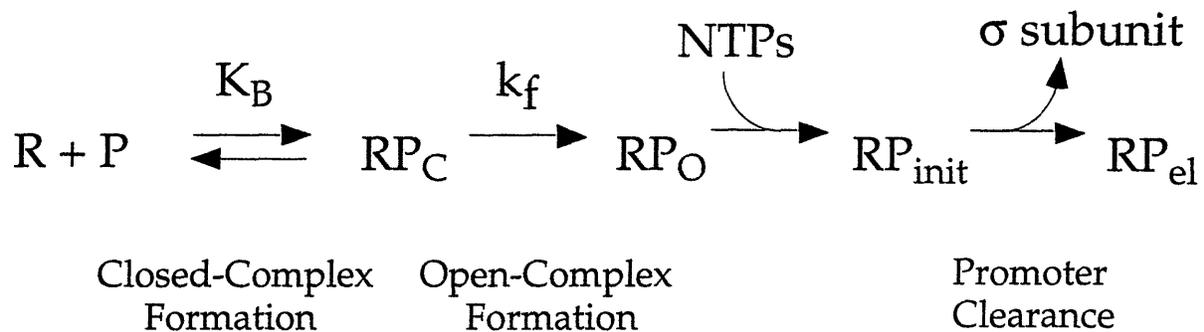
Figure 1. Model for prokaryotic transcription initiation. RNA Polymerase (R) recognizes a promoter-containing DNA fragment (P) and forms the closed complex (RP$_C$) which isomerizes to the open complex (RP$_O$). In the presence of nucleoside triphosphates (NTPs), transcription begins in the initiating complex (RP$_{init}$) and once a critical number of nucleotides are added, the enzyme leaves the promoter, loses the σ subunit, and forms the remainder of the transcript in the elongating complex (RP$_{el}$).

Regulation of prokaryotic transcription initiation

Prokaryotic regulators of transcription initiation are usually proteins that bind to specific DNA sites and exert their effects on the transcription apparatus either directly, through protein-protein contacts with RNA polymerase, or indirectly, by masking the RNA polymerase binding site or through alterations of the DNA structure. Several general

themes have emerged from a large body of work on various prokaryotic transcriptional regulators.

1) A multiprotein complex, composed of identical or nonidentical subunits, is often responsible for regulation.

2) Cooperativity, between the subunits in the multiprotein complex or between the regulator and RNA polymerase, is frequently observed.

3) Regulators can affect any step in transcription initiation, including stable complex formation with RNA polymerase, isomerization to the open complex, and clearance from the promoter.

4) Specific interactions with RNA polymerase are common and usually involve contacts with either the α or the σ subunit of RNA polymerase.

5) Alteration of the DNA structure induced by binding or a conformational change of a protein can play a major role in regulation of a promoter.

In this introduction, I will not attempt to cover all that is known about prokaryotic regulators of transcription initiation. Instead, I will focus on three specific examples of prokaryotic regulatory proteins. Although many other systems have been characterized, the three presented here, λ cI repressor, the cAMP receptor protein (CRP), and MerR each exhibit numerous aspects of the above themes of regulation, and yet each has made unique contributions to our understanding of gene regulation. In these systems, the binding sites for the protein are known, the steps at which the regulator affects transcription initiation are well characterized, the structure of the DNA can play a role, and interactions with RNA polymerase have been suggested based on mutational studies.

These three examples will serve as background for the experiments presented later in this thesis which are designed to investigate transcriptional control by P22 Arc repressor. Through these experiments, I address a number of the transcriptional themes that are manifest in the examples: Is cooperative DNA binding important for repression by Arc? Which steps in transcription initiation can Arc affect? Can only one of the two DNA binding subunits of the Arc multiprotein complex perform the task of repression? An understanding of the methods used to elucidate transcriptional regulation by λ repressor, CRP, and MerR sets the stage for understanding the rationale and the results of the experiments with Arc.

λ cI repressor

Overview

The λ cI repressor is a primary regulator of the lysis-lysogeny decision of bacteriophage λ and is one of the best characterized transcriptional regulators (Ptashne, 1986). The structures of the N-terminal DNA binding domain alone and in complex with operator DNA have been solved, facilitating analysis of its functions (Jordan & Pabo, 1988; Pabo & Lewis, 1982). λ repressor binds specifically to 17 base pair (bp) DNA sites, and six such sites are found in the λ genome, three in the right-operator region (O_R), and three in the left-operator region (O_L). The right operator sites overlap two divergent promoters, P_R and P_{RM} and regulation of these two promoters by λ repressor helps to decide between lysis and lysogeny. The discussion here will focus on O_R and regulation of P_R and P_{RM}.

Cooperativity

A major feature of regulation by λ repressor is that it binds cooperatively to DNA. Figure 2 shows a diagram of the O_R region, indicating the positions of the two promoters and the three binding sites for λ repressor and illustrating the cooperative DNA binding. One dimer of λ repressor binds to each 17 bp site, resulting in a cooperative DNA binding reaction that is second-order in monomer concentration (Ptashne, 1986). Furthermore, results of DNAse I footprinting experiments indicated that λ repressor demonstrates additional pairwise cooperativity (Johnson *et al.*, 1979). The binding of a repressor dimer to O_R1 increases the affinity of another repressor dimer bound at O_R2 so that both sites become occupied at a repressor concentration where O_R2 alone would not be bound. Similarly, λ repressor can bind cooperatively to O_R2 and O_R3 if O_R1 is mutated but cannot bind cooperatively to O_R1 and O_R3 if O_R2 is mutated (Johnson *et al.*, 1979).

λ repressor (236 amino acids per monomer) is composed of two domains, an amino (N) terminal domain and carboxy (C) terminal domain (Pabo *et al.*, 1979). The RecA protein of *E. coli* can mediate cleavage of λ repressor, separating the two domains (Roberts *et al.*, 1977; Sauer *et al.*, 1982). This cleavage prevents efficient function of λ repressor when conditions are appropriate for lytic growth. A proteolytic fragment of λ repressor that contains only the 92 amino acids of the N-terminal domain can dimerize, although more weakly, and bind to the operator sites (Pabo & Lewis, 1982; Pabo *et al.*,

15

1979). However, this fragment binds noncooperatively to each site, suggesting that protein-protein contacts mediated by the C-terminus of λ repressor are required for the pairwise cooperative DNA binding (Johnson et al., 1979).

$$P_R \rightarrow$$

-35   -10

O_R3 O_R2 O_R1

-10   -35

$\leftarrow P_{RM}$

C   C   C   C
N   N   N   N

O_R3 O_R2 O_R1
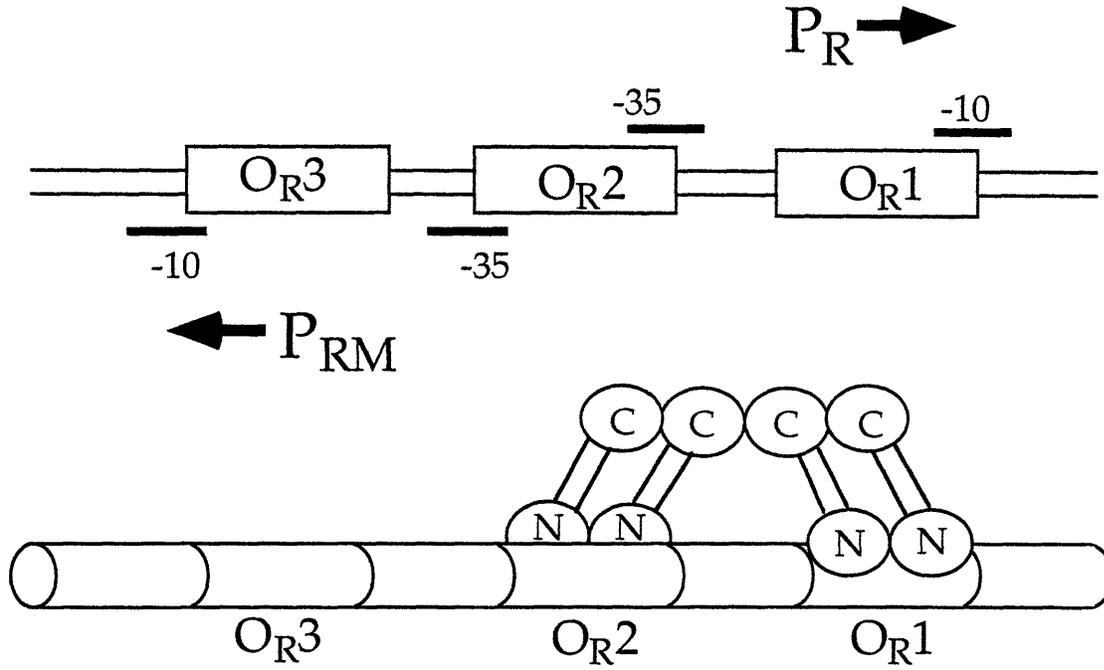
Figure 2. The top figure diagrams the positions of the three λ repressor dimer binding sites in the right operator region with respect to the P_R and P_{RM} promoters. The -35 and -10 promoter elements are indicated. The bottom figure illustrates pairwise cooperative binding of two λ repressor dimers bound to O_R1 and O_R2, a situation that would lead to repression of P_R and activation of P_{RM}. These figures were adapted from Johnson et al. (1979) and Ptashne (1986).

A wild-type λ repressor dimer will bind first to O_R1, the site with the highest affinity for λ repressor. A second dimer will then bind to O_R2, cooperatively stabilized by the dimer at O_R1. At higher repressor concentrations, the O_R3 site will fill. The sequential filling of the three binding sites has regulatory consequences for the P_R and P_{RM} promoters (Johnson et al., 1979; Maurer et al., 1980; Meyer et al., 1980; Meyer & Ptashne, 1980; Ptashne, 1986). When λ repressor occupies O_R1 and/or O_R2, transcription from P_R is repressed, preserving the lysogenic state by preventing transcription of genes that would switch the pathway toward lysis. Moreover, once the O_R2 site is occupied, transcription of the repressor gene from P_{RM} is activated. Occupancy of O_R3 at higher concentrations of λ repressor represses transcription from

P_{RM}, maintaining the concentration of λ repressor at a level that could be overcome by RecA-mediated cleavage of the repressor if lysis were eventually warranted.

The pairwise cooperative DNA binding of λ repressor is somewhat flexible. This is illustrated by the fact that repressor can bind cooperatively to sites separated by different numbers of bp (for example, the spacing between the O_R1 and the O_R2 sites is 7 bp, whereas the spacing between the O_R2 and O_R3 sites is 6 bp). However, it is not entirely flexible because dimers cannot bind cooperatively to O_R1 and O_R3 when O_R2 is mutated (Johnson et al., 1979). It was further shown that λ repressor dimers can bind cooperatively to sites separated by integral turns of the DNA helix, but not to sites on opposite sides of the helix. This type of cooperativity also required the presence of the C-terminal domain of λ repressor (Hochschild & Ptashne, 1986). Interestingly, constructs that position O_R1 5.9 turns of the DNA helix upstream of O_R2 result in an inability of λ repressor bound to O_R2 to stimulate transcription of P_{RM} whereas O_R1 positioned 5.5 turns upstream does not prevent a dimer bound at O_R2 from activating P_{RM} transcription (Hochschild & Ptashne, 1988). It is possible that the cooperative DNA binding between the two repressor dimers on the 5.9 turn construct results in improper positioning of the dimer at O_R2, preventing activation of P_{RM}. Alternatively, the distortion of the DNA that accompanies the cooperative binding of the widely separated sites could be responsible for the lack of P_{RM} activation.

Cooperative DNA binding of λ repressor has been further investigated by the isolation and characterization of mutants defective in pairwise cooperative DNA binding. Hochschild and Ptashne (1988) obtained one cooperativity-defective mutant and Whipple et al. (1994) isolated other mutants using a screen based on the fact that activation of P_{RM} is not obtained when repressor dimers bind to the 5.9 turn O_R1/O_R2 construct described above. Beckett et al. (1993) isolated cooperativity-defective mutants using a screen that requires cooperative DNA binding of λ repressor to two sites to obtain repression of a reporter gene. Benson et al. (1994) obtained a set of cooperativity-defective mutants using two phage superinfection selections, one selection that requires reduction of some aspect of binding to two sites (dimerization, intirinsic affinity, or cooperativity) followed by a second selection requiring only noncooperative DNA binding of λ repressor to one site for host survival. Each mutation obtained by these groups mapped to the C-terminal domain of λ repressor, confirming that specific protein-protein interactions in that domain mediate the cooperativity. The residues involved were G147, N148, S149, S159, E188, K192, R196, D197, S198, G199, F202, Y210, M212, S228, and T234. Only a few mutations also

reduced the ability of λ repressor monomers to dimerize (SN159, SN228, and TK234) (Whipple *et al.*, 1994) or were at position shown or predicted to affect the RecA-mediated cleavage of repressor (G147, S149, and K192) (Beckett *et al.*, 1993; Benson *et al.*, 1994; Gimble & Sauer, 1985; Gimble & Sauer, 1989; Whipple *et al.*, 1994). These results indicate that the cooperativity function can be separated from the two other major roles of the C-terminal domain, even though certain residues may be involved in more than one of those functions.

Whipple *et al.* (1994) further showed that the cooperative specificity of λ repressor can be switched to that of the homologous P22 c2 repressor by changing residues at six positions in λ repressor to the corresponding residues in c2 repressor (ND148, RM196, SA198, QR200, VK201, and QK204). They termed this hybrid repressor variant λv1-5; 148. λ repressor and c2 repressor have different DNA binding specificities, but their binding sites and the primary immunity regions of both bacteriophages are analogously organized. In addition, the λ and the P22 repressors have similar structures, an N-terminal domain that binds to DNA and a C-terminal domain that facilitates dimerization, RecA-mediated cleavage, and pairwise cooperativity. Furthermore, cooperative DNA binding of both λ repressor and P22 repressor is somewhat flexible since both can bind cooperatively to sites separated by several integral turns of the DNA helix (Hochschild & Ptashne, 1986; Valenzuela & Ptashne, 1989). Despite these similarities, the two proteins do not interact cooperatively with each other. The λv1-5; 148 variant can interact cooperatively with itself or with the P22 repressor, but not with λ repressor. This switch in specificity with just a small number of amino acid changes indicates that the structure of the C-terminal domain of the two repressors is quite similar, as predicted from sequence homology, and that a relatively small number of protein-protein interactions may be responsible for the specificity of the cooperative interactions. It would be interesting to compare these biochemical data with structural information of the C-terminal domain of λ repressor, which has so far eluded crystallographers.

One interesting question that has not been addressed is whether both C-terminal domains of each dimer are required for pairwise cooperativity, or whether one is more important than the other. An answer to this question might be obtained from an oriented heterodimer experiment (Zhou *et al.*, 1993a). In principle, altered DNA-binding specificity mutations paired in the same monomer with cooperativity defective mutations might appropriately position the cooperativity defective mutant monomer on only one half of a hybrid operator site. Appropriate constructs containing hybrid dimer binding sites could

then show if one position of the C-terminus is more important for cooperative binding of λ repressor and if the same types of interactions are used on adjacent sites and on sites separated by many turns of the helix.

Mechanisms of transcriptional regulation

The cooperative DNA binding of λ repressor is clearly an interesting and important aspect of its ability to regulate transcription. The hierarchy of intrinsic affinities for the three operator sites and the pairwise cooperativity result initially in the positioning of two dimers bound to O_R1 and O_R2. These two dimers repress transcription from P_R and the dimer bound at O_R2 activates transcription from P_{RM}. Much research has been conducted to elucidate the exact mechanism of transcriptional control of these two promoters by the pair of dimers bound cooperatively to O_R1 and O_R2.

Repression of P_R was investigated using an O_R2^- template to prevent activation of P_{RM}. Kinetic studies were performed *in vitro* to measure the rate of open-complex formation at various RNA polymerase concentrations in the presence or absence of λ repressor. Such experiments allow quantitation of the two main steps of open-complex formation, binding of RNA polymerase in the closed complex (K_B) and isomerizaton to the open complex (k_f). The presence of repressor specifically affected K_B, indicating that λ repressor bound at O_R1 primarily diminishes the ability of RNA polymerase to bind P_R in the closed complex (Hawley *et al.*, 1985). This competitive effect of λ repressor could result from bound repressor masking specific DNA bases that RNA polymerase must contact to form a specific complex, or it could result from other steric interference between the repressor and RNA polymerase.

Activation of P_{RM} involves specific contacts between λ repressor and a subunit of RNA polymerase. Meyer and Ptashne (1980) first demonstrated that activation of P_{RM} did not result simply from repression or removal of the divergent P_R promoter. In addition, activation of P_{RM} transcripton from an O_R1^-/O_R3^- template indicated that occupancy of O_R2 alone was sufficient and that pairwise cooperativity was not required structurally for activation of P_{RM} (Meyer *et al.*, 1980). Hawley and McClure (1982) showed that K_B for RNA polymerase at P_{RM} was essentially unaffected by the binding of repressor to O_R2 but k_f, the rate of isomerization to the open complex, was increased 11-fold (Hawley & McClure, 1982). Although these data are consistent with both direct and indirect effects of

19

λ repressor on RNA polymerase, many results eventually confirmed the direct-contact model.

The spatial relationship between O_R2 and P_{RM} is different than that between O_R2 and P_R, probably accounting for the differential regulation; O_R2 is one bp closer to P_R. In fact, if a bp is deleted between O_R2 and P_{RM}, λ repressor will no longer activate and will instead repress P_{RM} (Woody et al., 1993). The need for precise positioning of the O_R2 site for activation is consistent with the hypothesis that specific protein-protein contacts between λ repressor and RNA polymerase are responsible for activation. The isolation of mutants of λ repressor that bound O_R1 and O_R2 and repressed P_R but were unable to activate P_{RM} transcription strongly bolstered the direct-contact model (Guarente et al., 1982; Hochschild et al., 1983). Three positive control (pc) mutations were isolated (DN38, EK34, and GR43) and were located in the DNA binding domain on the solvent exposed surface of helix 2 and in the turn between helices 2 and 3 in the helix-turn-helix motif. One of the pc mutants, DN38, was studied in vitro and found to be incapable of increasing k_f (Hawley & McClure, 1983). The ability of the pc residues to activate transcription could be transferred to a heterologous helix-turn-helix in the Cro protein, which can also bind to O_R2 but which normally represses P_{RM} (Bushman & Ptashne, 1988). Furthermore, mutations were made at these same positions in the homologous 434 repressor, and they increase or decrease its ability to activate transcription of the 434 P_{RM} promoter, which is positioned analogously to the λ P_{RM} promoter (Bushman & Ptashne, 1988). The efficiency of activation by the hybrid 434 repressor depended on the acidic nature of certain residues. Similarly, in λ repressor, a glutamic acid residue at position 34 is the critical residue for activation (Bushman et al., 1989).

Two groups conducted genetic screens to determine which part of RNA polymerase is contacted by λ repressor. They focused on the α and the σ^{70} subunits of RNA polymerase because mutations in both had been isolated which result in an inability of RNA polymerase to respond to various other activators (Hu & Gross, 1985; Hu & Gross, 1988; Russo & Silhavy, 1992). Mutations in the σ^{70} subunit of RNA polymerase were isolated that cannot be strongly activated by λ repressor (Kuldell & Hochschild, 1994) and allele specific suppressors of one pc mutation were also found in the σ^{70} subunit of RNA polymerase (Li et al., 1994). The mutations lie in or near the putative helix-turn-helix DNA binding motif of σ^{70}. Presumably, only one of the two λ repressor monomers is properly positioned to contact RNA polymerase at these positions. The promoter proximal monomer is likely to be responsible for activation because of overlapping phosphate

20

contacts made by that monomer and RNA polymerase (Hochschild *et al.*, 1983), but this has not been tested directly. An understanding of how the interactions between λ repressor and the σ subunit of RNA polymerase affect isomerization will require a better structural understanding of RNA polymerase itself.

Conclusions

As described above, transcriptional regulation by λ repressor is well-characterized and involves many of the themes previously mentioned. Multiple dimers bound to DNA are required for the biological function of λ repressor, and the three dimer binding sites have different roles in regulation of the two promoters. Cooperative DNA binding of the dimers is essential but involves relatively few contacts between the C-terminal domains of the paired dimers. Moving the O_R1 and O_R2 dimer binding sites far apart can have negative regulatory consequences at P_{RM} even though cooperativity and binding to O_R2 are maintained, perhaps because of the distortion of the DNA that occurs around the promoter. λ repressor can act at discrete steps in transcription initiation, blocking closed-complex formation or stimulating isomerization, depending on the location of the binding site with respect to the promoter. Finally, specific contacts between λ repressor and RNA polymerase have been identified which are responsible for its ability to activate transcription.

cAMP Receptor Protein

Overview

The cAMP receptor protein (CRP; also known as the catabolite gene activator protein or CAP) is a versatile regulator of many genes in *E. coli* that respond to the level of cAMP, and thus to the level of glucose in the cell. CRP is a dimer that when bound to cAMP can bind specifically to a 22 bp operator. The structures of CRP-cAMP and CRP-cAMP in complex with its operator are available (McKay *et al.*, 1982; Schultz *et al.*, 1991). Each CRP monomer is composed of two domains, an N-terminal domain that binds cAMP, and a C-terminal domain that contains a helix-turn-helix DNA binding motif. Allosteric changes in CRP occur upon binding of one cAMP molecule per dimer, and these changes reposition the recognition helices leading to an enhancement of specific DNA binding (Kolb *et al.*, 1993). A single CRP binding site can be found at many positions with respect to the promoters that it regulates, with common positions centered at -41.5, -61.5, and -70.5. Upon binding to one of these operator sites, CRP bends the DNA dramatically by approximately 90° (Schultz *et al.*, 1991; Zinkel & Crothers, 1991). When bound to DNA, CRP can regulate transcription positively or negatively. These effects can be direct by affecting the activity of RNA polymerase or indirect by repositioning other transcription factors that subsequently exert their effects on the transcription machinery (Kolb *et al.*, 1993). This discussion will focus only on the direct positive effects of CRP on transcription at a variety of promoters.

Activation by CRP

An analysis of spacing requirements for CRP regulation suggested that the CRP site must always lie on the same face of the helix relative to the promoter for efficient activation, hinting at interactions between bound CRP and RNA polymerase (Gaston *et al.*, 1990). The finding that CRP and RNA polymerase bound cooperatively to the *lac* promoter also supported a direct interaction between the two protein complexes (Straney *et al.*, 1989). Positive control mutations, analogous to the pc mutants in λ repressor discussed above, mapped to an exposed loop of CRP at positions A156, M157, T158, H159, P160 G162, M163, and Q164 (Bell *et al.*, 1990; Eschenlauer *et al.*, 1991; Niu *et al.*, 1994; Zhou *et al.*, 1994a; Zhou *et al.*, 1993b). Contact between CRP and RNA polymerase was finally detected directly both on promoter DNA by crosslinking and in the absence of promoter

DNA by fluorescence polarization (Chen *et al.*, 1994; Heyduk *et al.*, 1993). Thus, direct contact between CRP and RNA polymerase may be critical for its role in regulating many of its target promoters.

Most of the promoters that are directly activated by CRP can be grouped into two major classes based mainly on the position of the CRP binding sites with respect to the promoter (see Figure 3) but also on which steps of transcription initiation are affected by CRP (Kolb *et al.*, 1993; Zhou *et al.*, 1994b). The class I sites are positioned upstream of the -35 RNA polymerase recognition element, usually centered around -61.5, -72.5, -82.5, or -92.5. The class II binding sites are positioned at -41.5 and overlap the -35 region. Although experimental results indicate that CRP contacts RNA polymerase at both class I and class II promoters, many lines of evidence suggest that the mechanism of CRP activation at these two types of promoters is basically different.
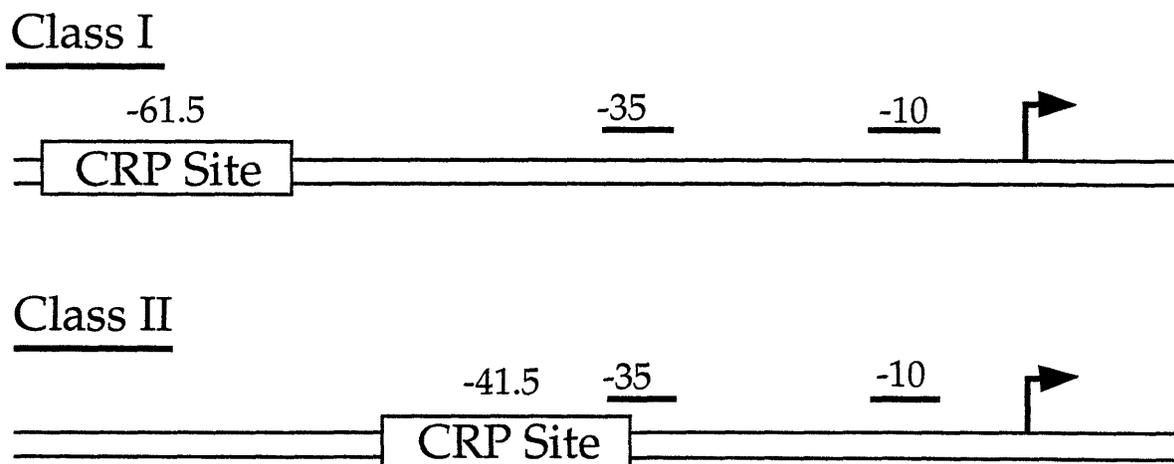
Class I

Class II

Figure 3. The two major classes of CRP-dependent promoters are indicated. Class I promoters have a CRP site positioned well upstream of the -35 region whereas class II promoters have a CRP site overlapping the -35 promoter element.

At class I promoters such as the *lac* promoter, where the binding site is positioned at -61.5, CRP has been shown to activate transcription primarily by stabilizing the binding of RNA polymerase in the closed complex, thereby increasing the K_B term (Gaston *et al.*, 1990; Malan *et al.*, 1984). Zhou *et al.* (1994) showed that alanine substitution of positive control residue T158 resulted in the largest reduction in activation of two class I promoters

23

with CRP sites centered at -61.5 and -72.5. Alanine substitution of G162 and P160 also reduced activation but with smaller effects. It was further shown that only the promoter proximal monomer of the CRP dimer needed to have wild-type residues at the activating loop for the dimer to function as an activator at class I promoters (Zhou et al., 1994a; Zhou et al., 1993b). Thus, exact postioning of the activation loop seems to be critical for proper interaction with RNA polymerase.

Both the α and the σ subunits of RNA polymerase may be involved in activation by CRP at class I promoters. Residues in the activation loop of CRP can be crosslinked to the C-terminal 100 amino acids of the α subunit of RNA polymerase (Chen et al., 1994). Furthermore, RNA polymerase holoenzymes containing C-terminal truncation mutants or specific point mutants of the α subunit of RNA polymerase cannot be activated by wild-type CRP at class I promoters (Chen et al., 1994; Igarashi & Ishihama, 1991; Russo & Silhavy, 1992). The C-terminal domain of the α subunit has recently been implicated in specific DNA binding of a sequence upstream of some promoters (Blatter et al., 1994; Ross et al., 1993). Therefore, the role of CRP at class I promoters may be to stabilize the α subunit on the DNA. Alternatively or additionally, the interactions with the α subunit may stabilize the binding of the σ subunit to the DNA. This is suggested because some of the same σ^{70} mutations that reduced activation by λ repressor (discussed above) also reduced CRP-activation of a class I promoter (Kuldell & Hochschild, 1994). Perhaps the contact of the α subunit by CRP indirectly affects the conformation of the σ subunit, resulting in better binding to the promoter DNA. The solution structure of the α C-terminal domain has recently been solved (Jeon et al., 1995), but a more complete mechanistic understanding of how CRP activates transcription by interacting with the α subunit will require more structural detail about the holoenzyme.

At class II promoters, such as the galP1 promoter, where the CRP site is centered at -41.5, CRP can activate transcription by affecting both the binding of RNA polymerase in the closed complex and isomerization to the open complex (Gaston et al., 1990; Goodrich & McClure, 1992). It was shown that at class II promoters, CRP uses the same activation loop of residues 156 to 164, but that it uses those residues in a different way than at class I promoters. Whereas residue T158 was most critical for activation at class I promoters, the following positions were most essential for activation of a class II promoter and are listed in order of importance; G162, P160, M163, and T158 (Zhou et al., 1994a). Furthermore, unlike at class I promoters, it is the promoter distal monomer of the CRP dimer that is responsible for activation of class II promoters (Zhou et al., 1994b). Because the

importance of the activation loop residues differ between the two classes of promoters, it may be that CRP does not use the same contacts with the α subunit and interacts with a different part of RNA polymerase. Consistent with this is the observation that the α subunit C-terminal truncation mutants that could not be activated by CRP at class I promoters can still respond to CRP at class II promoters (Igarashi & Ishihama, 1991). Recent experiments with truncation mutants of the σ^{70} subunit suggest that the CRP-contact point for class II promoters may be located in the C-terminus of σ^{70} near the helix-turn-helix motif (Kumar et al., 1994).

Although there are many more details concerning CRP activation at these two classes of promoters, a compelling feature is that when CRP binds in different positions with respect to the promoter, it can have different qualitative effects on transcription activation. Since CRP is predicted to contact different regions of the RNA polymerase when bound to the two sites, it is reasonable to ask whether two DNA bound CRP dimers can synergistically activate transcription of a promoter. Two groups showed that such synergy can occur, consistent with the prediction of different contact points between CRP and RNA polymerase (Busby et al., 1994; Joung et al., 1993). Joung et al. (1994) further showed that synergistic activation also occurred as a result of the binding of CRP to a class I site and λ repressor to a single binding site positioned analogously to O_R2 at position -42 (Joung et al., 1994). Multiple contact points between regulators and E. coli RNA polymerase is similar to the situation in eukaryotic transcription in which large sets of regulators interact with the basal transcription machinery (Tjian & Maniatis, 1994).

At some promoters, CRP seems to activate by different mechanisms. When CRP is bound at -70.5 upstream of the malT promoter, it has little affect on open-complex formation but stimulates clearance of RNA polymerase from the promoter by reducing abortive recycling (Menendez et al., 1987). At the uhpT promoter, where CRP binds at -103.5, activation was not reduced by mutations in the activator loop described above (Merkel et al., 1995), suggesting that activation occurs by a different mechanism which may or not be similar to the mechanism at the malT promoter. Intrinsically bent DNA can substitute for CRP sites at certain promoters both in vitro and in vivo (Bracco et al., 1989; Gartenberg & Crothers, 1991), and at the fur promoter, activation is obtained when the CRP site at -70 is replaced by a site for RepA, a heterologous DNA bending protein (Perez-Martin & Espinosa, 1993). Both types of results suggest a role primarily for DNA bending by CRP and not for CRP contact of RNA polymerase at those promoters. The different mechanisms of activation at the malT and uhpT promoters discussed above could also

potentially be explained if DNA bending activates transcription at those sites. Bending of the DNA could bring upstream DNA sequences into contact with RNA polymerase, or a bend could induce topological changes in the DNA around the promoter, thus affecting the activity of RNA polymerase indirectly (Perez-Martin et al., 1994). At some promoters, either DNA bending or protein-protein contacts between CRP and RNA polymerase might suffice for activation, whereas at other promoters, both activities of CRP might be required for maximum activation.

Conclusions

CRP is an extremely versatile transcription factor that uses both its ability to distort its binding site and its ability to contact RNA polymerase at multiple points to activate transcription. It can act at multiple steps in transcription initiation, affecting binding of RNA polymerase, isomerization, and promoter clearance. Specific interactions between CRP and RNA polymerase have been identified which are important for its ability to activate certain promoters. At promoters where contact between RNA polymerase and CRP is required, only one monomer of the CRP dimer is positioned properly to be directly involved in activation. Finally, one of the most striking features of activation by CRP is that the promoter architecture seems to be a defining element for determining the mechanism of activation. Although CRP has much broader regulatory activites than λ repressor, both exemplify many of the same transcriptional themes.

MerR

Overview

The *mer* operon of *Tn21* or *Tn501*, found in gram-negative bacteria, helps the host to detoxify the mercuric ion (Hg(II)) through expression of a mercuric reductase and mercuric transport proteins (Summers, 1992). The MerR protein, a member of this operon, is a 144 residue protein composed of two domains, one containing a helix-turn-helix DNA binding motif and one containing residues that bind to Hg(II). When bound to Hg(II), MerR (Hg-MerR) can activate transcription of the detoxification genes at the $P_{TP(C)AD}$ promoter. MerR also represses the transcription of its own gene from the P_R promoter in both the presence and absence of Hg(II), resulting in production of a constant

26

amount of MerR protein. These two promoters regulated by MerR are divergent (see Figure 4). The binding site for MerR is located between the -35 and -10 promoter elements of $P_{TP(C)AD}$ and around the +1 start site of transcription from P_R. A MerR dimer binds to this site both in the presence and absence of Hg(II) (Summers, 1992). Therefore, unlike CRP, specific DNA binding by MerR is not dependent on ligand binding. Instead, as will be discussed in detail below, the role of the inducer in activation is downstream of binding to the operator. The following discussion will focus only on activation and repression of the $P_{TP(C)AD}$ promoter by MerR.

$$P_{TP(C)AD}$$

$$P_R$$
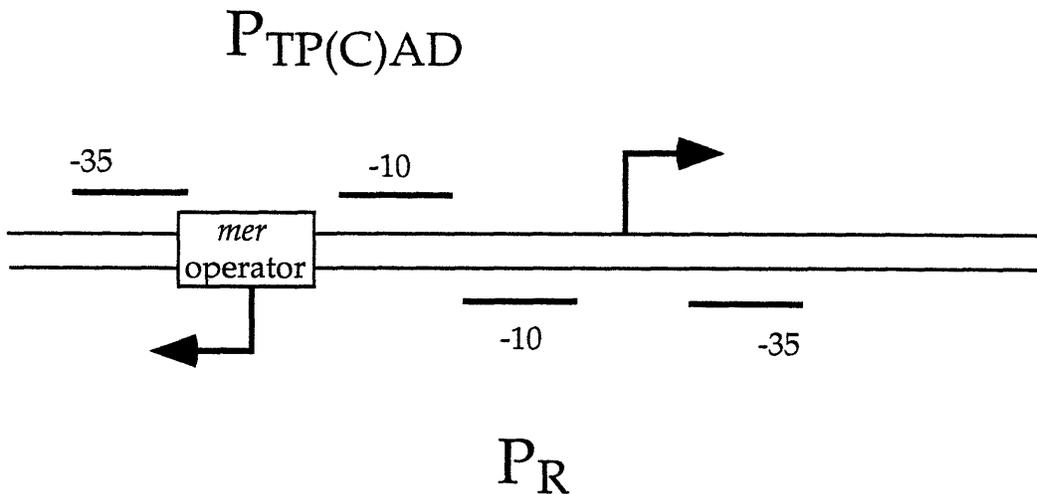
Figure 4. The mer Operon. Diagram shows the position of the *mer* operator with respect to the two promoters that it regulates. The discussion will concentrate on the top promoter $P_{TP(C)AD}$ that is repressed in the absence of Hg(II) and activated in the presence of Hg(II).

Activation of $P_{TP(C)AD}$ by MerR

The MerR operator is in an unusual position with respect to the $P_{TP(C)AD}$ promoter, located between the two RNA polymerase recognition elements (see Figure 4) (Collado-Vides *et al.*, 1991). MerR binds to this site both in the presence and absence of Hg(II) with approximately the same affinity (Heltzel *et al.*, 1990; O'Halloran *et al.*, 1989). In spite of the proximity of the operator to the RNA polymerase recognition elements, RNA polymerase can still bind to $P_{TP(C)AD}$ in the closed complex in the absence of Hg(II). Therefore, the $P_{TP(C)AD}$ promoter exists in a preinduction complex containing MerR and RNA polymerase that is poised to respond to the presence of Hg(II) (Heltzel *et al.*, 1990;

27

O'Halloran *et al.*, 1989). When Hg(II) binds to MerR, a conformational change occurs in the protein that results in alterations in the bending and the winding of the operator DNA, but does not otherwise alter the position of the protein on the DNA. Hg-MerR bends the DNA less than MerR, and the DNA in the Hg-MerR complex is underwound by about 33°. These distortions are evident from changes in footprint patterns, alterations in gel mobilities of the complexes, and from linking number changes (Ansari *et al.*, 1995; Ansari *et al.*, 1992; Frantz & O'Halloran, 1990; Heltzel *et al.*, 1990; O'Halloran *et al.*, 1989). The P$_{TP(C)AD}$ promoter has a nonoptimal 19 bp spacing between the -35 and -10 elements. The changes in the structure of the DNA induced by Hg-MerR would realign those elements, making a more favorable promoter geometry and leading to activation of the promoter (Ansari *et al.*, 1995; Ansari *et al.*, 1992; Frantz & O'Halloran, 1990; Heltzel *et al.*, 1990; O'Halloran *et al.*, 1989; Summers, 1992). In support of this hypothesis, changing the spacing between the P$_{TP(C)AD}$ recognition elements to the optimal 17 bp results in increased promoter activity and lack of dependence on Hg-MerR (Lund & Brown, 1989; Parkhill & Brown, 1990).

Mutation of any one of the residues involved in binding the Hg(II) ion (C82, C117, and C126) results in an activation-defective but not a repression-defective phenotype (Livrelli *et al.*, 1993). In addition, mutation of A60 can result in a specific activation-defective phenotype. The DNA distortion seen upon Hg(II) induction of wild-type MerR is not seen with AT60 and AV60 mutants, consistent with a role for DNA bending and unwinding of the operator in activation (Livrelli *et al.*, 1993; Ross *et al.*, 1989). These activation-defective mutants may be unable to adopt the conformation of MerR required to distort the DNA. In contrast, a constitutively activating MerR protein is obtained with two double mutants (SC86/AV89 or AV89/SL131). These mutants exhibit the DNA distortion in the operator even in the absence of Hg(II), providing yet more evidence supporting a direct role for the distortion in activation (Comess *et al.*, 1994; Parkhill *et al.*, 1993). These mutants may be locked in the conformation required to actively bend and unwind or stabilize the distorted DNA.

Activation by MerR could require specific contacts between MerR and RNA polymerase, although no mutants of the appropriate phenotype have been described. The expected phenotype would be proper DNA distortion by the protein in the presence of Hg(II) but lack of activation of P$_{TP(C)AD}$. It may be easier to screen for such mutants starting with one of the constitutively active double mutants discussed above. A

complementary approach would be to screen for mutations in subunits of RNA polymerase that render it incapable of being activated by MerR.

Activation by MerR apparently also requires certain bp in the operator DNA. Lee *et al.* (1993) showed that transversion of position -22 from G to T resulted in an activation-deficient phenotype but did not derepress the divergent P_R promoter, indicating that MerR was still bound to the operator. These results suggest that slightly different DNA contacts may be made in the uninduced and induced complexes and that the operator mutation destabilized a contact necessary either for inducing the DNA distortion or for binding to the distorted DNA (Lee *et al.*, 1993). Alternatively, the mutant operator may simply be unable to adopt conformation required for activation.

Repression of $P_{TP(C)AD}$ by MerR

One of the most interesting aspects of regulation by MerR is that it actively represses transcription from $P_{TP(C)AD}$ in the absence of Hg(II) by stably sequestering RNA polymerase in a closed complex that is ready for induction by Hg(II) (Heltzel *et al.*, 1990; O'Halloran *et al.*, 1989). MerR mutants that cannot bind to the DNA do not result in stable closed complex formation (Livrelli *et al.*, 1993). The cooperative DNA binding could result from protein-protein contacts between MerR and RNA polymerase. In principle, the same stabilizing contacts or different contacts between MerR and RNA polymerase could actively prevent isomerization to the open complex. Consistent with this hypothesis, several MerR mutants have been characterized which result in a repression-defective phenotype even though their DNA binding affinities are unaltered. Comess *et al.* (1994) observed a repression defective phenotype (*i.e.* a high expression of a reporter gene in the absence of Hg(II) but further induction upon addition of Hg(II)) when certain pairs of acidic residues in MerR were changed to either glutamine or asparagine. One pair of mutations, EQ83/EQ84, was sufficient to confer a mild repression-defective phenotype, but the effect of these mutations was amplified in the presence of other pairs of mutations (either DN68/EQ69, EQ77/DN78, or DN68/EQ69). The same group then targeted a small region (positions 81 to 92) for random oligonucleotide mutagenesis. The mutation SC87 was isolated as a repression-defective mutant from that library (Comess *et al.*, 1994).

The EQ77/DN78/EQ83/EQ84 MerR mutant was purified and shown to have the same affinity for operator DNA as wild-type MerR both in the presence and absence of

Hg(II), showing that the repression-defective phenotype in the absence of Hg(II) does not result from an inability to bind the DNA. However, the DNA binding properties of the SC87 mutant were not characterized. On the basis of the repression-defective phenotype and the wild-type binding affinity of the quadruple mutant, it was proposed that these mutants cannot properly contact RNA polymerase to prevent transcription (Comess et al., 1994). However, an alternative explanation is that these mutants can partially distort the DNA in the absence of Hg(II), but not as strongly or properly as the constitutively activated mutants described in the previous section. The ability of either mutant protein to distort the DNA in the absence of Hg(II) was not examined. Thus, these two possibilities would have to be distinguished before one can ascribe a protein-protein interaction function to the residues in question. A footprinting screen for MerR mutants that can bind DNA but cannot stably sequester RNA poymerase at the promoter would be difficult but may reveal other possible positions of contact between MerR and RNA polymerase. Screening for mutations in RNA polymerase that result in noncooperative binding with MerR would also be difficult but potentially worthwhile since the ability of MerR to stably sequester RNA polymerase in a closed preinduction complex is unique among regulators studied thus far.

Conclusions

The MerR protein regulates transcription of the $P_{TP(C)AD}$ promoter by novel mechanisms both in the presence and absence of Hg(II). Both repression and activation of $P_{TP(C)AD}$ by this protein are complex. This is the first system for which stable sequestering of RNA polymerase in the closed complex by an activator is seen, and the first for which a clear realignment of the promoter recognition elements by a DNA distortion is a component of activation. The binding of MerR between the recognition elements may position the protein to interact with multiple parts of RNA polymerase as well as to distort the DNA, and more work is needed to define possible contact points between MerR and RNA polymerase. In addition, more structural information about MerR is needed to fully understand the multiple functions of this versatile regulator.

Summary

The three regulators discussed above, λ repressor, CRP, and MerR share many of the common features of transcriptional regulation discussed initially. However, each has a

few particularly interesting features that stand out. Transcriptional regulation by λ repressor is strongly influenced by cooperative DNA binding to multiple sites, resulting in the precise positioning of repressor dimers for the repression and activation functions. Activation by CRP bound to disparately spaced sites is possible because CRP can interact with multiple regions of RNA polymerase and/or use its ability to distort DNA to regulate different steps in transcription initiation, including the promoter clearance step at which only a few other regulators are known to act (Lee & Goldfarb, 1991; Narayan *et al.*, 1994; Goodrich & Tjian, 1994). MerR can either repress or activate transcription of a promoter when bound to the same site, depending on a conformationally induced DNA distortion in the operator, which is located in the unusual position between the promoter recognition elements. Thus, although similar transcriptional themes run throughout the discussions, each system has provided particular contributions to our understanding of transcription initiation and regulation. As will be seen in this thesis, investigation of transcriptional regulation by Arc has also provided interesting variations on the commonly seen themes of regulation.

Arc

Focus of thesis

Arc repressor, the subject of this thesis, regulates transcription from two overlapping divergent promoters, P_{ant} and P_{mnt}, in the *immunity I* operon of bacteriophage P22 (Figure 5) (Susskind & Youderian, 1983). This operon consists of three genes, two of which, *antirepressor* and *arc*, are transcribed together from P_{ant}. The third gene, *mnt*, is transcribed from the P_{mnt} promoter. The main function of the *immunity I* operon is the production and regulation of the expression of Antirepressor protein. Antirepressor antagonizes the P22 c2 repressor (analogous to the λ repressor) during early lytic growth or superinfection. However, Antirepressor expression is not desired during lysogeny, when its expression would lead to lysis. Nor is Antirepressor expression warranted after commitment to the lytic pathway, when its production is a waste of resources and is in fact antagonistic to phage growth at high levels of expression. Thus, two regulators, Arc (for antirepressor control) and Mnt (for maintenance of lysogeny), exist to prevent expression of Antirepressor at different times during the life cycle of the phage. The role of the Mnt protein is to prevent transcription of P_{ant} during lysogeny and to activate its own transcription from P_{mnt} (Susskind & Youderian, 1983; Vershon *et al.*, 1987a). The role of

31

the Arc protein is to prevent transcription from both P_{ant} and P_{mnt} during late lytic growth (Susskind & Youderian, 1983; Vershon *et al.*, 1987b). In this thesis, I investigate Arc regulation of both P_{ant} and P_{mnt}.

$$P_{ant}$$

| mnt | | O_{arc} | O_{mnt} | arc | antirepressor |

$$P_{mnt}$$

$$P_{ant} \longrightarrow$$

-35   -10

TAGTGTATTGACAT GATAGA AGCAC TCTACT ATATTCTCAATAGG
ATCACATAACTGTA CTATCT TCGTG AGATGA TATAAGAGTTATCC

-10   -35

arc binding sites
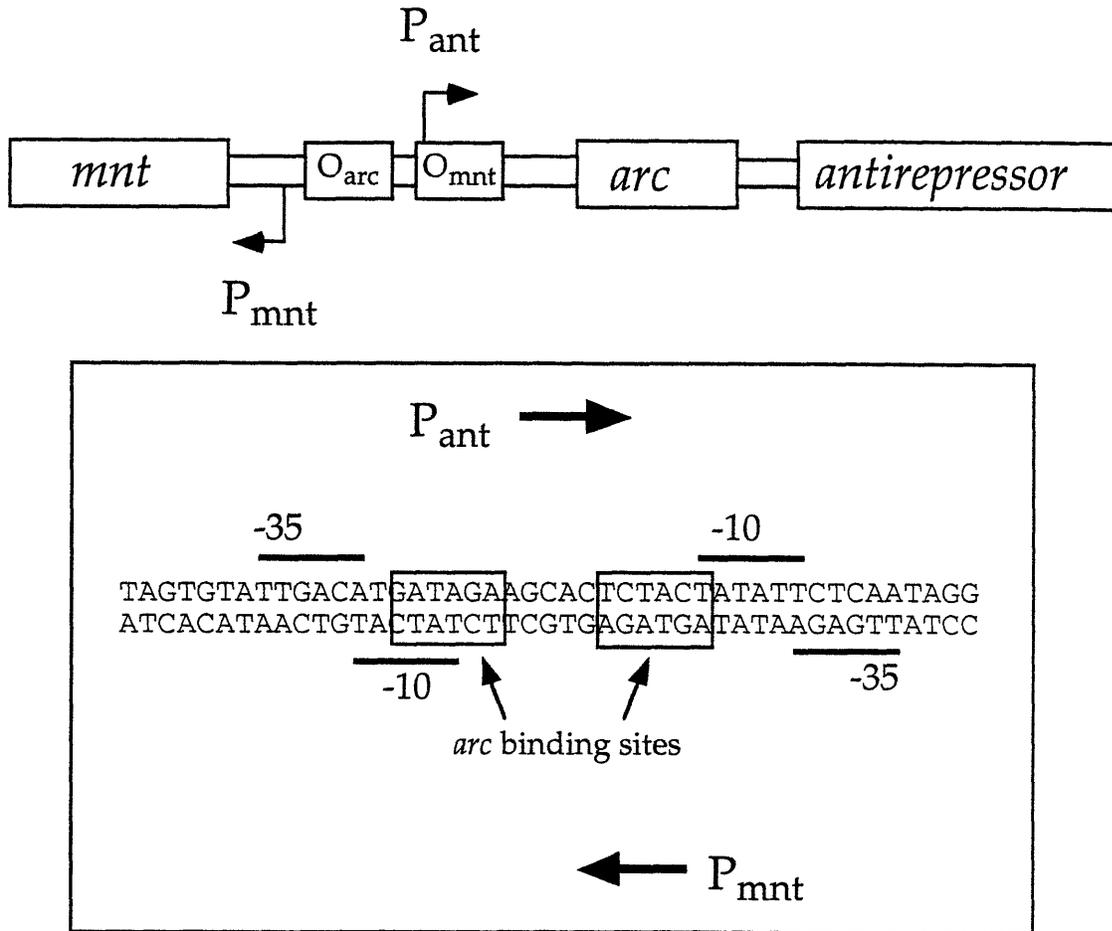
$$\longleftarrow P_{mnt}$$

Figure 5. The *immunity I* operon of bacteriophage P22. The top diagram illustrates the overall structure of the operon and the positions of the two promoters, P_{ant} and P_{mnt}, and the O_{arc} and O_{mnt} operators. Below is a close-up of the *arc* operator region showing the position of the *arc* operator with respect to the two promoters. The two *arc* binding sites are boxed.

Arc is a member of the ribbon-helix-helix family of regulatory proteins (Bowie & Sauer, 1990; Phillips, 1991; Raumann *et al.*, 1994a). It is a small protein (53 residues per monomer) that folds as a dimer, with the two monomers intertwining and forming a β-

sheet with one β-strand from each monomer (Bowie & Sauer, 1989a; Breg *et al.*, 1990; Vershon *et al.*, 1985). Arc uses this β-sheet to interact specifically with operator DNA (Raumann *et al.*, 1994b; Vershon *et al.*, 1986b). Two Arc dimers bind to the operator, one to each of the half-sites (Figure 5), and the two dimers interact cooperatively through protein-protein contacts to stabilize the DNA-bound tetramer (Brown *et al.*, 1990; Brown & Sauer, 1993; Raumann *et al.*, 1994b). In the *immunity I* operon, the *arc* operator is located in between and slightly overlapping the -35 and -10 elements of the two promoters that it regulates, in a position similar to that of the *mer* operator described above (Vershon *et al.*, 1987b).

The structures of the Arc dimer and the Arc tetramer-operator complex have been solved, facilitating a detailed understanding of this system (Breg *et al.*, 1990; Raumann *et al.*, 1994b). In addition, a wealth of biochemical information exists on the folding and DNA binding properties of Arc (Bowie & Sauer, 1989a; Bowie & Sauer, 1989b; Bowie & Sauer, 1989c; Brown *et al.*, 1990; Brown *et al.*, 1994; Brown & Sauer, 1993; Milla *et al.*, 1993; Milla *et al.*, 1994; Milla *et al.*, 1995; Milla & Sauer, 1994; Milla & Sauer, 1995; Robinson & Sauer, 1996; Schildbach *et al.*, 1995; Silva *et al.*, 1992; Smith & Sauer, 1995; Waldburger *et al.*, 1995). The studies presented in this thesis expand our knowledge by focusing on the cooperative DNA binding of Arc and the role of Arc in transcription initiation.

Chapter 2 asks whether cooperativity between the two dimers is required for repression of P_{ant}; mutants in both the protein and the operator DNA that cannot support cooperative binding of two Arc dimers are described. Chapter 3 demonstrates that Arc has multiple abilities in regulating transcription initiation; it can simultaneously slow open-complex formation and enhance promoter clearance when bound to the P_{ant} -35 proximal half-site. Chapter 4 asks if the two Arc dimers in the DNA bound tetramer are functionally equivalent; experiments are presented which test whether a single Arc dimer bound in either of the two half-site positions can repress transcription of P_{ant} and/or P_{mnt} promoter variants. Chapter 5 summarizes the main points of chapters 2-4, placing the information into context with the examples of regulators discussed in the introduction, and presents preliminary results for possible future research on cooperative DNA binding and transcriptional regulation by Arc.

References

Adhya, S. & Garges, S. (1990). Positive control. *J. Biol. Chem.*, **265**, 10797-10800.

Ansari, A.Z., Bradner, J.E. & O'Halloran, T.V. (1995). DNA-bend modulation in a repressor-to-activator switching mechanism. *Nature*, **374**, 371-375.

Ansari, A.Z., Chael, M.L. & O'Halloran, T.V. (1992). Allosteric underwinding of DNA is a critical step in positive control of transcription by Hg-MerR. *Nature*, **355**, 87-89.

Beckett, D., Burz, D.S., Ackers, G.K. & Sauer, R.T. (1993). Isolation of λ repressor mutants with defects in cooperative operator binding. *Biochemistry*, **32**, 9073-9079.

Bell, A., Gaston, K., Williams, R., Chapman, K., Kolb, A., Buc, H., Minchin, S., Williams, J. & Busby, S. (1990). Mutations that alter the ability of the *Escherichia coli* cyclic AMP receptor protein to activate transcription. *Nucleic Acids Res.*, **18**, 7243-7250.

Benson, N., Adams, C. & Youderian, P. (1994). Genetic selection for mutations that impair the co-operative bindng of lambda repressor. *Mol. Microbiol.*, **11**, 576-579.

Blatter, E.E., Ross, W., Tang, H., Gourse, R.L. & Ebright, R.H. (1994). Domain organization of RNA polymerase α subunit: C-terminal 85 amino acids constitute a domain capable of dimerization and DNA binding. *Cell*, **78**, 889-896.

Borukhov, S., Lee, J. & Goldfarb, A. (1991). Mapping of a contact for the RNA 3' terminus in the largest subunit of RNA polymerase. *J. Biol. Chem.*, **266**, 23932-23935.

Bowie, J.U. & Sauer, R.T. (1989a). Equilibrium dissociation and unfolding of the Arc repressor dimer. *Biochemistry*, **28**, 7139-7143.

Bowie, J.U. & Sauer, R.T. (1989b). Identification of C-terminal extensions that protect proteins from intracellular proteolysis. *J. Biol. Chem.*, **264**,

Bowie, J.U. & Sauer, R.T. (1989c). Identifying determinants of folding and activity for a protein of unknown structure. *Proc. Natl. Acad. Sci.*, **86**, 2152-2156.

Bowie, J.U. & Sauer, R.T. (1990). TraY proteins of F and related episomes are members of the Arc and Mnt repressor family. *J. Mol. Biol.*, **211**, 5-6.

Bracco, L., Kotlarz, D., Kolb., A., Diekmann, S. & Buc, H. (1989). Synthetic curved DNA sequences can act as transcriptional activators in *Escherichia coli. EMBO J.*, **8**, 4289-4296.

Breg, J.N., van Opheusden, H.J., Burgering, M.J.M., Boelens, R. & Kaptein, R. (1990). Structure of Arc repressor in solution: evidence for a family of β-sheet DNA-binding proteins. *Nature*, **346**, 586-589.

Brown, B.M., Bowie, J.U. & Sauer, R.T. (1990). Arc repressor is tetrameric when bound to operator DNa. *Biochemistry*, **29**, 11189-11195.

Brown, B.M., Milla, M.E., Smith, T.L. & Sauer, R.T. (1994). Scanning mutagenesis of the Arc repressor as a functional probe of operator recognition. *Nature Struct. Biol.*, **1**, 164-168.

Brown, B.M. & Sauer, R.T. (1993). Assembly of the Arc repressor-operator complex: cooperative interactions between DNA-bound dimers. *Biochemistry*, **32**, 1354-1363.

Buc, H. & McClure, W.R. (1985). Kinetics of open complex formation between Escherichia coli RNA polymerase and the *lac UV5* promoter. Evidence for a sequential mechanism involving three steps. *Biochemistry*, **24**, 2712-2723.

Busby, S. & Ebright, R.H. (1994). Promoter structure, promoter recognition, and transcription activation in prokaryotes. *Cell*, **79**, 743-746.

Busby, S., West, D., Lawes, M. & Webster, C. (1994). Transcription activation by the *Escherichia coli* cyclic AMP receptor protein: receptors bound in tandem at promoters can interact synergistically. *J. Mol. Biol.*, **241**, 341-352.

Bushman, F.D. & Ptashne, M. (1988). Turning λ cro into a transcriptional activator. *Cell*, **54**, 191-197.

Bushman, F.D., Shang, C. & Ptashne, M. (1989). A single glutamic acid residue plays a key role in the transcriptional activation function of lambda repressor. *Cell*, **58**, 1163-1171.

Chamberlin, M.J. (1974). The selectivity of transcription. *Annu. Rev. Biochem.*, **43**, 721-775.

Chen, Y., Ebright, Y.W. & Ebright, R.H. (1994). Identification of the target of a transcription actiator protein by protein-protein photocrosslinking. *Science*, **265**, 90-92.

Collado-Vides, J., Magasanik, B. & Gralla, J.D. (1991). Control site location and tanscriptional regulation in *Escherichia coli*. *Microbiol. Rev.*, **55**, 371-394.

Comess, K.M., Shewchuk, L.M., Ivanetich, K. & Walsh, C.T. (1994). Construction of a synthetic gene for the metalloregulatory protein MerR and analysis of regionally mutated proteins for transcriptional regulation. *Biochemistry*, **33**, 4175-4186.

Cowing, D.W., Mecsas, J., Record, M.T., Jr. & Gross, C.A. (1989). Intermediates in the formation of the open complex by RNA polymerase holoenzyme containing the sigma factor σ^{32} at the *groE* promoter.*J. Mol. Biol.*, **210**, 521-530.

Eschenlauer, A. C., & Reznikoff, W. S. (1991). *Escherichia coli* catabolite gene activator protein mutants defective in positive control of *lac* operon transcription. *J. Bacteriol.* **173**: 5024-5029.

Frantz, B. & O'Halloran, T.V. (1990). DNA distortion accompanies transcriptional activation by the metal-responsive gene-regulatory protein MerR. *Biochemistry*, **29**, 4747-4751.

Gartenberg, M.R. & Crothers, D.M. (1991). Synthetic DNA bending sequences increase the rate of *in vitro* transcription initiation at the *Escherichia coli lac* promoter. *J. Mol. Biol.*, **219**, 217-230.

Gaston, K., Bell, A., Kolb, A., Buc, H. & Busby, S. (1990). Stringent spacing requirements for transcription activation by CRP. *Cell*, **62**, 733-743.

Geiduschek, E.P. & Tocchini-Valentini, G.P. (1988). Transcripion by RNA polymerase III. *Annu. Rev. Biochem.*, **57**, 873-914.

Gimble, F.S. & Sauer, R.T. (1985). Mutations in bacteriophage λ repressor that prevent RecA-mediated cleavage. *J. Bacteriol.*, **162**, 147-154.

Gimble, F.S. & Sauer, R.T. (1989). λ repressor mutants that are better substrates for RecA-mediated cleavage. *J. Mol. Biol.*, **206**, 29-39.

Goodrich, J.A. & McClure, W.R. (1992). Regulation of open complex formaiton at the *Escherichia coli* galactose operon promoters: Simultaneous interaction of RNA polymerase, *gal* repressor and CAP/cAMP. *J. Mol. Biol.*, **224**, 15-29.

Goodrich, J. A., & Tjian, R. (1994). Transcription factors IIE and IIH and ATP hydrolysis direct promoter clearance by RNA polymerase II. *Cell*, **77**, 145-156.

Gralla, J.D. (1989). Bacterial gene regulation from distant DNA sites. *Cell*, **57**, 193-195.

Gralla, J.D. (1991). Transcriptional control--lessons from an *E. coli* promoter data base. *Cell*, **66**, 415-418.

Grosschedl, R., Giese, K. & Pagel, J. (1994). HMG domain proteins: architectural elements in the assembly of nucleoprotein structures. *TIG*, **10**, 94-98.

Guarente, L. (1988). UASs and Enhancers: common mechanism of transcriptional activation in yeast and mammals. *Cell*, **52**, 303-305.

Guarente, L., Nye, J.S., Hochschild, A. & Ptashne, M. (1982). Mutant λ phage repressor with a specific defect in its positive control function. *Proc. Natl. Acad. Sci.*, **79**, 2236-2239.

Harley, C.B. & Reynolds, R.P. (1987). Analysis of E. coli promoter sequencs. *Nucleic Acids Res.*, **15**, 2343-2361.

Hawley, D.K., Johnson, A.D. & McClure, W.R. (1985). Functional and physical characterization of transcription initiation complexes in the bacteriophage λ O_R region. *J. Biol. Chem.*, **260**, 8618-8626.

Hawley, D.K. & McClure, W.R. (1982). Mechanism of activation of transcription initiation from the λ P_{RM} promoter. *J. Mol. Biol.*, **157**, 493-525.

Hawley, D.K. & McClure, W.R. (1983). The effect of a lambda repressor mutation on the activation of transcription initiation form the lambda P_{RM} promoter. *Cell*, **32**, 327-333.

Helmann & Chamberlin (1988). Structure and function of bacterial sigma factors. *Annu. Rev. Biochem.*, **57**, 839-872.

Heltzel, A., Lee, I.W., Totis, P.A. & Summers, A.O. (1990). Activator-dependent preinduction binding of a σ-70 RNA polymerase at the metal regulated *mer* promoter. *Biochemistry*, **29**, 9572-9584.

Heyduk, T., Lee, J.C., Ebright, Y.W., Blatter, E.E., Zhou, Y. & Ebright, R.H. (1993). CAP interacts with RNA polymerase in solution in the absence of promoter DNA. *Nature*, **364**, 548-549.

Hochschild, A., Irwin, N. & Ptashne, M. (1983). Repressor structure and the mechanism of positive control. *Cell*, **32**, 319-325.

Hochschild, A. & Ptashne, M. (1986). Cooperative binding of λ repressors to sites separated by integral turns of the DNA helix. *Cell*, **44**, 681-687.

Hochschild, A. & Ptashne, M. (1988). Interaction at a distance between λ repressors disrupts gene activation. *Nature*, **336**, 353-357.

Hu, J.C. & Gross, C.A. (1985). Mutations in the sigma subunit of *E. coli* RNA polymerase which affect positive control of transcription. *Mol. Gen. Genet.*, **199**, 7-13.

Hu, J.C. & Gross, C.A. (1988). Mutations in rpoD that increase expression of genes in the *mal* regulon of Escherichia coli K-12. *J. Mol. Biol.*, **203**, 15-27.

Igarashi, K., Fujita, N. & Ishihama, A. (1991). Identification of a subunit assembly domain in the alpha subunit of *Escherichia coli* RNA polymerase. *J. Mol. Biol.*, **218**, 1-6.

Igarashi, K. & Ishihama, A. (1991). Bipartite functional map of the *E. coli* RNA polymerase α subunit: Involvement of the C-terminal region in transcription activation by cAMP-CRP. *Cell*, **65**, 1015-1022.

Ishihama, A. (1988). Promoter selectivity of prokaryotic RNA polymerases. *TIG*, **4**, 282-286.

Jeon, Y.H., Negishi, T., Shirakawa, M.,Yamazaki, T., Fujita, N., Ishihama, A. & Kyogoku, Y. (1995). Solution structure of the activator contact domain of the RNA polymerase α subunit. *Science*, **270**, 1495-1497.

Jin, D.J. & Gross, C.A. (1991). RpoB8, a Rifampicin-resistant termination-proficient RNA polymerase, has an increased K_m for purine nucleotides during transcription elongation. *J. Biol. Chem.*, **266**, 14478-14485.

Johnson, A.D., Meyer, B.J. & Ptashne, M. (1979). Interactions between DNA-bound repressors govern regulation by the λ phage repressor. *Proc. Natl. Acad. Sci.*, **76**, 5061-5065.

Jordan, S.R. & Pabo, C.O. (1988). Structure of the lambda complex at 2.5 Å resolution. *Science*, **242**, 893-899.

Joung, J.K., Koepp, D.M. & Hochschild, A. (1994). Synergistic activation of transcription by bacteriophage λ cI protein and *E. coli* cAMP receptor protein. *Science*, **265**, 1863-1865.

Joung, J.K., Le, L.U. & Hochschild, A. (1993). Synergistic activation of transcription by *Escherichia coli* cAMP receptor protein. *Proc. Natl. Acad. Sci.*, **90**, 3083-3087.

Kashlev, M., Lee, J., Zalenskaya, K., Nikiforov, V. & Goldfarb, A. (1990). Blocking of the initiation-to-elongation transition by a transdominant RNA polymerase mutation. *Science*, **248**, 1006-1009.

Kimura, M. & Ishihama, A. (1995). Functional map of the alpha subunit of *Escherichia coli* RNA polymerase: amino acid substitution within the amino-terminal assembly domain. *J. Mol. Biol.*, **254**, 342-349.

Kolb, A., Busby, S., Buc, H., Garges, S. & Adhya, S. (1993). Transcriptional regulation by cAMP and its receptor protein. *Annu. Rev. Biochem.*, **62**, 749-795.

Kuldell, N. & Hochschild, A. (1994). Amino acid substitutions in the -35 recognition motif of σ70 that result in defects in phage λ repressor-stimulated transcription. *J. Bacteriol.*, **176**, 2991-2998.

Kumar, A., Grimes, B., Fujita, N., Makino, K., Malloch, R.A., Hayward, R.S. & Ishihama, A. (1994). Role of the sigma70 subunit of *Eshcerichia coli* RNA polymerase in transcription activation. *J. Mol. Biol.*, **235**, 405-413.

Laurenson, P. & Rine, J. (1992). Silencers, silencing, and heritable transcriptional states. *Microbiol. Rev.*, **56**, 543-560.

Lee, J. & Goldfarb, A. (1991). *lac* repressor acts by modifying the initial transcribing complex so that it cannot leave the promoter. *Cell*, **66**, 793-798.

Lee, I.W., Livrelli, V., Park, S.-J., Totis, P. & Summers, A.O. (1993). In vivo DNA-protein interactions at the divergent mercury resistance (*mer*) promoters. II. Represor/activator (MerR)-RNA polymerase interaction with *merOP* mutants. *J. Biol. Chem.*, **268**, 2632-2639.

Li, M., Moyle, H. & Susskind, M.M. (1994). Target of the transcriptional activation function of phage λ cI protein. *Science*, **263**, 75-77.

Livrelli, V., Lee, I.W. & Summers, A.O. (1993). *In vivo* DNA-protein interactions at the divergent mercury resistance (*mer*) promoters. I. Metalloregulatory protein MerR mutants. *J. Biol. Chem.*, **268**, 2623-2631.

Lund, P. & Brown, N. (1989). Up-promoter mutatons in the positively-regulated *mer* promoter of TN501. *Nucleic Acids Res.*, **17**, 5517-5527.

Malan, T.P., Kolb, A., Buc, H. & McClure, W.R. (1984). Mechanism of CRP-cAMP activation of *lac* operon transcription initiation: Activation of the P1 promoter. *J. Mol. Biol.*, **180**, 881-909.

Matthews, K.S. (1992). DNA looping. *Microbiol. Rev.*, **56**, 123-136.

Maurer, R., Meyer, B.J. & Ptashne, M. (1980). Gene regulation at the right operator (O_R) of bacteriophage λ. I. O_R3 and autogenous negative control by repressor. *J. Mol. Biol.*, **139**, 147-161.

McClure, W.R. (1985). Mechanism and control of transcription initiation in prokaryotes. *Annu. Rev. Biochem.*, **54**, 171-204.

McKay, D.B., Weber, I.T. & Steitz, T.A. (1982). Structure of the catabolite gene activator protein at 2.9-Å resolution. *J. Biol. Chem.*, **257**, 9518-9524.

Menendez, M., Kolb, A. & Buc, H. (1987). A new target for CRP action at the *malT* promoter. *EMBO J.*, **6**, 4227-4234.

Merkel, T.J., Dahl, J.L., Ebright, R.H. & Kadner, R.J. (1995). Transcription activation at the *Escherichia coli uhpT* promoter by the catabolite gene activator protein. *J. Bacteriol.*, **177**, 1712-1718.

Meyer, B.J., Maurer, R. & Ptashne, M. (1980). Gene regulation at the right operator (O_R) of bacteriophage λ. II. O_R1, O_R2, and O_R3: their roles in mediating th effects of repressor and cro. *J. Mol. Biol.*, **139**, 163-194.

Meyer, B.J. & Ptashne, M. (1980). Gene regulation at the right operator (O_R) of bacteriophage λ. III. l repressor directly activates gene transcription. *J. Mol. Biol.*, **139**, 195-205.

Milla, M.E., Brown, B.M. & Sauer, R.T. (1993). P22 Arc repressor: enhanced expression of unstable mutants by addition of polar C-terminal sequences. *Protein Science*, **2**, 2198-2205.

Milla, M.E., Brown, B.M. & Sauer, R.T. (1994). Protein stability effects of a complete set of alanine substitutions in Arc repressor. *Nature Struct. Biol.*, **1**, 518-523.

Milla, M.E., Brown, B.M., Waldburger, C.D. & Sauer, R.T. (1995). P22 Arc repressor: transition state properties inferred from mutational effects on the rates of protein unfolding and refolding. *Biochemistry*, **34**, 13914-13919.

Milla, M.E. & Sauer, R.T. (1994). P22 Arc repressor: folding kinetics of a single-domain, dimeric protein. *Biochemistry*, **33**, 1125-1133.

Milla, M.E. & Sauer, R.T. (1995). Critical side-chain interactions at a subunit interface in the Arc repressor dimer. *Biochemistry*, **34**, 3344-3351.

Narayan, S., Widen, S. G., Beard, W. A., & Wilson, S. H. (1994). RNA Polymerase II transcription: rate of promoter clearance is enhanced by a purified activating transcription factor cAMP response element binding protein. *J. Biol. Chem.*, **269**, 12755-12763.

Niu, W., Zhou, Y., Dong, Q., Ebright, Y. & Ebright, R.H. (1994). Characterization of the activating region of *Escherichia coli* catabolite gene activator protein (CAP). I. Saturation and alanine-scanning mutagenesis. *J. Mol. Biol.*, **243**, 595-602.

North, A.K., Klose, K.E., Stedman, K.M. & Kustu, S. (1993). Prokaryotic enhancer-binding proteins reflect eukaryote-like modularity: the puzzle of nitrogen regulatory protein C. *J. Bacteriol.*, **175**, 4267-4273.

O'Halloran, T.V., Frantz, B., Shin, M.K., Ralston, D.M. & Wright, J.G. (1989). The MerR heavy metal receptor mediates positive activation in a topologically novel transcription complex. *Cell*, **56**, 119-129.

Pabo, C.O. & Lewis, M. (1982). The operator-binding domain of λ repressor: structure and DNA recognition. *Nature*, **298**, 443-447.

Pabo, C.O., Sauer, R.T., Sturtevant, J.M. & Ptashne, M. (1979). The λ repressor contains two domains. *Proc. Natl. Acad. Sci.*, **76**, 1608-1612.

Parkhill, J., Ansari, A.Z., Wright, J.G., Brown, N.L. & O'Halloran, T.V. (1993). Construction and characterization of a mercury-independent MerR activator (MerRAC): transcriptional activation in the absence of Hg(II) is accompanied by DNA distortion. *EMBO J.*, **12**, 413-421.

Parkhill, J. & Brown, N.L. (1990). Site-specific insertion and deletion mutants in the *mer* promoter-operator region of Tn501; the nineteen base-pair spacer is essential for normal induction of the promoter by MerR. *Nucleic Acids Res.*, **18**, 5157-5162.

Perez-Martin, J. & Espinosa, M. (1993). Protein-induced bending as a transcriptional switch. *Science*, **260**, 805-807.

Perez-Martin, J., Rojo, F. & De Lorenzo, V. (1994). Promoters responsive to DNA bending: a common theme in prokaryotic gene expression. *Microbiol. Rev.*, **58**, 268-290.

Phillips, S.E.V. (1991). Specific β-sheet interactions. *Curr. Opin. Struct. Biol.*, **1**, 89-98.

Popham, D.L., Szeto, D., Keener, J. & Kustu, S. (1989). Function of a bacterial activator protien that binds to transcriptional enhancers. *Science*, **243**, 629-635.

Ptashne, M. (1986). *A Genetic Switch.* Cell Press, Cambridge, MA.

Raumann, B.E., Brown, B.M. & Sauer, R.T. (1994a). Major groove DNA recognition by β-sheets: the ribbon-helix-helix family of gene regulatory proteins. *Curr. Opin. Struct. Biol.*, **4**, 36-43.

Raumann, B.E., Rould, M.A., Pabo, C.O. & Sauer, R.T. (1994b). DNA recognition by β-sheets in the arc repressor-operator crystal structure. *Nature*, **367**, 754-757.

Roberts, J.W., Roberts, C.W. & Mount, D.W. (1977). Inactivation and proteolytic cleavage of phage λ repressor *in vitro* in an ATP-dependent reaction. *Proc. Natl. Acad. Sci.*, **74**, 2283-2287.

Robinson, C.R. & Sauer, R.T. (1996). Covalent attachment of Arc repressor subunits by a peptide linker enhances affinity for operator DNA. *Biochemistry*, **35**, 109-116.

Roe, J.-H., Burgess, R.R. & Record, M.T., Jr. (1984). Kinetics and mechanism of the interaction of *Escherichia coli* RNA polymerase with the λ P_R promoter. *J. Mol. Biol.*, **176**, 495-521.

Ross, W., Gosink, K.K., Salomon, J., Igarashi, K., Zou, C., Ishihama, A., Severinov, K. & Gourse, R.L. (1993). A third recognition element in bacterial promoters: DNA binding by the α subunit of RNA polymerase. *Science*, **262**, 1407-1413.

Ross, W., Park, S.-J., Summers, A. O. (1989). Genetic analysis of transcriptional activation and repression in the Tn21 *mer* operon. *J. Bacteriol.*, **171**, 4009-4018.

Russo, F.D. & Silhavy, T.J. (1992). Alpha: the cinderella subunit of RNA polymerase. *J. Biol. Chem.*, **267**, 14515-14518.

Sasse-Swight, S. & Gralla, J.D. (1990). Role of eukaryotic-type functional domains found in the prokaryotic enhancer receptor factor σ54. *Cell*, **62**, 945-954.

Sauer, R.T., Ross, M.J. & Ptashne, M. (1982). Cleavage of the λ and P22 repressors by recA protein. *J. Biol. Chem.*, **257**, 4458-4462.

Schickor, P., Metzger, W., Werel, W., Lederer, H. & Heumann, H. (1990). Topography of intermediates in transcription initiation of *E. coli*. *EMBO J.*, **9**, 2215-2220.

Schildbach, J.F., Milla, M.M., Jeffrey, P.D., Raumann, B.E., & Sauer, R.T. (1995) Crystal structure, folding, and operator binding of the hyperstable Arc repressor mutant PL8. *Biochemistry*, **34**, 1405-1412.

Schultz, S.C., Shields, G.C. & Steitz, T.A. (1991). Crystal Structure of a CAP-DNA complex: the DNA is bent by 90°. *Science*, **253**, 1001-1007.

Siebenlist, U., Simpson, R.B., & Gilbert, W.A. (1980). E. coli RNA polymerase ineracts homologously with two different promoters. *Cell*, **20**, 269-281.

Siegele, D.A., Hu, J.C. & Gross, C.A. (1988). Mutations in rpoD, the gene encoding the σ70 subunit of Escherichia coli RNA polymerase, that increase expression of the lac operon in the absence of CAP-cAMP. *J. Mol. Biol.*, **203**, 29-37.

Siegele, D.A., Hu, J.C., Walter, W.A. & Gross, C.A. (1989). Altered promoter recognition by mutant forms of the σ70 subunit of *Escherichia coli* RNA polymerase. *J. Mol. Biol.*, **206**, 591-603.

41

Silva, J.L., Silveira, C.F., Correia, J.A. & Pontes, L. (1992). Dissociation of a native dimer to a molten globule monomer. Effects of pressure and dilution on the association equilibrium of Arc repressor. *J. Mol. Biol.*, **223**, 545-555.

Smith, T.L. & Sauer, R.T. (1995). P22 Arc Repressor: Role of cooperativity in repression and binding to operators with altered half-site spacing. *J. Mol. Biol.*, **249**, 729-742.

Somers, W. S. & Phillips, S. E. V. (1992). Crystal structure of the *met* repressor-operator complex at 2.8 Å resolution reveals DNA recognition by β-strands. *Nature*, **359**, 387-393.

Spassky, A., Kirkegaard, K. & Buc, H. (1985). Changes in the DNA structure of *lac UV5* promoter during formaiton of an open complex with *Escherichia coli* RNA polymerase. *Biochemistry*, **24**, 2723-2731.

Straney, D.C. & Crothers, D.M. (1987). Comparison of the open complexes formed by RNA polymerse at the *Escherichia coli lac UV5* promoter. *J. Mol. Biol.*, **193**, 279-292.

Straney, D.C., Straney, S.B. & Crothers, D.M. (1989). Synergy between *Escherichia coli* CAP protein and RNA polymerase in the *lac* promoter open complex. *J. Mol. Biol.*, **206**, 41-57.

Suh, W.C., Leirmo, S. & Record, M.T., Jr. (1992). Roles of Mg^{2+} in the mechanism of formation and dissociation of open complexes between Escherichia coli RNA polymerase and the lPR promoter: kinetic evidence for a second open complex requiring Mg^{2+}. *Biochemistry*, **31**, 7815-7825.

Summers, A.O. (1992). Untwist and shout: a heavy metal-responsive transcriptional regulator. *J. Bacteriol.*, **174**, 3097-3101.

Susskind, M.M. & Youderian, P., ed. *Bacteriophage P22 antirepressor and its control.* Lambda II, ed. Hendrix, R.W., *et al.* 1983, Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY.

Tjian, R. & Maniatis, T. (1994). Transcriptional activation: a complex puzzle with few easy pieces. *Cell*, **77**, 5-8.

Valenzuela, D. & Ptashne, M. (1989). P22 repressor mutants deficient in co-operative binding and DNA loop formation. *EMBO J.*, **8**, 4345-4350.

Vershon, A.K., Bowie, J.U. & Sauer, R.T. (1986b). Isolation and analysis of Arc repressor mutants: evidence for an unusual mechanism of DNA binding. *Proteins: Structure, Function, and Genetics*, **1**, 302-311.

Vershon, A.K., Liao, S.-M., McClure, W.R. & Sauer, R.T. (1987a). Bacteriophage P22 Mnt repressor: DNA binding and effects on transcription *in vitro*. *J. Mol. Biol.*, **195**, 311-322.

Vershon, A.K., Liao, S.-M., McClure, W.R. & Sauer, R.T. (1987b). Interaction of the bacteriophage P22 Arc repressor with operator DNA. *J. Mol. Biol.*, **195**, 323-331.

Vershon, A.K., Youderian, P., Susskind, M.M. & Sauer, R.T. (1985). The bacteriophage P22 Arc and Mnt repressors. Overproduction, purification, and properties. *J. Biol. Chem.*, **260**, 12124-12129.

von Hippel, P.H., Bear, D.G., Morgan, W.D. & McSwiggen, J.A. (1984). Protein-nucleic acid interactions in transcription: a molecular analysis. *Annu. Rev. Biochem.*, **53**, 389-446.

Waldburger, C., Gardella, T., Wong, R. & Susskind, M.M. (1990). Changes in conserved region 2 of Escherichia coli σ70 affecting promoter recognition. *J. Mol. Biol.*, **215**, 267-276.

Waldburger, C.D., Schildbach, J.F. & Sauer, R.T. (1995). Are buried salt bridges important for protein stability and conformational specificity? *Nat. Struct. Biol.*, **2**, 122-128.

Whipple, F.W., Kuldell, N.H., Cheatham, L.A. & Hochschild, A. (1994). Specificity determinants for the interaction of λ repressor and P22 repressor dimers. *Genes & Dev.*, **8**, 1212-1223.

Wolffe, A.P. (1994). Transcription: in tune with the histones. *Cell*, **77**, 13-16.

Woody, S.T., Fong, R.S.-C. & Gussin, G. (1993). Effects of a single base-pair deletion in the baceriophage λ PRM Promoter. *J. Mol. Biol.*, **229**, 37-51.

Young, R.A. (1991). RNA polymerase II. *Annu. Rev. Biochem.*, **60**, 689-715.

Zhou, Y., Busby, S. & Ebright, R.H. (1993a). Identification of the functional subunit of a dimeric transcription activator protein by use of oriented heterodimers. *Cell*, **73**, 375-379.

Zhou, Y., Merkel, T.J. & Ebright, R.H. (1994a). Characterization of the activating region of Escherichia coli catabolite gene activator protein (CAP). II. Role at class I and class II CAP-dependent promoters. *J. Mol. Biol.*, **243**, 603-610.

Zhou, Y., Pendergrast, P.S., Bell, A., Williams, R., Busby, S. & Ebright, R.H. (1994b). The functional subunit of a dimeric transcription activator protein depends on promoter architecture. *EMBO J.*, **13**, 4549-4557.

Zhou, Y., Zhang, X. & Ebright, R.H. (1993b). Identification of the activating region of catabolite gene activator protein (CAP): Isolation and characterization of mutants of CAP specifically defective in transcription activation. *Proc. Natl. Acad. Sci.*, **90**, 6081-6085.

Zinkel, S.S. & Crothers, D.M. (1991). Catabolite activator protein-induced DNA bending in transcription initiation. *J. Mol. Biol.*, **219**, 201-215.

Chapter 2

P22 Arc Repressor: Role of Cooperativity in Repression and Binding to Operators with Altered Half-Site Spacing

Introduction

Many important biological processes, including the regulation of transcription, are mediated by proteins which bind to specific DNA sequences in cooperation with other proteins (for reviews see Ptashne, 1986; Tjian & Maniatis, 1992; Buratowski, 1994). However, in only a few cases is such cooperativity well understood in both structural and functional terms. The Arc repressor of bacteriophage P22 is a member of the ribbon-helix-helix family of DNA binding proteins (Phillips, 1991; Raumann et al., 1994a) and binds to a 21 base pair (bp) operator as a tetramer (Brown et al., 1990). In the cocrystal structure of the Arc-operator complex (Raumann et al., 1994b), each dimer contacts a six bp sequence in each DNA half-site and also interacts with the adjacently bound dimer (Figure 1a). These cooperative interactions stabilize the tetrameric complex kinetically and thermodynamically (Brown et al., 1990; Brown & Sauer, 1993). Biologically, Arc functions as part of the immunity I operon to repress transcription of the *antirepressor* gene and its own gene initiated from the P_{ant} promoter during lytic phage growth (Susskind & Youderian, 1983).

The dimer-dimer interface in the Arc-operator complex is formed by the side chains of Met1, Arg31, Ser35, Gln39, and portions of the protein main chain from adjacent dimers (Raumann et al., 1994b). Although the functional side-chain groups of Ser35 and Gln39 make some dimer-dimer contacts, these appear to be relatively unimportant as alanine substitutions at either position have little effect on operator binding (Brown et al., 1994). Substituting alanine for Met1 reduces cooperativity but only moderately (Brown et al., 1994). The side chains of Arg31 and Arg31', by contrast, make a critical set of hydrogen bonds across the interface to the backbone carbonyl oxygens of Asn29' and Asn29 (Figure 1b); mutating Arg31 to Ala (RA31) drastically reduces the operator affinity of the Arc dimer and results in non-cooperative binding of dimers to the operator (Brown et al., 1994). However, the RA31 mutation also disrupts a wild-type salt-bridge interaction between Arg31, Glu36, and Arg40, causing a dramatic destabilization of the Arc dimer (Breg et al., 1990; Brown et al., 1994; Milla et al., 1994; Raumann et al., 1994b). The RA31 mutant is inactive *in vivo*, but it is difficult to know if this phenotype is caused by its cooperativity defect, its stability defect, or a combination of the two.

Here, we show that replacing Ser35 with Arg or Leu disrupts cooperative binding but maintains near wild-type stability of the dimer (see Figure 1b for the position of

45

Ser35 in the dimer-dimer interface). By analyzing the phenotypes of these mutants *in vitro* and *in vivo*, we show that cooperativity is necessary to allow operator occupancy and repression at low protein concentrations *in vivo* but is not structurally required for repression. We also test the effects of operator spacing variants on the cooperativity of Arc binding and demonstrate that the ability of Arc dimers to interact cooperatively is critically dependent on the spacing between DNA half-sites. This indicates that the Arc dimers need to be positioned precisely by DNA binding to allow formation of the Arg31-mediated protein-protein interactions.

Results

(a) *Rationale of Mutant Design*

To determine the importance of dimer-dimer cooperativity in Arc function, we sought mutations that would prevent cooperativity without causing significant stability defects. In the wild-type cocrystal structure, Ser35 is part of the dimer-dimer interface but is almost fully solvent exposed in the dimer structure alone (Breg *et al.*, 1990; Raumann *et al.*, 1994b). As a result, we reasoned that introducing larger side chains such as Leu, Ile or Arg at this position might block cooperative interactions sterically without significantly affecting protein stability or operator half-site binding. The Ser35→Leu (SL35) and Ser35→Arg (SR35) mutations were constructed by cassette mutagenesis in an *arc* gene containing six C-terminal histidines (*arc-st6*) and the mutant proteins and Arc-st6 were purified to homogeneity for biochemical studies (see Methods).

(b) *Expression and activities in vivo*

The SR35-st6 and SL35-st6 mutants were found to be expressed as well as Arc-st6 when expressed from the P_{tac} promoter in strain UA2F following induction with IPTG (data not shown). The repressor activities of these proteins *in vivo* were monitored by the assay systems shown in Figure 10 and described in Methods. Wild-type Arc is active in both the multiple copy (*rpsl*) and single copy (*cat*) assays under inducing (IPTG) and non-inducing conditions (Table 1). The SR35 and SL35 mutants are inactive in both assays under non-inducing conditions and display only partial activity under inducing conditions (Table 1). Thus, the SR35-st6 and SL35-st6 proteins are unable to repress P_{ant}-mediated transcription when expressed at low levels in the cell and have only partial repressor activity when expressed at higher levels.

(c) *Stability assays*

Since Arc is a mixture of denatured monomers and folded dimers at the low concentrations at which DNA binding is observed (Bowie & Sauer, 1989a), the equilibrium constant for dimer dissociation and unfolding (K_u) of each mutant protein must be determined before its DNA binding affinity can be calculated. Values of K_u can be obtained from denaturation studies and then used in equation 1 (see Materials and Methods) to calculate the concentration of dimer at each total protein concentration in DNA binding experiments. Figure 2 shows urea and thermal dissociation/denaturation curves for Arc-st6, SR35-st6, and SL35-st6. The SR35 mutant is less stable than Arc-st6

47

and the SL35 mutant is slightly more stable than Arc-st6 in both assays. The equilibrium constants for unfolding/dissociation calculated for each protein from these experiments are listed in Table 2. Table 2 also lists the melting temperatures (t_m), enthalpies of unfolding/dissociation (ΔH_u), free energies for unfolding/dissociation (ΔG_u), and dependencies of unfolding/dissociation free energies on urea concentration (m -values).

(d) *Equilibrium DNA binding*

The binding of Arc-st6 and the mutant proteins to DNA fragments bearing a single operator half-site (L1) or the intact operator (O1) (see Figure 9) was assayed by DNA mobility shift experiments as shown in Figures 3 and 4. After corrections for differences in dimer stabilities, the half-site affinities calculated for the wild-type and mutant dimers are within experimental error ($K_1 \approx 0.2$–0.4 nM; Table 3). As shown in Figure 5a, the binding curves of Arc-st6 and SL35-st6 for the L1 half-site operator are nearly superimposible.

Although the SR35 and SL35 mutants have essentially wild-type affinity for an operator half-site, they bind significantly more poorly than wild-type Arc to the intact O1 operator (Figures 4 and 5b). The equilibrium constant (K_2) for dissociation of the mutant tetrameric complexes to give free operator and two dimers increases approximately 300-fold for SR35 and 600-fold for SL35 (Table 3). If these decreases in affinity are due solely to changes in the cooperativity, then the cooperative free energies for the SR35 and SL35 mutants are reduced by approximately 2.2 kcal/mol dimer and 2.5 kcal/mol dimer, respectively. Consistent with reductions in cooperative binding energy, a dimer-DNA complex is observed with the SR35 and SL35 mutants but is not observed with the wild-type protein (Figure 4). The SL35-st6 intact operator binding curve is also less steep than the Arc-st6 intact operator binding curve (Figure 5b), as expected for a less cooperative DNA binding reaction.

(e) *Dissociation rate experiments*

Because an Arc dimer dissociates from a single half-site with a half-life of about 2 sec whereas the DNA-bound tetramer stabilized by cooperative interactions has a half-life of about 1 h (Brown & Sauer, 1993), measuring the dissociation kinetics of the tetramer-DNA complex also provides a sensitive method that can be used to assess changes in cooperativity. Arc mutants defective in cooperativity should have fast dissociation rates. The dissociation rates of the wild-type and mutant tetramer-operator

complexes were determined using DNA mobility shift assays (not shown). Both the SR35-operator complex and SL35-operator complex have half-lives of less than 20 sec (Table 3), consistent with a significant reduction in cooperativity in the mutant tetrameric complexes.

(f) *Repression activity assays in vitro*

In principle, the inactivity of the SR35 and SL35 mutant proteins in the cell could be caused by an inability to bind the operator strongly enough to achieve full occupancy and/or could result from an inability of the operator-bound mutants to prevent RNA polymerase from binding P_{ant} and initiating transcription. To test this latter possibility, the ability of the SR35 and SL35 mutants to prevent formation of open complexes of RNA polymerase and the P_{ant} promoter was assayed *in vitro*. RNA polymerase bound in the open complex protects promoter regions from digestion with DNase I (Straney & Crothers, 1987), and thus the ability of Arc or its variants to block open-complex formation can be assayed by measuring the extent of protection of the promoter by RNA polymerase in the presence of increasing concentrations of Arc. Figure 6 shows footprinting assays of this type for Arc-st6 and SR35-st6. Although more SR35-st6 is required to occupy the DNA (as expected because of its cooperativity defect), SR35-st6 is capable of preventing open complex formation in a manner proportional to its occupancy of the operator. Similar results were obtained with the SL35 mutant (data not shown). As a result, it seems unlikely that the inactivity of the mutants *in vivo* can be explained by an inherent inability of these proteins to repress transcription.

(g) *Footprinting*

To probe the interaction of the SR35 and SL35 mutants with operator DNA, footprinting experiments were conducted. The footprints of both mutant proteins on both strands of the operator were similar to those produced by Arc-st6 in hydroxyl radical and 1,10-phenanthroline-copper experiments (Figure 7). Wild-type Arc-st6 and each mutant protects the operator from cleavage by hydroxyl radicals at positions between and flanking the two DNA half-sites, suggesting that each protein interacts with the operator DNA in a fundamentally similar manner (Figure 7a). The 1,10-phenanthroline-copper cleavages with Arc-st6 and the two mutants also indicate similar DNA structures in the complexes: enhanced cleavages are seen in the centers of the half-sites where the minor groove is widened in the cocrystal structure, and protection from cleavage is seen between and flanking the half-sites. Arc-st6 and SR35-st6 always showed a strong

49

enhancement of 1,10-phenanthroline-copper cleavage in the half-sites. With the SL35-st6 protein, the degree of cleavage in the half-sites by this reagent varied slightly but was always enhanced above the control reaction without protein (Figure 7b).

(h) *Interactions with operator spacing variants*

In the wild-type 21 bp *arc* operator (O1), the centers of the half-sites are separated by 11 bp or about one turn of the DNA helix (Figure 9). Deletion of a single bp in the center of the operator has been shown to reduce binding of wild-type Arc (Vershon *et al.*, 1989). To assess in greater detail the importance of the half-site spacing in Arc binding and cooperativity, variant operators with spacings increased by 1, 3, 10, and 11 bp or decreased by 1 or 3 bp were constructed (Figure 9); we refer to these variants as +1, +3, +10, +11, -1, and -3, respectively. The binding of both Arc-st6 and SL35-st6 to each DNA fragment was assayed. Comparing the binding of Arc-st6 to the different operators allows determination of whether the different half-site spacings and orientations affect dimer-dimer cooperativity. Comparing the binding of Arc-st6 and SL35-st6 to each DNA fragment allows evaluation of the need for the wild-type cooperative interface.

In equilibrium DNA mobility shift experiments, Arc-st6 and SL35-st6 bound to each of the +1, +3, +10, +11, -1, and -3 mutant operators with a half-maximal concentration of approximately 1-3 nM (data not shown). For comparison, the half maximal concentration for cooperative binding of Arc-st6 to the wild-type O1 operator is about 0.2 nM and that for non-cooperative binding of the SL35 mutant is about 2 nM, indicating that these variant operators show a reduction in binding affinity comparable to that exhibited by the cooperativity defective protein mutant SL35 on the intact O1 operator (see Figure 5b). In dissociation rate experiments, the half-lives of the Arc-st6 complexes with the altered spacing operators were all less than 30 sec (data not shown) compared with a half-life of about 1 h on the wild-type operator. Thus, changing the operator spacing by even 1 bp causes significant reductions in the ability of Arc-st6 dimers to interact cooperatively when bound to DNA.

Figure 8 shows DNA mobility shifts with Arc-st6 and SL35-st6 using the intact operator fragment (O1), an operator half-site fragment that contains nonspecific flanking DNA (L2) and the operator fragments with altered spacings. Several points are evident: (*i*) For each spacing variant (and the L2 operator) a dimer-DNA complex is observed with wild-type Arc-st6 as well as with the SL35 mutant. This result is consistent with the cooperativity defects inferred from the reduced binding affinity and

rapid dissociation rates of Arc-st6 from the altered spacing operators (see above). (*ii*) Arc-st6 but not SL35-st6 forms tetrameric complexes on the L2 half-site operator, suggesting that tetramer formation on a DNA fragment bearing only one half-site is dependent on cooperative interactions using the normal interface between a specifically bound dimer and a nonspecifically bound dimer. (*iii*) Each spacing variant is capable of forming tetramer-DNA complexes with both Arc-st6 and SL35-st6 at nM protein concentrations. The Arc-st6 tetramer complexes could result from dimers bound specifically to the two half-sites or from a dimer bound specifically to one half-site interacting cooperatively with a nonspecifically bound dimer (as discussed above for the L2 operator). The latter possibility is not likely for the SL35-st6 tetramer complexes, however, because of the inability of this mutant to interact cooperatively with another dimer. Thus, it appears that both half-sites in the spacing variants can indeed be simultaneously occupied by specifically bound dimers. This is expected if there is no strong negative cooperativity since the two *arc* half-sites have similar sequences and dimer affinities (Brown and Sauer, 1993). (*iv*) The +10 operator shows three bound complexes with Arc-st6 and SL35-st6. These complexes could reflect alternative conformations of the dimer or tetramer complexes or the presence of a third, non-specifically bound dimer in a hexameric complex. (*v*) Many of the tetramer complexes with the varied spacing operators have mobilities different from those of the wild-type operator complex. This could indicate structural variations in DNA bending or distortion in the complexes. (*vi*) A comparison of Arc-st6 and SL35-st6 on each template shows differences in the intensities and mobilities of the tetramer bands in certain cases (+1, -1, +3, and +10). This could arise if some residual dimer-dimer interactions persist in the Arc-st6 complexes.

Discussion

To obtain Arc mutants which were reasonably stable but defective in cooperativity, we substituted larger amino acids for Ser35, one of the side chains in the dimer-dimer interface of the protein-DNA complex. The SR35 mutant is less stable than wild type, while the SL35 mutant is slightly more stable. However, both mutants give hydroxyl radical and 1,10-phenanthroline-copper footprints similar to those of wild-type Arc. Moreover, after correction for the small differences in dimer stability, both mutants have essentially wild-type affinities for operator half-sites. Thus, the observed 300-600 fold reductions in the affinity of these mutant dimers for intact operator DNA must result predominantly from cooperativity defects. Significant cooperativity defects of the SR35

and SL35 mutant dimers are also indicated by the rapid dissociation of the mutant protein-operator complexes and by the presence of dimer-operator intermediates in the equilibrium binding reactions. We assume that the larger mutant side chains at position 35 sterically prevent the normal close approach of the two DNA-bound dimers or interfere in some other way with the Arg31-mediated hydrogen bonds across the dimer-dimer interface.

Although both the SR35 and SL35 mutant proteins are expressed as well as wild-type Arc-st6 in the cell, they are inactive when expressed at low levels and only weakly active when expressed at high levels in both single and multiple copy assays for repression *in vivo*. These results show that dimer-dimer cooperativity is important for Arc repression of P_{ant} *in vivo*. In principle, cooperative interactions between the dimers might be required to distort the DNA in a manner that prevents RNA polymerase open complex formation, to create a particular protein-protein interface recognized by RNA polymerase, or to contribute energetically to DNA binding and saturation of the operator at low protein concentrations in the cell. Because the DNA-bound mutants are capable of preventing open complex formation by RNA polymerase *in vitro*, however, the first two possibilities are unlikely. The partial activity seen with the SR35 and SL35 mutants at high levels of expression in the cell is also most consistent with a purely energetic contribution of cooperativity to Arc binding and repression. Therefore, the inactivity of both mutants at low protein concentrations *in vivo* appears to result from an inability to saturate the operator.

Our conclusion that dimer-dimer cooperativity is needed in the Arc system primarily to allow saturation of the operator *in vivo* is supported by the properties of Arc mutants in which Arg31 (the primary cooperativity residue) and its salt bridge partners Glu36 and Arg40 are replaced by hydrophobic residues (e.g., Met31-Tyr36-Leu40 or MYL; Waldburger *et al.*, 1995). The MYL mutant is defective in cooperative binding to the intact *arc* operator *in vitro* but is much more stable than wild-type Arc. Because folding/dimerization and DNA binding are coupled reactions, the cooperativity and stability effects cancel, and the MYL mutant binds the operator *in vitro* with the same half-maximal concentration as wild-type Arc. Moreover, the MYL mutant is fully active *in vivo* (± IPTG) as assayed by the same methods used here (Waldburger *et al.*, 1995). Thus, the properties of the MYL, SR35, and SL35 mutants all suggest that the main role of dimer-dimer cooperativity in repression is to increase occupancy of the Arc operator.

The role of dimer-dimer cooperativity in maintaining the structure of the DNA in the Arc-operator complex can also be addressed with the cooperativity mutants studied here. In the Arc-DNA cocrystal structure, the operator DNA is bent and slightly unwound, and the minor groove is widened in the center of each half-site and narrowed between the half-sites (Raumann et al., 1994b). The Arc dimers in the complex contact each other over the center of the operator where the largest bend and the greatest narrowing of the minor groove is observed, suggesting that cooperative interactions between the two dimers might help to maintain this bent and distorted DNA structure. In the wild-type complex, 1,10-phenanthroline-copper enhancements in the centers of the operator half-sites and strong hydroxyl radical and 1,10-phenanthroline-copper protections between the half-sites are consistent with distortion of the minor groove of the Arc-DNA complex in solution (Figure 7 and Brown et al., 1994). The cooperativity mutants, however, give hydroxyl radical and 1,10-phenanthroline-copper footprints similar to those of wild-type Arc-st6 (Figure 7), indicating that strong dimer-dimer contacts are not needed for these observed DNA distortions. Therefore, the bending and distortion of the DNA seen in the cocrystal structure may result not from dimer-dimer interactions but rather from individual dimer-DNA interactions.

The cooperative contacts between operator-bound Arc dimers consist of hydrogen bonds which by their nature are highly dependent on distance and orientation. As such, these dimer-dimer interactions would be expected to be sensitive to changes in the spacing of operator half-sites, since each one bp change results in a 3.4 Å translation along and a 36° rotation around the helical axis of the DNA. In fact, each spacing variant analyzed (+1, +3, +10, +11, -1, and -3) appeared to be bound by wild-type Arc-st6 with reduced cooperativity as shown by the presence of dimer-operator complexes, by reduced binding affinity, and by rapid dissociation of the protein-DNA complexes. Even those spacing variants which maintain the half-sites on the same side of the DNA helix (+10 and +11) did not display strong cooperative binding. Thus, the wild-type cooperativity contacts are not sufficiently strong to permit energetically significant dimer-dimer contacts if the spacing and orientation of the operator half-sites is altered. Moreover, with the spacing variants tested, no other part of the wild-type protein is able to provide protein-protein contacts that can substitute for the wild-type cooperativity contacts.

The regulation of transcription frequently involves protein-protein contacts between proteins bound to separate DNA sites, and two major classes of such interactions have emerged: those which are flexible and do not strongly depend on the spacing

53

between the DNA binding sites and those for which the spacing is an integral part of specific binding and regulation. Examples of the first class of flexible interactions include the binding of Lac repressor to distant operator sites (Mossing & Record, 1986; Oehler et al., 1990), the pairwise cooperativity of λ cI and HKO22 repressor dimers (Hochschild & Ptashne, 1986; Mao et al., 1994), the interaction of NtrC with σ^{54}-containing RNA polymerase holoenzymes (Popham et al., 1989; Su et al., 1990), and the interactions of eukaryotic enhancer-binding proteins with the basal transcription machinery through coactivator proteins (Dynlacht et al., 1991; Goodrich et al., 1993; Arias et al., 1994; Chen et al., 1994; Kwok et al., 1994).

The second class of regulators requires specific spacings between DNA binding sites. Examples include cooperative binding of MetJ dimers to adjacent binding sites (Phillips et al., 1989), interaction between LexA monomers bound to neighboring half-sites (Kim & Little, 1992; Oertel-Buchheit et al., 1993), contact between λ cI repressor and RNA polymerase at the P_{RM} promoter (Guarente et al., 1982; Woody et al., 1993; Li et al., 1994), DNA binding of the MCM1/α2 complex in yeast (Smith & Johnson, 1992), and DNA binding of nuclear hormone receptor family members including the glucocorticoid receptor (Luisi et al., 1991) and the retinoid X receptors and their heterodimerization partners (Mangelsdorf et al., 1991; Naar et al., 1991; Umesono et al., 1991; Yu et al., 1991; Durand et al., 1992). Proteins in each of these systems, like Arc, recognize the fixed spacing between DNA binding sites as well as the sequence of each binding site, thereby using the energy and specificity of the protein-protein contacts as an essential part of DNA binding specificity. Thus, Arc provides a simple model system in which the role of cooperativity in determining DNA binding specificity can be understood in detailed structural terms.

Materials and Methods

(a) Buffers and solutions

SDS sample buffer is 62.5 mM Tris-HCl (pH 6.8), 3% sodium dodecyl sulfate, 5% β-mercaptoethanol, 10 % glycerol, and 0.1% bromophenol blue. Storage buffer is 50 mM Tris-HCl (pH 7.5), 250 mM KCl, and 0.2 mM EDTA. Urea denaturation buffer is 50 mM Tris-HCl (pH 7.5), 100 mM KCl, 0.1 mM EDTA, and various concentrations of Urea (0 to 5 M in stepwise increments). DNA mobility shift buffer is 10 mM Tris-HCl (pH 7.5), 100 mM KCl, 3 mM $MgCl_2$, 0.1 mM EDTA, 100 μg/ml bovine serum albumin, and 0.02% Nonidet P-40. Loading buffer for the equilibrium DNA mobility shift assays

is 50% glycerol, 10 mM Tris-HCl (pH 8.0), and 1 mM EDTA (pH 8.0). Competitor buffer is 10 mM Tris-HCl (pH 7.5), 100 mM KCl, 3 mM MgCl$_2$, 0.1 mM EDTA, 100 μg/ml bovine serum albumin, 0.02% Nonidet P-40, 10% glycerol, and 2.0 or 0.2 μM unlabeled *arc* operator DNA (O1 fragment). Footprinting buffer is 10 mM Tris-HCl (pH 7.5), 100 mM KCl, 10 mM MgCl$_2$, 1.5 mM CaCl$_2$, 0.1 mM EDTA, 100 μg/ml bovine serum albumin, and 0.02% Nonidet P-40. Footprinting quench solution is 2.5 M ammonium acetate, 20 mM EDTA, and 10 μg/ml salmon sperm DNA. Footprinting loading solution is 85% formamide, 10 mM NaOH, 1 mM EDTA, 0.1% bromophenol blue, and 0.1% xylene cyanole.

(b) *Construction of Arc mutants.*

Plasmid pSA600, which encodes Arc-st6 (the st6 tail adds six histidines to the C-terminus) (Milla *et al.*, 1993), was digested with BglII and MluI. These restriction sites overlap codons 31-32 and 39-41, respectively, in the *arc-st6* gene. Two mutagenic oligonucleotides were synthesized using an Applied Biosystems 381A DNA instrument and were annealed and ligated to the pSA600 backbone: 5'-GATCTGTGAAT(AC)(TG)CGAGATTTATCAA-3' and 5'-CGCGTTGATAAATCTCG(AC)(TG)ATTCACA-3'. The bases in parentheses indicate mixtures of two bases at single positions, enabling recovery of codons for Ser, Arg, Leu, and Ile at position 35 in the Arc sequence. The sequences of 19 isolates were determined by dideoxy strand terminating sequencing (Sanger *et al.*, 1977). The Ser35, Arg35, and the Leu35 alleles were recovered 7, 9, and 3 times, respectively, but the Ile35 allele was not recovered.

(c) *Protein purification*

Each Arc protein used in these studies contains six histidines added to the C-terminus for ease of purification. Previous studies have shown that these additional C-terminal residues have no significant effect on the stability or DNA binding properties of the protein (Milla *et al.*, 1993). Arc-st6 and the mutant SR35-st6 and SL35-st6 proteins were purified using nickel chelate chromatography (Qiagen) followed by SP-Sephadex chromatography as described (Milla *et al.*, 1993). The concentrations of Arc-st6 and the mutants (in moles of monomer equivalents per liter) were calculated using an extinction coefficient of 6756 M^{-1} cm^{-1} at 280 nm (Brown *et al.*, 1990). RNA polymerase holoenzyme was prepared from *Escherichia coli* strain MG1655 as described (Hager *et al.*, 1990), and its concentration was determined using an extinction coefficient of 2.79 x

10^5 M^{-1} cm^{-1} at 280 nm (calculated from the extinction coefficient of Lowe *et al.* (1979) assuming a molecular weight of 450,000 for RNA polymerase σ^{70} holoenzyme).

(d) *Arc repression assays in vivo*

Arc repression *in vivo* was assayed in strain UA2F (Vershon *et al.*, 1986a) containing the multicopy plasmid pSA600 (Milla *et al.*, 1993). Kanamycin resistance and Lac repressor (*lac I^q*) are encoded by an F' episome in this strain, and Arc-st6 or mutants are expressed from a P_{tac} promoter in pSA600. Since repression of P_{tac} by Lac repressor is leaky, some Arc is expressed even in the absence of isopropyl-β-D-thiogalactoside (IPTG). Two assays for Arc function *in vivo* were used (Figure 10). The first involves Arc repression of a multicopy reporter gene for the wild-type *rpsl* ribosomal subunit gene, driven by a P_{ant} promoter weakened by point mutation. Expression of the wild-type *rpsl* gene confers dominant streptomycin sensitivity to strain UA2F, which is normally resistant because it contains a chromosomal copy of the strA allele of *rpsl* (Bowie & Sauer, 1989b; Mossing *et al.*, 1991). For this assay, cells were grown at 37 °C on LB-agar plates containing 100 μg/ml ampicillin, 25 μg/ml kanamycin, 50 μg/ml streptomycin, and either no IPTG or 2 μg/ml IPTG. Cells containing active Arc survive under these conditions. The degree of function of each Arc protein was assessed in the streptomycin assay as either strong (a lawn of bacterial growth), moderate (patches of small colonies), weak (a few small colonies), or inactive (no colonies).

The second assay for activity in the cell involves Arc repression of a single copy chloramphenicol acetyl transferase (*cat*) gene expressed from a wild-type P_{ant} promoter (Figure 10; Vershon *et al.*, 1986a). Cells were grown at 37 °C on LB-agar plates containing 100 μg/ml ampicillin, 25 μg/ml kanamycin, 75 μg/ml cloramphenicol, and either no IPTG or 2 μg/ml IPTG. Cells containing active Arc do not survive in this plate assay. The degree of function of each Arc protein was assessed as either strong (no colonies), moderate (a few small colonies), weak (patches of small colonies), or inactive (a lawn of growth).

(e) *Expression assays*

The intracellular expression of the mutant proteins in strain UA2F was compared to that of Arc-st6 after induction with 100 μg/ml IPTG for 2.5 h at 37 °C. After this time, cells were lysed by boiling in SDS sample buffer, DNA was removed by centrifugation,

and samples were electrophoresed on 16% polyacrylamide Tris-Tricine SDS gels (Schagger & von Jagow, 1987). Arc was then visualized by Coomassie staining.

(f) Denaturation assays

Thermal and urea denaturation experiments were performed and analyzed essentially as described (Milla et al., 1993; Milla et al., 1994). All measurements were made with an AVIV 60DS spectropolarimeter equipped with a Hewlett-Packard temperature controller. The final concentration of protein in each experiment was 5 μM. Thermal denaturation experiments were performed in storage buffer and were monitored by changes in ellipticity at 222 nm. Urea denaturation experiments were performed at 20 °C in urea denaturation buffer and were monitored by changes in ellipticity at 234 nm.

(g) DNA fragments

The sequences of the operator DNA fragments used for binding assays are shown in Figure 9. The O1, L1, and L2 operator fragments were gifts of Bronwen Brown and had been purified by FPLC chromatography using a Pharmacia MonoQ column (Brown & Sauer, 1993). Oligonucleotides for the remaining operators were synthesized and purified by gel electrophoresis on denaturing gels. One oligonucleotide strand of each operator fragment was labeled with ^{32}P using T4 polynucleotide kinase and was then annealed with a two-fold excess of the complementary oligonucleotide (Brown & Sauer, 1993).
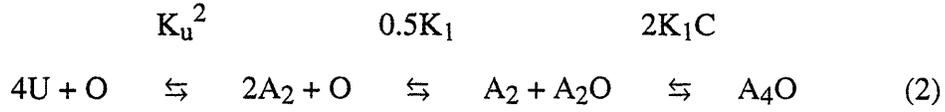
(h) Equilibrium DNA mobility shift assays

DNA mobility shift experiments were carried out as described (Brown & Sauer, 1993) using a DNA concentration of 10-20 pM. Binding reactions were incubated for at least 2 h at 20 °C in DNA mobility shift buffer. A one-tenth volume of loading buffer was added, and the reactions were loaded onto 7% acrylamide 0.5X TBE gels. Autoradiographs of gels were quantified by scanning densitometry.

At the nanomolar concentrations at which DNA binding is measured, Arc and its variants are an equilibrium mixture of inactive, denatured monomers (U) and active dimers (A$_2$) (Bowie & Sauer, 1989a). The equilibrium constants (K$_u$) for dimer dissociation/unfolding were determined from urea denaturation experiments (see above), and the concentrations of monomer and dimer at each total protein concentration (P$_t$) were calculated from from the equations:

$$[U] = \frac{-K_u + \sqrt{K_u^2 + 8K_uP_t}}{4} \qquad\qquad [A_2] = \frac{[U]^2}{K_u} \qquad\qquad (1)$$

The equilibria involved in the overall operator binding reaction are:

$$4U + O \quad\overset{K_u^2}{\leftrightarrows}\quad 2A_2 + O \quad\overset{0.5K_1}{\leftrightarrows}\quad A_2 + A_2O \quad\overset{2K_1C}{\leftrightarrows}\quad A_4O \qquad\qquad (2)$$

where O is free operator DNA, A_2O represents a dimer bound to either of the two operator half-sites, A_4O represents two dimers bound to the operator, K_1 is the microscopic equilibrium constant for dissociation of a dimer from a single half-site (assuming that both half-sites have the same affinity; Brown & Sauer, 1993), and C is the cooperativity constant. It is also useful to define an equilibrium constant (K_2) for dissociation of two dimers from the operator:

$$K_2 = K_1^2C = \frac{[A_2]^2[O]}{[A_4O]} \qquad\qquad (3)$$

Values of K_1 for Arc variants were determined by Scatchard analysis of DNA mobility shift experiments monitoring binding to the L1 DNA fragment (see Figure 9) containing a single operator half-site (O_h) as described (Brown et al., 1990; Brown & Sauer, 1993). The equilibrium expression for K_1 is:

$$K_1 = \frac{[A_2][O_h]}{[A_2O_h]} \qquad\qquad (4)$$

where O_h represents the half-site operator and A_2O_h represents the dimer-bound DNA. To calculate K_1, the intensities of the free DNA ($[O_h]$) and dimer-bound DNA ($[A_2O_h]$) were determined for each protein concentration and used to calculate the binding parameter Θ_d:

$$\Theta_d = \frac{[A_2O]}{[A_2O] + [O_h]} = \frac{1}{1 + \frac{[O_h]}{[A_2O]}} = \frac{1}{1 + \frac{K_1}{[A_2]}} \qquad\qquad (5)$$

The value of K_1 was then calculated from the slope, determined by linear regression, of the Scatchard form of equation 5:

$$\frac{\Theta_d}{[A_2]} = \frac{1}{K_1} - \frac{\Theta_d}{K_1} \qquad\qquad (6)$$

where [A$_2$] can be calculated from equation 1. In these experiments, residual free DNA was present even when the amount of shifted protein-L1 complex had clearly reached saturation (see Figure 3). This residual unshifted DNA is either incompetent to bind protein (possibly because it is not properly annealed) or represents complexes that dissociate upon entering the gel or early during electrophoresis. Because Θ_d reaches a maximum of 0.7-0.8 instead of 1.0, the binding curves for each data set were normalized by dividing the Θ_d obtained at each protein concentration by the maximum value of Θ_d. After this correction, the calculated binding curves matched the experimental data very well. Two observations suggest that the equilibrium constants determined from these DNA mobility shift experiments are valid. First, similar equilibrium dissociation constants for the L1-Arc-st6 complex are obtained by mobility shift assays and by competition assays (Brown & Sauer, 1993). Second, the equilibrium constant for Arc-st6 obtained from the mobility shift assays is within two-fold of that obtained by quantitative DNase I footprinting assays under similar binding conditions (data not shown).

Values of K$_2$ for Arc variants were determined from DNA mobility shift assays using the O1 DNA fragment (see Figure 9) containing the intact operator. The intensities of the free DNA ([O]) and tetramer-bound DNA ([A$_4$O]) were determined for each protein concentration and used to calculate the binding parameter Y:

$$Y = \frac{[A_4O]}{[A_4O] + [O]} = \frac{1}{1 + \frac{[O]}{[A_4O]}} = \frac{1}{1 + \frac{K_2}{[A_2]^2}} \tag{7}$$

The value of K$_2$ was calculated from the slope, determined by linear regression, of the Scatchard from of equation 7:

$$\frac{Y}{[A_2]^2} = \frac{1}{K_2} - \frac{Y}{K_2} \tag{8}$$

where [A$_2$] can be calculated from equation 1. In some DNA mobility shift assays (see Figure 4), there are two free DNA bands but only the upper band is shifted by Arc binding. The lower band, presumably unannealed or improperly annealed DNA, was not included in any calculations. It is important to note that the analysis described above is also valid for the cooperativity mutants which show dimer-bound intermediates ([A$_2$O])

in the binding assays. Values of the cooperativity constant (C) for each mutant protein were calculated from the measured values of K_1 and K_2 using equation 3. When values of K_1 and C were used to calculate the expected fractions of all DNA species (i.e., [O], [A$_2$O], and [A$_4$O]), as a function of protein concentration, the fits to experimental data were good and could be brought to within experimental error by changes of two-fold or less in the value of C.

i) *Dissociation rate experiments*

Arc at a concentration (20 nM) sufficient to bind at least 90% of the operator was incubated with 6000-7000 cpm ^{32}P-labeled DNA in 200 μl DNA mobility shift buffer. The dissociation reaction was initiated by adding an equal volume of competitor buffer, and at different times, 30 μl aliquots were loaded onto 7% acrylamide 0.5X TBE gels that were running at 300 V. After 10 min, the voltage was turned down to 150 V. In control reactions, premixing the labeled DNA and competitor buffer before addition of Arc resulted in no observable binding. The dissociation rate constant (k_d) for the complexes was determined by fitting data to the equation:

$$\ln\left(\frac{\Theta}{\Theta_o}\right) = -k_d t \qquad (9)$$

where Θ represents the DNA fraction bound by Arc at time t and Θ_o represents the DNA fraction bound by Arc at time zero.

(j) *Footprinting*

An end-labeled 130 bp DNA fragment containing the wild-type immunity I region of bacteriophage P22 was generated by the polymerase chain reaction using an EcoRI fragment of pMS200 (Youderian *et al.*, 1982) as a template and the primers 5'-TCTAGCCATGCCATCACTCC-3' and 5'-TGAACTGCGGCATTTTGCTC-3'. Only one of the primers was end-labeled with ^{32}P using T4 polynucleotide kinase, resulting in labeling of either the top or the bottom strand.

The ability of Arc-st6, SR35-st6, and SL35-st6 to prevent RNA polymerase open complex formation with the P_{ant} promoter was assayed *in vitro* by footprinting. Duplicate sets of reactions were performed in footprinting buffer to allow the binding of the Arc variants and the binding of RNA polymerase to be assessed. Arc was preincubated with the DNA (approximately 250 pM) for 20 min at 37 °C. RNA

polymerase (or buffer for the duplicate reactions) was then added to a concentration of 30 nM and the incubation was continued at 37 °C for an additional 40 min. DNase I was then added to a final concentration of 52 ng/ml for 1 min, and the digestion was stopped by the addition of a two-fold excess of footprinting quench solution.

Hydroxyl radical and 1,10-phenanthroline-copper cleavage reactions were also performed to compare the operator binding of Arc-st6 and the mutants. Both reagents cleave the DNA via a reaction initiated on the deoxyribose moiety in the minor groove, but 1,10-phenanthroline-copper is more sensitive to distortions in the geometry of the minor groove (Sigman *et al.*, 1991). The binding reactions for these experiments contained 70 nM protein and approximately 200-500 pM DNA in footprinting buffer. Reactions were incubated for at least 30 min at room temperature (22 ± 1 °C). Hydroxyl radical and 1,10-phenanthroline-copper cleavages were performed following general procedures (Dixon *et al.*, 1991; Sigman *et al.*, 1991) although the times of the reactions and the concentrations of the reagents were adjusted to give appropriate levels of DNA cleavage.

All footprinting reactions were extracted with phenol/chloroform, ethanol precipitated, washed with 75% ethanol, and dried. The pellets were resuspended in footprinting loading solution and run on 6% or 8% polyacrylamide gels containing 8.3 M urea and 1X TBE. Gels were exposed to a phosphor screen overnight, and traces of the footprints were obtained using a phosphorimager and the ImageQuant program (Molecular Dynamics). Bands were identified by comparison with Maxam-Gilbert guanine sequencing reactions.

References

Arias, J., Alberts, A.S., Brindle, P., Claret, F.X., Smeal, T., Karin, M., Feramisco, J. & Montminy, M. (1994). Activation of cAMP and mitogen responsive genes relies on a common nuclear factor. *Nature*, **370**, 226-229.

Bowie, J.U. & Sauer, R.T. (1989a). Equilibrium dissociation and unfolding of the Arc repressor dimer. *Biochemistry*, **28**, 7139-7143.

Bowie, J.U. & Sauer, R.T. (1989b). Identification of C-terminal extensions that protect proteins from intracellular proteolysis. *J. Biol. Chem.*, **264**, 7596-7602.

Breg, J.N., van Opheusden, J.H.J., Burgering, M.J.M., Boelens, R. & Kaptein, R. (1990). Structure of Arc repressor in solution: evidence for a family of β-sheet DNA-binding proteins. *Nature*, **346**, 586-589.

Brown, B.M., Bowie, J.U. & Sauer, R.T. (1990). Arc repressor is tetrameric when bound to operator DNA. *Biochemistry*, **29**, 11189-11195.

Brown, B.M., Milla, M.E., Smith, T.L. & Sauer, R.T. (1994). Scanning mutagenesis of the Arc repressor as a functional probe of operator recognition. *Nature Struct. Biol.*, **1**, 164-168.

Brown, B.M. & Sauer, R.T. (1993). Assembly of the Arc repressor-operator complex: cooperative interactions between DNA-bound dimers. *Biochemistry*, **32**, 1354-1363.

Buratowski, S. (1994). The basics of basal transcription by RNA polymerase II. *Cell*, **77**, 1-3.

Chen, J.-L., Attardi, L.D., Verrijzer, C.P., Yokomori, K. & Tjian, R. (1994). Assembly of recombinant TFIID reveals differential coactivator requirements for distinct transcriptional activators. *Cell*, **79**, 93-105.

Dixon, W.J., Hayes, J.J., Levin, J.R., Weidner, M.F., Dombroski, B.A. & Tullius, T.D. (1991). Hydroxyl radical footprinting. In *Methods in Enzymology: Protein-DNA Interactions* (Sauer, R.T., ed.), pp. 380-413, Academic Press, Inc., San Diego, CA.

Durand, B., Saunders, M., Leroy, P., Leid, M. & Chambon, P. (1992). All-trans and 9-cis retinoic acid induction of CRABPII is mediated by RAR-RXR heterodimers bound to DR1 and DR2 repeated motifs. *Cell*, **71**, 73-85.

Dynlacht, B.D., Hoey, T. & Tjian, R. (1991). Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. *Cell*, **66**, 563-576.

Goodrich, J.A., Hoey, T., Thut, C.J., Admon, A. & Tjian, R. (1993). Drosophila TAF$_{II}$40 interacts with both a VP16 activation domain and the basal transcription factor TFIIB. *Cell*, **75**, 519-530.

Guarente, L., Nye, J.S., Hochschild, A. & Ptashne, M. (1982). Mutant λ phage repressor with a specific defect in its positive control function. *Proc. Natl. Acad. Sci.*, **79**, 2236-2239.

Hager, D.A., Jin, D.J. & Burgess, R.R. (1990). Use of Mono Q high-resolution ion-exchange chromatogaphy to obtain highly pure and active *Escherichia coli* RNA polymerase. *Biochemistry*, **29**, 7890-7894.

Hochschild, A. & Ptashne, M. (1986). Cooperative binding of λ repressors to sites separated by integral turns of the DNA helix. *Cell*, **44**, 681-687.

Kim, B. & Little, J.W. (1992). Dimerization of a specific DNA-binding protein on the DNA. *Science*, **255**, 203-206.

Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. App. Cryst.*, **24**, 946-950.

Kwok, R.P.S., Lundblad, J.R., Chrivia, J.C., Richards, J.P., Bächinger, H.P., Brennan, R.G., Roberts, S.G.E., Green, M.R. & Goodman, R.H. (1994). Nuclear protein CBP is a coactivator for the transcription factor CREB. *Nature*, **370**, 223-226.

Li, M., Moyle, H. & Susskind, M.M. (1994). Target of the transcriptional activation function of phage λ cI protein. *Science*, **263**, 75-77.

Lowe, P.A., Hager, D.A. & Burgess, R.R. (1979). Purification and properties of the σ subunit of *Escherichia coli* DNA-dependent RNA polymerase. *Biochemistry*, **18**, 1344-1352.

Luisi, B.F., Xu, W.X., Otwinowski, Z., Freedman, L.P., Yamamoto, K.R. & Sigler, P.B. (1991). Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. *Nature*, **352**, 497-505.

Mangelsdorf, D.J., Umesono, K., Kliewer, S.A., Borgmeyer, U., Ong, E.S. & Evans, R.M. (1991). A direct repreat in the cellular retinol-binding protein type II gene confers differential regulation by RXR and RAR. *Cell*, **66**, 555-561.

Mao, C., Carlson, N.G. & Little, J.W. (1994). Cooperative DNA-protein interactions: effects of changing the spacing between adjacent binding sites. *J. Mol. Biol.*, **235**, 532-544.

Milla, M.E., Brown, B.M. & Sauer, R.T. (1993). P22 Arc repressor: enhanced expression of unstable mutants by addition of polar C-terminal sequences. *Protein Science*, **2**, 2198-2205.

Milla, M.E., Brown, B.M. & Sauer, R.T. (1994). Protein stability effects of a complete set of alanine substitutions in Arc repressor. *Nature Struct. Biol.*, **1**, 518-523.

Mossing, M.C., Bowie, J.U. & Sauer, R.T. (1991). A streptomycin selection for DNA-binding activity. In *Methods in Enzymology: Protein-DNA Interactions* (Sauer, R.T., ed.), pp. 604-619, Academic Press, Inc., San Diego, CA.

Mossing, M.C. & Record, Jr., M. T. (1986). Upstream operators enhance repression of the *lac* promoter. *Science*, **233**, 889-892.

Näär, A.M., Boutin, J.-M., Lipkin, S.M., Yu, V.C., Holloway, J.M., Glass, C.K. & Rosenfeld, M.G. (1991). The orientation and spacing of core DNA-binding motifs dictate selective transcriptional responses to three nuclear receptors. *Cell*, **65**, 1267-1279.

Oehler, S., Eismann, E.R., Kramer, H. & Müller-Hill, B. (1990). The three operators of the lac operon cooperate in repression. *EMBO J.*, **9**, 973-979.

Oertel-Buchheit, P., Schmidt-Dörr, T., Granger-Schnarr, M. & Schnarr, M. (1993). Spacing requirements between LexA operator half-sites can be relaxed by fusing the LexA DNA binding domain with some alterative dimerization domains. *J. Mol. Biol.*, **229**, 1-7.

Phillips, S.E.V. Specific β-sheet interactions. (1991). *Curr. Op. Struct. Biol.*, **1**, 89-98.

Phillips, S.E.V., Manfield, I., Parsons, I., Davidson, B.E., Rafferty, J.B., Somers, W.S., Margarita, D., Cohen, G.N., Saint-Girons, I. & Stockley, P.G. (1989). Cooperative tandem binding of *met* repressor of *Escherichia coli*. *Nature*, **341**, 711-715.

Popham, D.L., Szeto, D., Keener, J. & Kustu, S. (1989). Function of a bacterial activator protein that binds to transcriptional enhancers. *Science*, **243**, 629-635.

Ptashne, M. (1986). *A Genetic Switch*. Cell Press, Cambridge, MA.

Raumann, B.E., Brown, B.M. & Sauer, R.T. (1994a). Major groove DNA recognition by β-sheets: the ribbon-helix-helix family of gene regulatory proteins. *Curr. Opin. Struct. Biol.*, **4**, 36-43.

Raumann, B.E., Rould, M.A., Pabo, C.O. & Sauer, R.T. (1994b). DNA recognition by β-sheets in the Arc repressor-operator crystal structure. *Nature*, **367**, 754-757.

Sanger, F., Nicklen, S. & Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci.*, **74**, 5463-5467.

Schagger, H. & von Jagow, G. (1987). Tricine-sodium dodecyl sulfate-polyacrylamide gel elecrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.*, **166**, 368-379.

Sigman, D.S., Kuwabara, M.D., Chen, C.-H.B. & Brucie, T.W. (1991). Nuclease activity of 1,10-phenanthroline-copper in study of Protein-DNA interactions. In *Methods in Enzymology: Protein-DNA Interactions* (Sauer, R.T., ed.), pp. 414-433, Academic Press, Inc., San Diego.

Smith, D.L. & Johnson, A.D. (1992). A molecular mechanism for combinatorial control in yeast: MCM1 protein sets the spacing and orientation of the homeodomains of an α2 dimer. *Cell*, **68**, 133-142.

Straney, S.B. & Crothers, D.M. (1987). Lac repressor is a transient gene-activating protein. *Cell*, **51**, 699-707.

Su, W., Porter, S., Kustu, S. & Echols, H. (1990). DNA-looping and enhancer activity: association between DNA-bound NtrC activator and RNA polymerase at the bacterial *glnA* promoter. *Proc. Natl. Acad. Sci.*, **87**, 5504-5508.

Susskind, M.M. & Youderian, P. (1983). Bacteriophage P22 Antirepressor and its control. In *Lambda II* (Hendrix, R.W., Roberts, J.W., Stahl, F.W., & Weisberg, R. A., eds.), pp. 347-363. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Tjian, R. & Maniatis, T. (1994). Transcriptional activation: a complex puzzle with few easy pieces. *Cell*, **77**, 5-8.

Umesono, K., Murakami, K.K., Thompson, C.C. & Evans, R.M. (1991). Direct repreats as selective response elements for the thyroid hormone, retinoic acid, and vitamin D$_3$ receptors. *Cell*, **65**, 1255-1266.

Vershon, A.K., Blackmer, K. & Sauer, R.T. (1986). Mutagenesis of the Arc repressor using synthetic primers with random nucleotide substitutions. In *Protein Engineering: Applications in Science, Medicine, and Engineering* (Inouye, M. & Sarma, R., eds.), pp. 243-256, Academic Press, Orlando, FL.

Vershon, A.K., Kelley, R.D., & Sauer, R.T. (1989). Sequence-specific binding of Arc repressor to DNA. *J. Biol. Chem.*, **264**, 3267-3273.

Waldburger, C.D., Schildbach, J.F. & Sauer, R.T. (1995). Are buried salt bridges important for protein stability and conformational specificity? *Nature Struct. Biol.*, **2**, 122-128.

Woody, S.T., Fong, R.S.-C. & Gussin, G. (1993). Effects of a single base-pair deletion in the baceriophage λ P$_{RM}$ Promoter. *J. Mol. Biol.*, **229**, 37-51.

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

Yu, V.C., Delsert, C., Andersen, B., Holloway, J.M., Devary, O.V., Näär, A.M., Kim, S.Y., Boutin, J.-M., Glass, C.K. & Rosenfeld, M.G. (1991). RXRβ: a coregulator that enhances binding of retinoic acid, thyroid hormone, and vitamin D receptors to their cognate response elements. *Cell*, **67**, 1251-1266.

Protein	Str^R -IPTG	Str^R +IPTG	Cm^R -IPTG	Cm^R +IPTG
Arc-st6	+++	+++	+++	+++
SR35-st6	-	+/-	-	+/-
SL35-st6	-	+	-	+

Table 1. Summary of Repression Activities *In Vivo*.

The activities of each protein in tests for function monitored by resistance to streptomycin (Str^R assay) or chloramphenicol (Cm^R assay) are shown. Repression activities in strain UA2F are classified as strong (+++), moderate (+), weak (+/-), and inactive (-). These assays are depicted in Figure 10 and described in the Methods.

Protein	K_u (M)	ΔG_u (kcal/mol dimer)	m-value (kcal/mol-M)	t_m (°C)	ΔH_u (kcal/mol dimer)
Arc-st6	5.8×10^{-9}	11.0	1.66	58.4	60.0
SR35-st6	4.6×10^{-8}	9.8	1.67	53.2	51.2
SL35-st6	4.1×10^{-9}	11.1	1.58	60.6	53.3

Table 2. Urea and Thermal Denaturation Parameters

K_u and ΔG_u are the equilibrium constant and free energy change for dissociation/unfolding of each protein dimer (5 μM Arc, 20 °C, pH 7.5, 100 mM KCl, and 0.1 mM EDTA). The m-value is the slope of a plot of ΔG_u vs. [urea]. t_m is the temperature at which 50% of the protein is denatured (5 μM Arc, pH 7.5, 250 mM KCl, and 0.2 mM EDTA). ΔH_u is the enthalpy of dimer dissociation/unfolding (at $T=t_m$, 5 μM Arc, pH 7.5, 250 mM KCl, and 0.2 mM EDTA). The values shown are the averages of two experiments.

Protein	K_1 (M)	K_2 (M^2)	$t_{1/2}$
Arc-st6	1.6 ($\pm$ 1.0) x 10^{-10}	8.8 ($\pm$ 6.8) x 10^{-23}	78 min
SR35-st6	4.2 ($\pm$ 2.7) x 10^{-10}	2.7 ($\pm$ 1.8) x 10^{-20}	18 sec
SL35-st6	4.6 ($\pm$ 3.0) x 10^{-10}	5.3 ($\pm$ 1.9) x 10^{-20}	7 sec

Table 3. Equilibrium and Kinetic Parameters for DNA Binding.
The average ($\pm$ standard deviation) of three or more experiments is listed for K_1 (L1 half-site affinity) and K_2 (O1 intact operator affinity; K_2 is defined as $K_1{}^2C$ in equation 3 where C is the cooperativity parameter). $t_{1/2}$ is the average half-life of the tetramer-O1 operator complex from two experiments. Binding was measured by DNA mobility shift experiments performed at 20 °C in buffer containing 10 mM Tris-HCl (pH 7.5), 100 mM KCl, 3 mM MgCl$_2$, 0.1 mM EDTA, 100 μg/ml bovine serum albumin, and 0.02% Nonidet P-40.

Figure 1. (a) The Arc tetramer bound to operator DNA (Raumann *et al.*, 1994b). Arc is shown as a ribbon trace and the operator DNA as a backbone trace. This view is down the two-fold axis of the complex. The cooperative interface is boxed. (b) Close-up view of the cooperative interface. Arg31 and Arg31' make key cooperative hydrogen bonds (dotted lines) with the backbone carbonyls of Asn29' and Asn29. These hydrogen bonds stabilize the bound tetramer. Mutation of Ser35 and Ser35' to Arg or Leu causes a significant cooperativity defect. This figure was prepared using the program MOLSCRIPT (Kraulis, 1991).
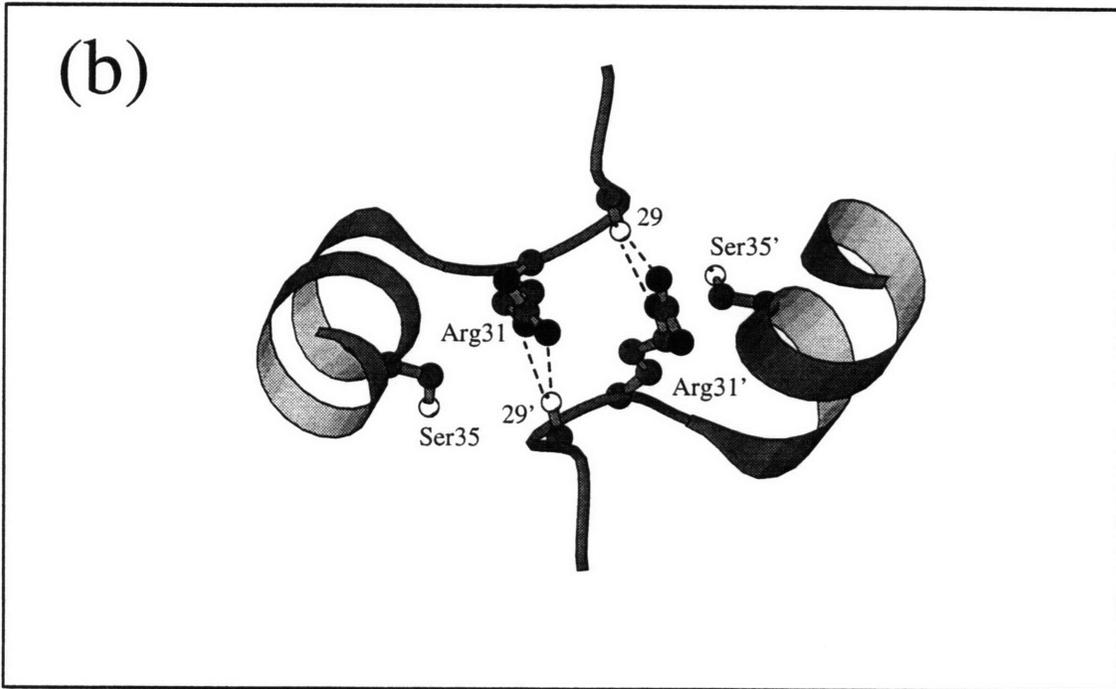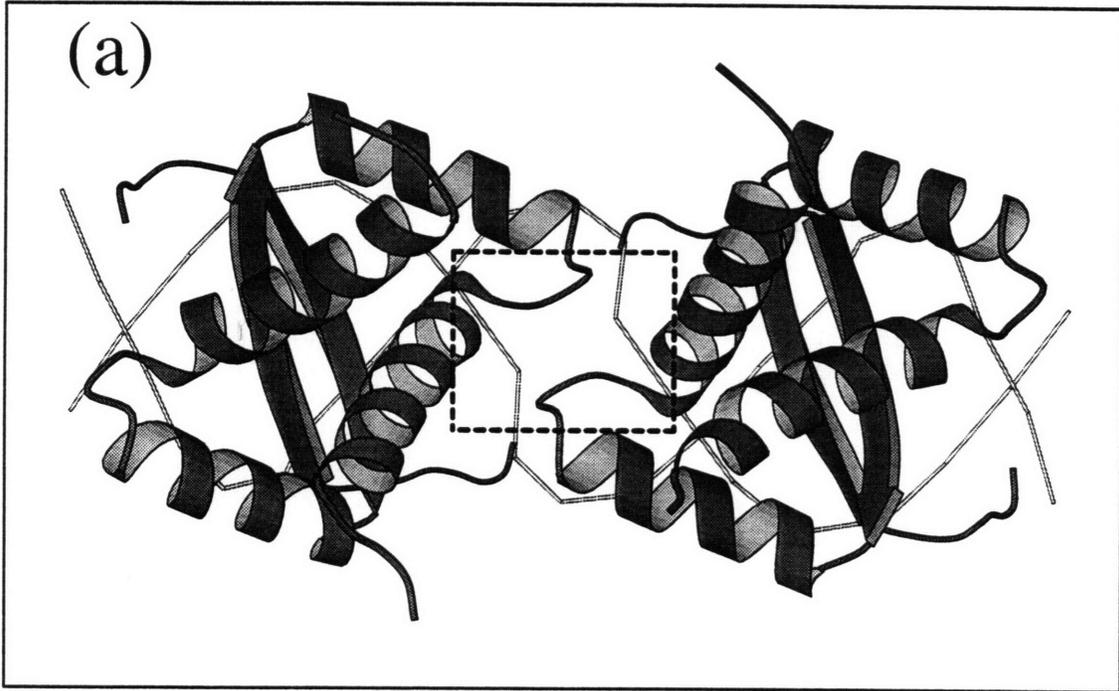
(a)

(b)

Figure 2. (a) Urea denaturation curves and (b) thermal denaturation curves for Arc-st6 (circles), SR35-st6 (diamonds), and SL35-st6 (squares). The results from a single experiment are shown for each protein. The solid lines in (a) show the fits of theoretical curves to the experimental data.
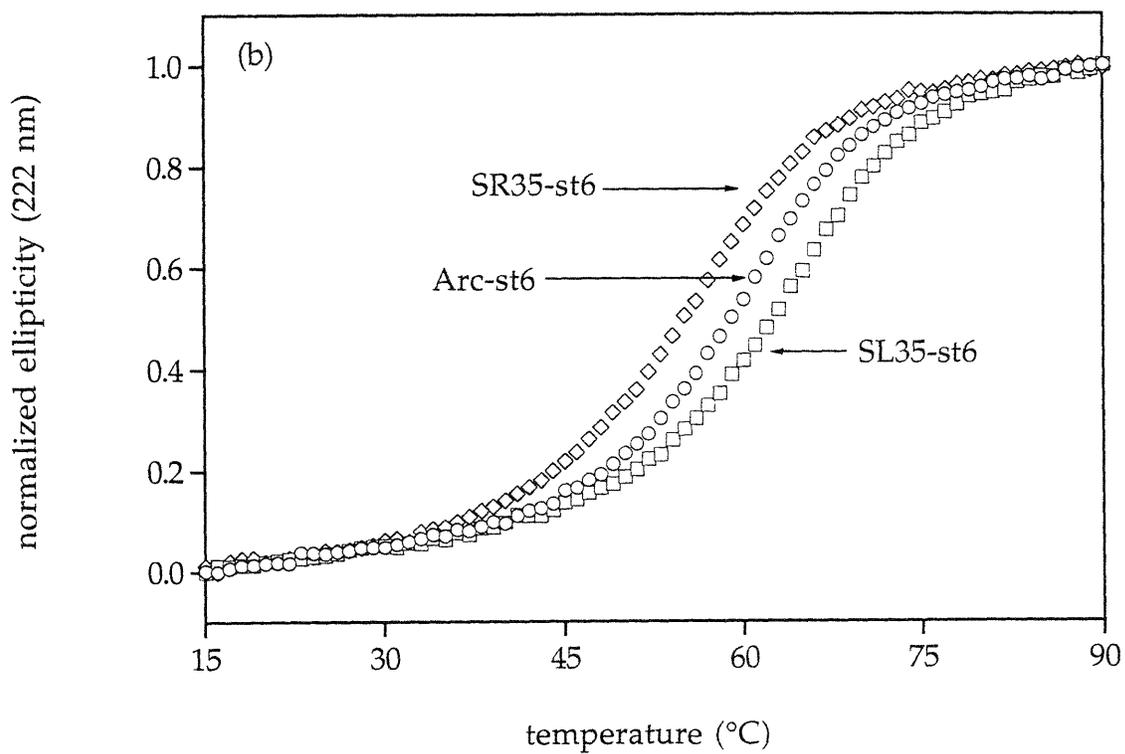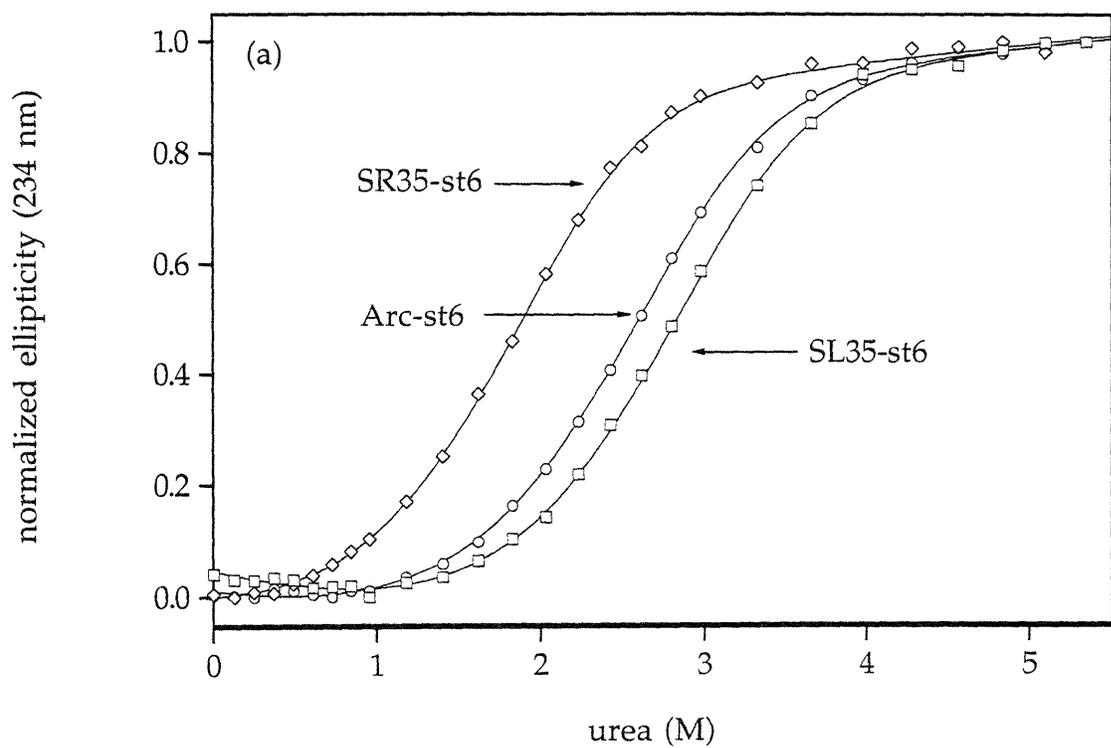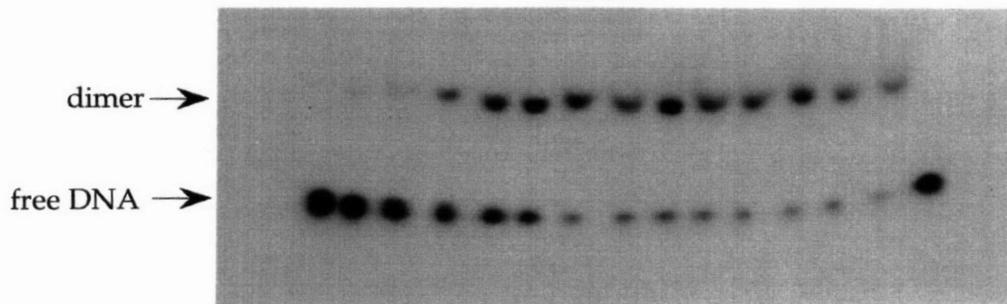
Figure 3. DNA mobility shift assays with the L1 half-site fragment (see Figure 3). The protein concentrations were increased from left to right in 1.7-fold increments starting from 1.5×10^{-10} M. The right-most lane in each gel contained no protein.

Arc-st6

dimer →

free DNA →

SR35-st6

dimer →

free DNA →

SL35-st6

dimer →

free DNA →
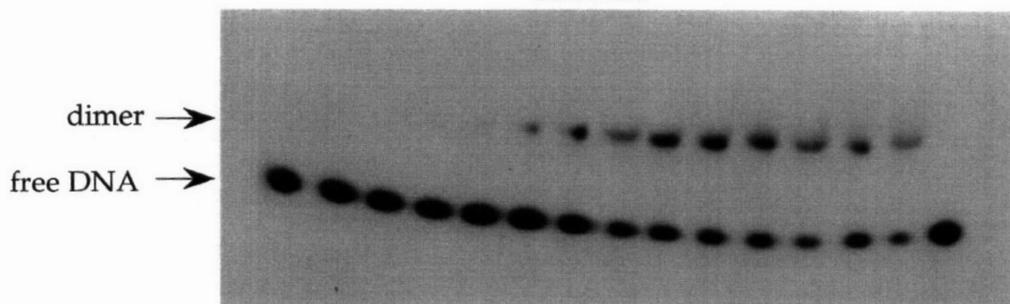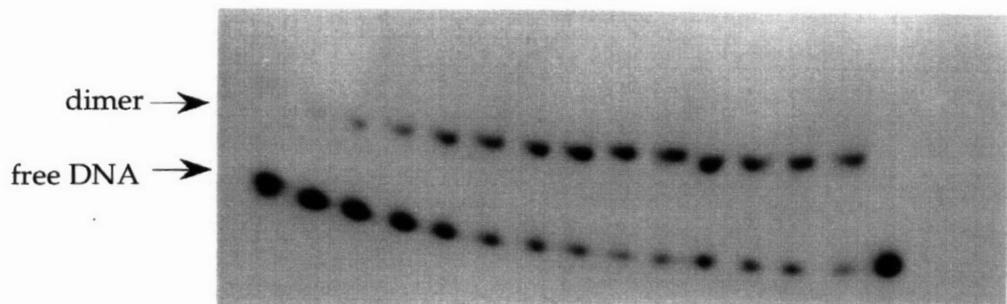
Figure 4. DNA mobility shift assays using the O1 DNA fragment containing the intact operator (see Figure 3). The protein concentrations were increased from left to right in 1.7-fold increments starting from 6.8 x10^{-11} M. The right-most lane in each gel contained no protein. The lower band of free DNA that does not shift is likely to be unannealed single stranded DNA.

Arc-st6

tetramer →

free DNA →

SR35-st6

tetramer →

dimer →

free DNA →

SL35-st6

tetramer →

dimer →

free DNA →
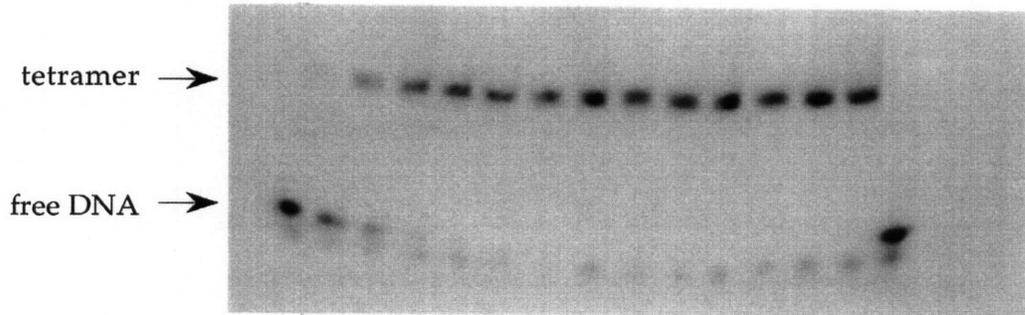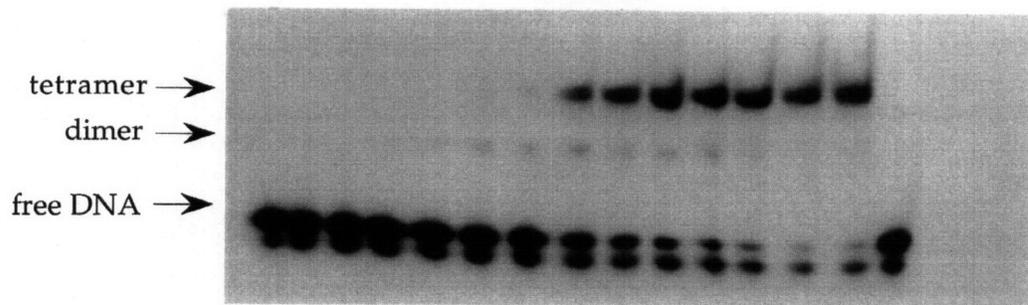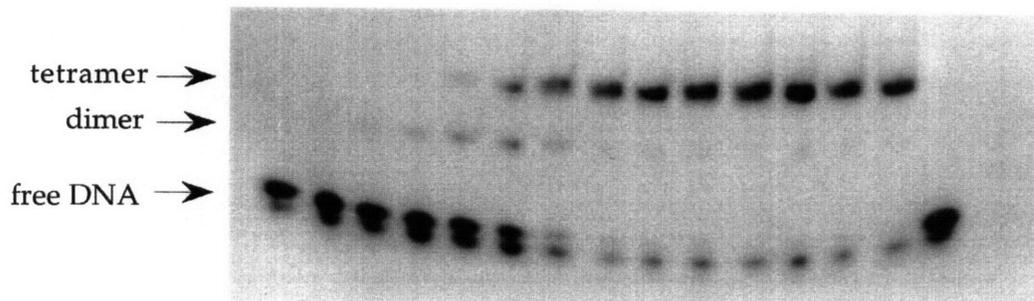
Figure 5. Binding curves for Arc-st6 (closed circles) and the SL35-st6 cooperativity mutant (open squares). (a) Binding to the L1 half-site fragment. The theoretical curves for Arc-st6 (dotted line) and SL35-st6 (solid line) were generated using K_1 values of 1.4 x 10^{-10} M and 2.2 x 10^{-10} M, respectively. (b) Binding curves to the O1 intact operator fragment. The theoretical curves for Arc-st6 (dotted line; $K_1 = 1.6$ x 10^{-10} M and C = 1.5 x 10^{-3}) and SL35-st6 (solid line; $K_1 = 4.6$ x 10^{-10} M and C = 2.9 x 10^{-1}) were calculated using the equation:

$$\Theta_t = \frac{[A_4O]}{[O] + [A_2O] + [A_4O]} = \frac{1}{\dfrac{K_1^2 C}{[A_2]^2} + \dfrac{2K_1 C}{[A_2]} + 1}$$

where Θ_t represents the fraction of tetramer-bound operator DNA, and $[A_2]$ can be calculated from the total Arc concentration using equation 1. Note that the SL35-st6 binding curve is shifted to higher protein concentrations and is less steep because of its reduced cooperative binding.

(a) L1 Half-Site Operator
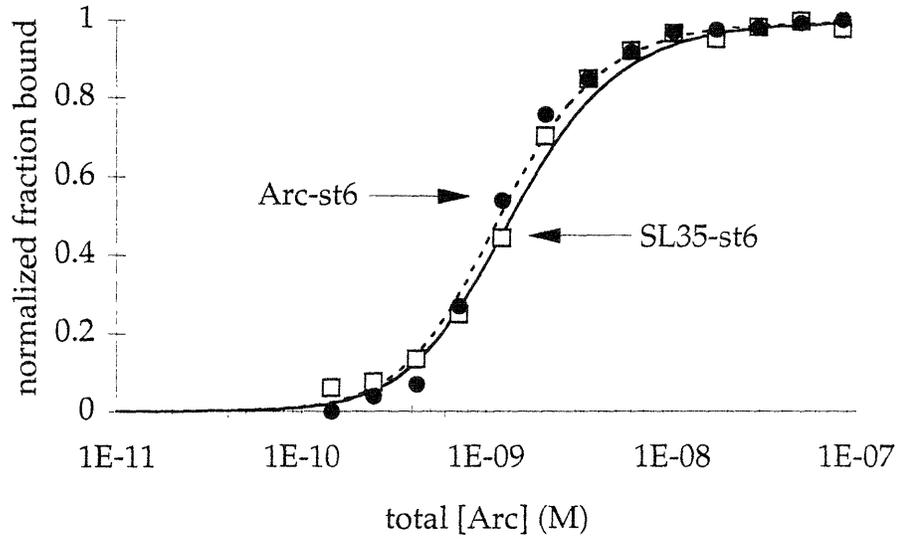
(b) O1 Intact Operator
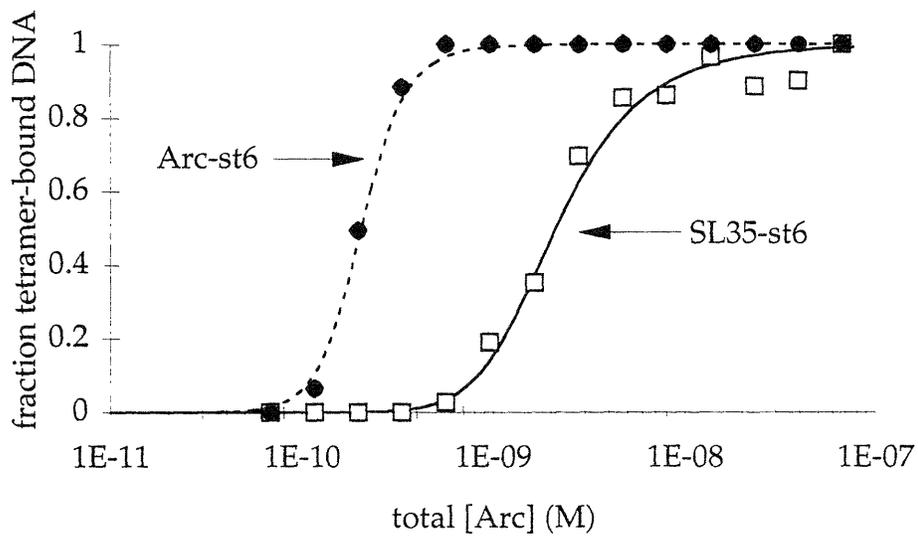
Figure 6. Open complex formation by RNA polymerase (RNAP) on the P_{ant} promoter assayed by DNase I footprinting in the presence of (a) Arc-st6 and (b) SR35-st6. The top strand of the DNA was labeled with ^{32}P. The Arc-st6, SR35-st6, and RNAP footprints are indicated by brackets. Note that the concentrations of both Arc-st6 and SR35-st6 required to saturate the operator are higher than in Figure 6 because these assays were conducted at a higher temperature (37°C) and with more operator DNA (250 pM).
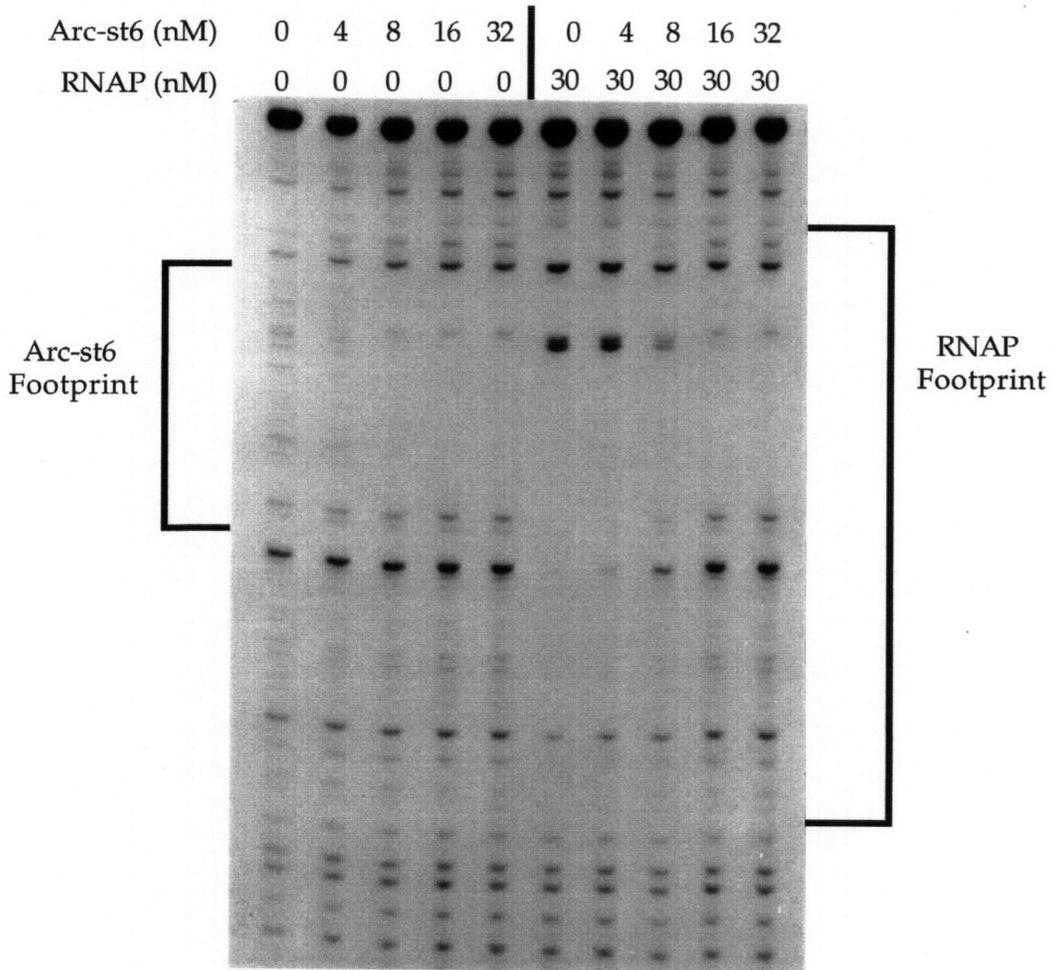
| Arc-st6 (nM) | 0 | 4 | 8 | 16 | 32 | 0 | 4 | 8 | 16 | 32 |
| RNAP (nM) | 0 | 0 | 0 | 0 | 0 | 30 | 30 | 30 | 30 | 30 |

Arc-st6
Footprint

RNAP
Footprint
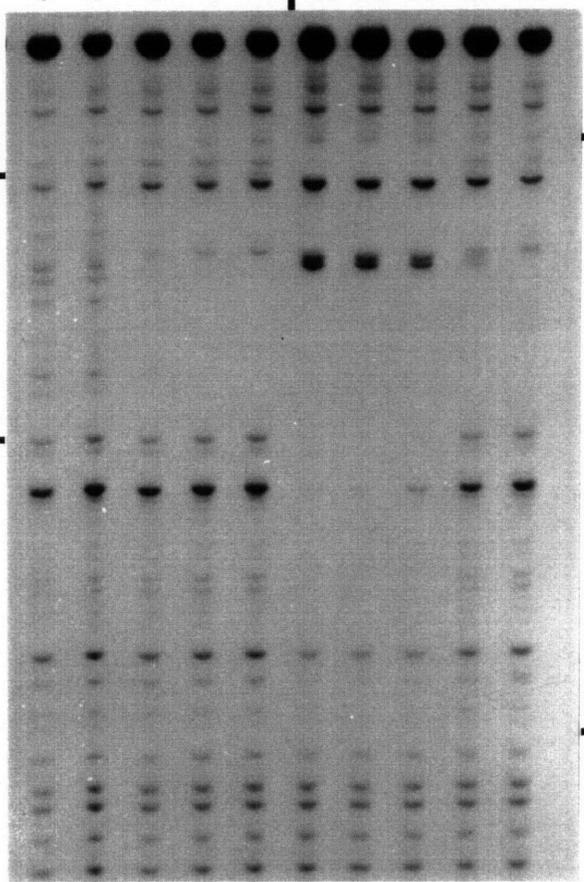
Figure 7. Densitometric traces of the (a) hydroxyl radical footprints and (b) 1,10-phenanthroline-copper footprints of Arc-st6, SR35-st6, and SL35-st6 on the top strand of the wild-type operator DNA. In both (a) and (b), the operator is shown below the traces: the hexamer sequences contacted in the cocrystal are boxed, and the operator is underlined with an oval at the central bp. In (a), bases significantly protected from cleavage by hydroxyl radicals are indicated with small squares. In (b), enhanced cleavages by 1,10-phenanthroline-copper are represented by arrows, with larger arrows indicating greater enhancement; bases that are unprotected but not greatly enhanced are indicated with small circles. The protection patterns shown for each protein on the bottom strand of the DNA were obtained from separate experiments for which the traces are not shown.

Arc-st6

SR35-st6

SL35-st6

DNA alone

TGACAT GATAGA AGCAC TCTACT ATATTC
ACTGTA CTATCT TCGTG AGATGA TATAAG

Arc-st6

SR35-st6

SL35-st6

DNA alone

```
TGACAT GATAGA AGCAC TCTACT ATATTC
ACTGTA CTATCT TCGTG AGATGA TATAAG
```
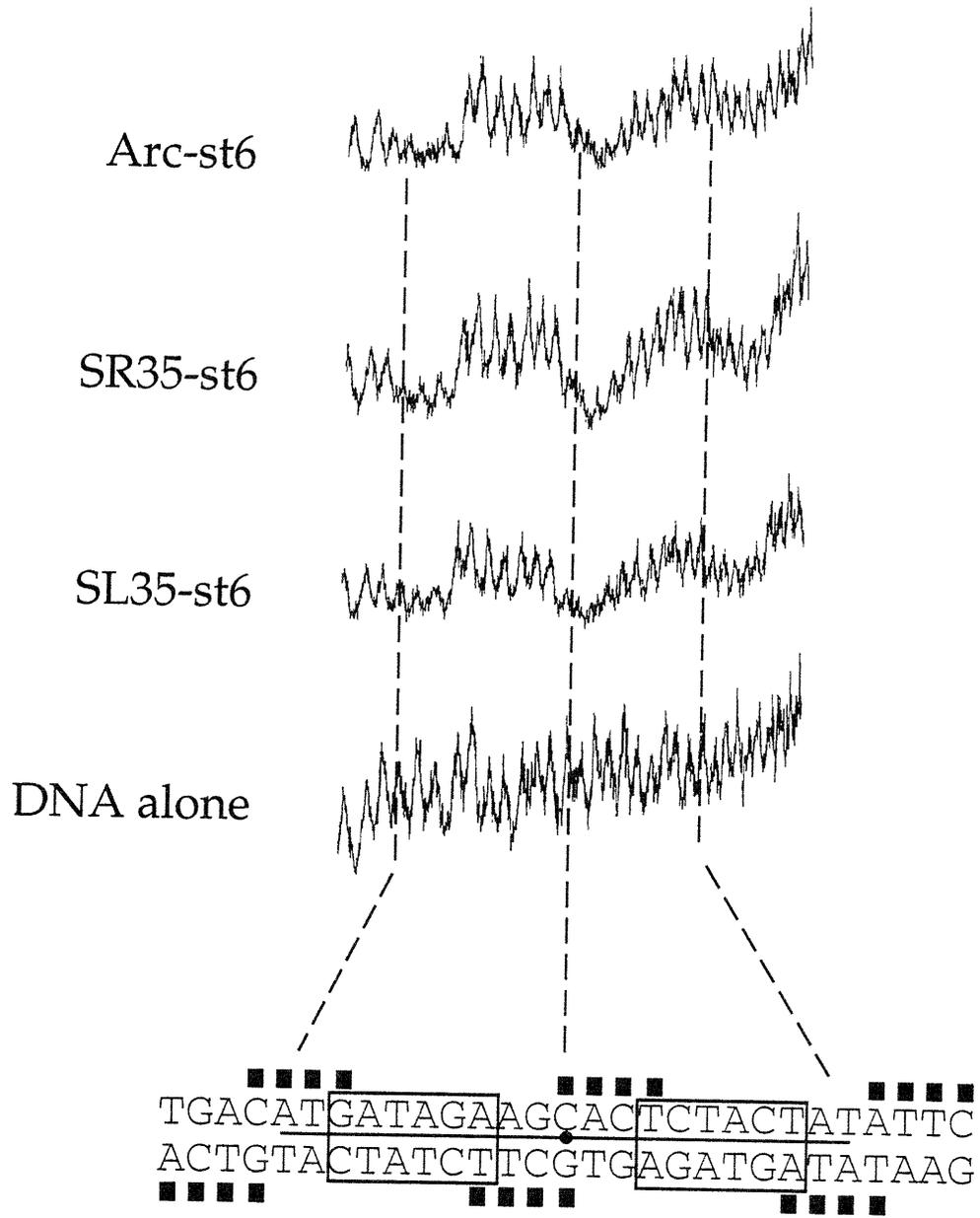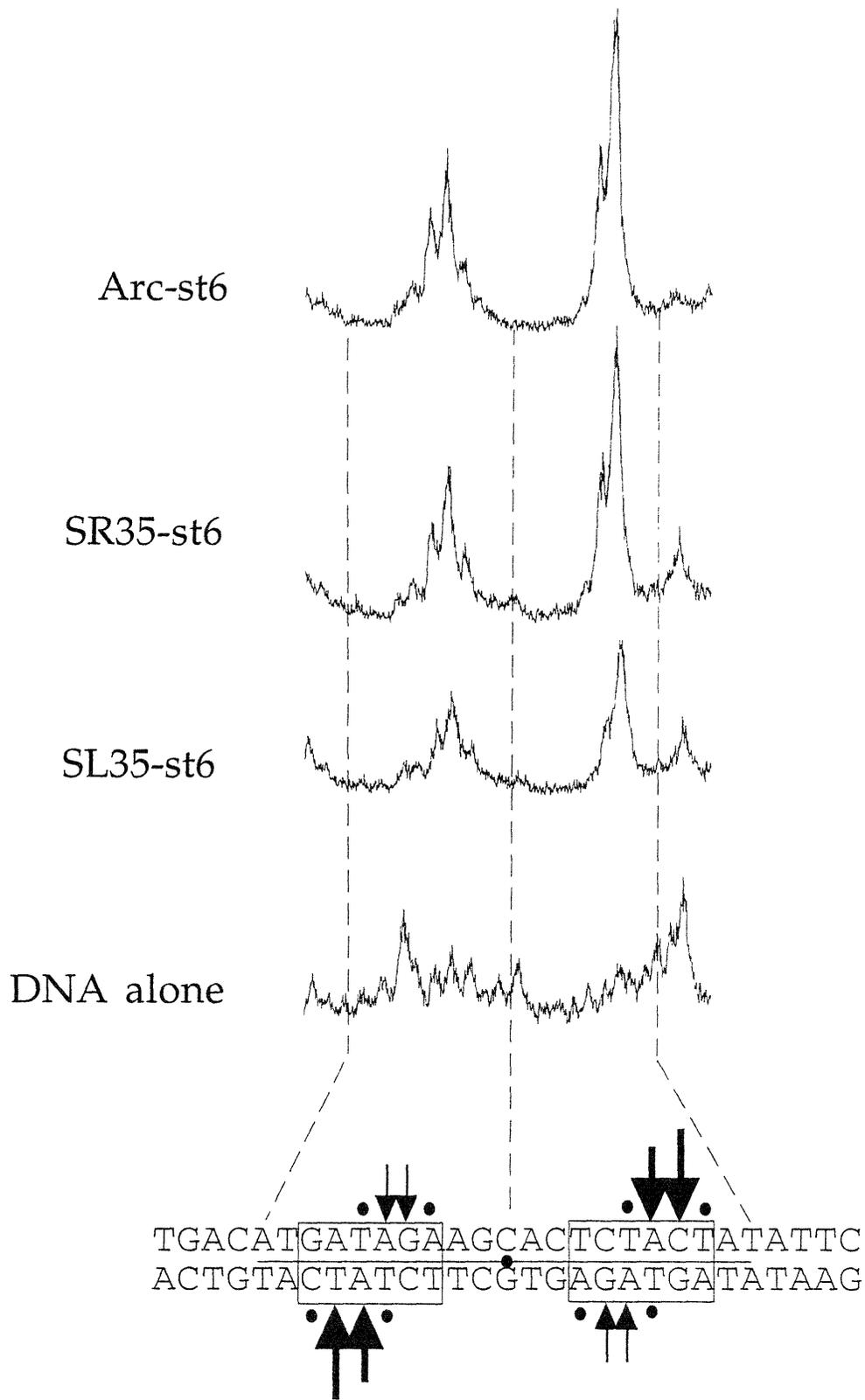
Figure 8. DNA mobility shift assays with the O1, L2, and altered spacing DNA fragments and Arc-st6 and SL35-st6. The concentrations of Arc-st6 and SL35-st6 used for each DNA fragment were: O1 operator (1 nM), L2 half-operator (3 nM), +1 operator (2 nM), -1 operator (2 nM), +3 operator (2 nM), -3 operator (2 nM), +10 operator (1 nM), and +11 operator (1 nM). See Figure 9 for the sequences of the DNA fragments.
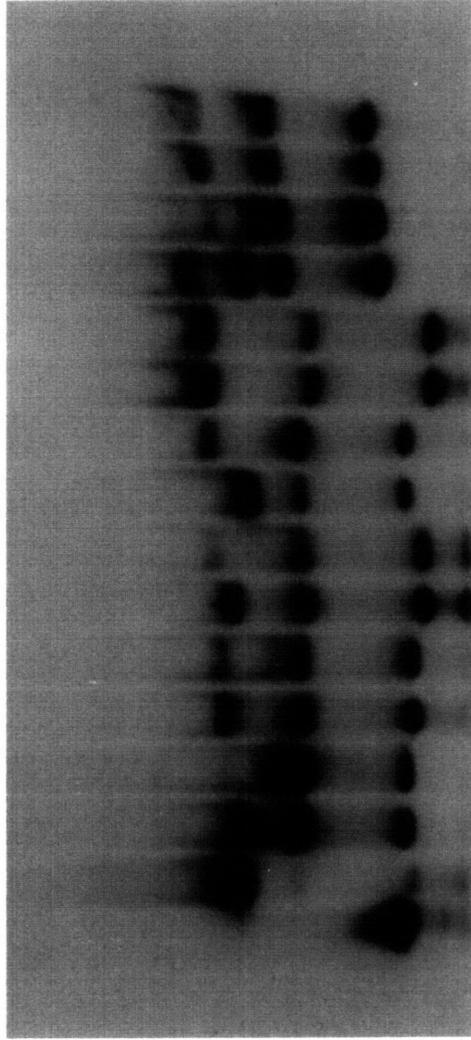
DNA:	O1	L2	+1	-1	+3	-3	+10	+11
Protein:	none WT Arc	WT Arc SL35	WT Arc SL35	WT Arc SL35	WT Arc SL35	WT Arc SL35	WT Arc SL35	WT Arc SL35

Figure 9. Wild-type and variant operator DNA fragments used for binding assays. The six base-pair half-site sequences contacted by each Arc dimer in the cocrystal are boxed. In operators with two half-sites, the number of base-pairs separating the centers of the two hexamers is indicated.

O1
ATCGAT GATAGA AGCAC TCTACT ATCG   Wild-Type
TAGCTA CTATCT TCGTG AGATGA TAGC   Operator
11 bp

L1
TGACAT GATAGA AGCACGT
ACTGTA CTATCT TCGTGCA

L2
TGACAT GATAGA AGCACGTGCCTCGATTC
ACTGTA CTATCT TCGTGCACGGAGCTAAG

Half-Site Operators

-3
ACTGAC GATAGA AC TCTACT GTTT
TGACTG CTATCT TG AGATGA CAAA
8 bp

-1
ATCGAT AGTAGA AGAC TCTACT ATCG
TAGCTA TCATCT TCTG AGATGA TAGC
10 bp

+1
ATCGAT GATAGA AGCACA TCTACT ATCG
TAGCTA CTATCT TCGTGT AGATGA TAGC
12 bp

+3
ATCGAT GATAGA AGCACAGC TCTACT ATCG
TAGCTA CTATCT TCGTGTCG AGATGA TAGC
14 bp

+10
ATCGAT GATAGA AGCACGGTAACGCAC TCTACT ATCG
TAGCTA CTATCT TCGTGCCATTGCGTG AGATGA TAGC
21 bp

+11
ATCGAT GATAGA AAGCACGGTAACGCAC TCTACT ATCG
TAGCTA CTATCT TTCGTGCCATTGCGTG AGATGA TAGC
22 bp

Altered Spacing Operators

Figure 10. Plasmid and chromosomal fusions of the P_{ant} promoter to reporter genes (*rpsl+* on pSA600 and *cat* on the chromosome) used to assay Arc activity in strain UA2F. P_{ant} can be repressed by Arc which is expressed from pSA600.

Chapter 3

Dual Regulation of Distinct Steps in Transcription Explains a Novel Repressor to Activator Switch

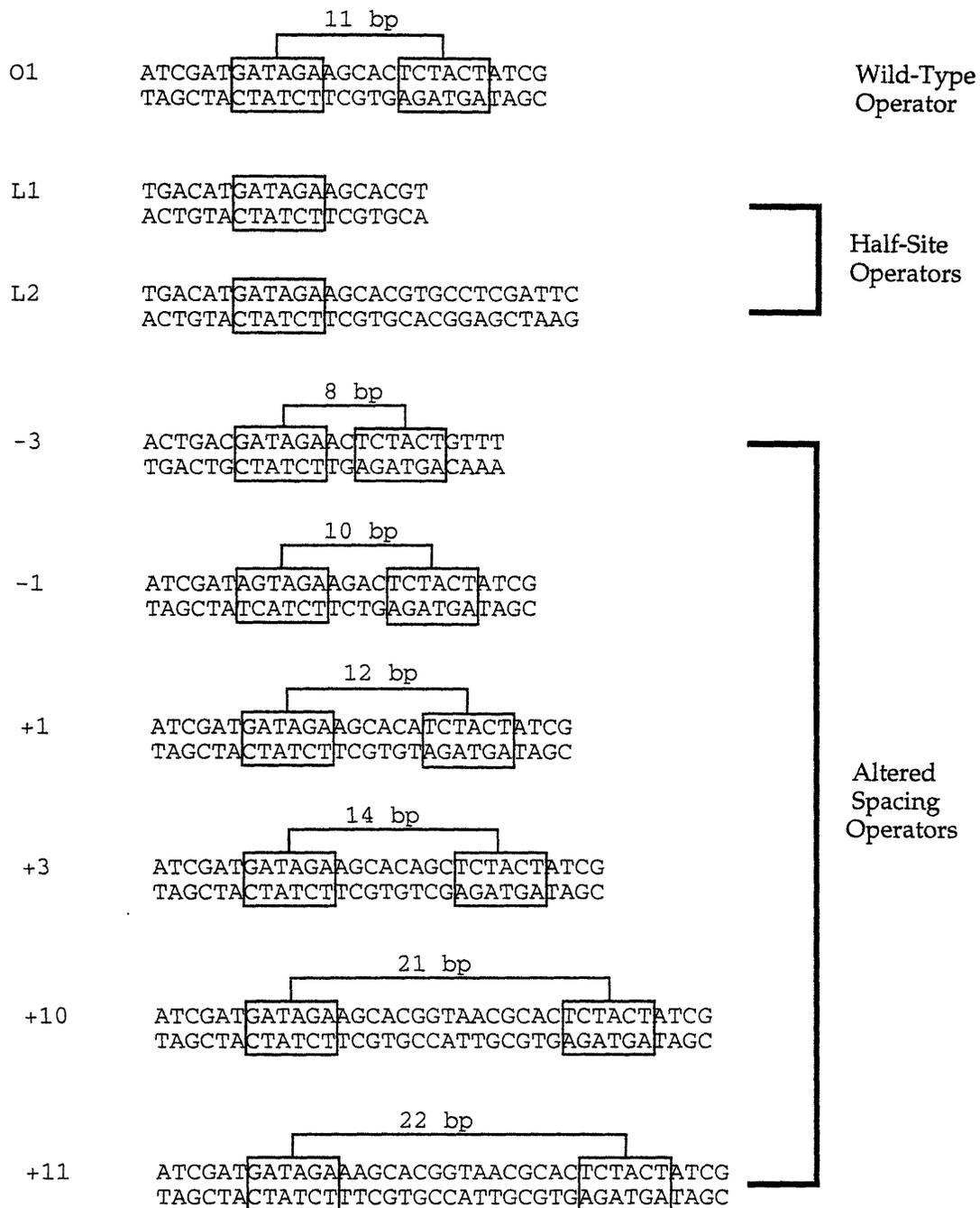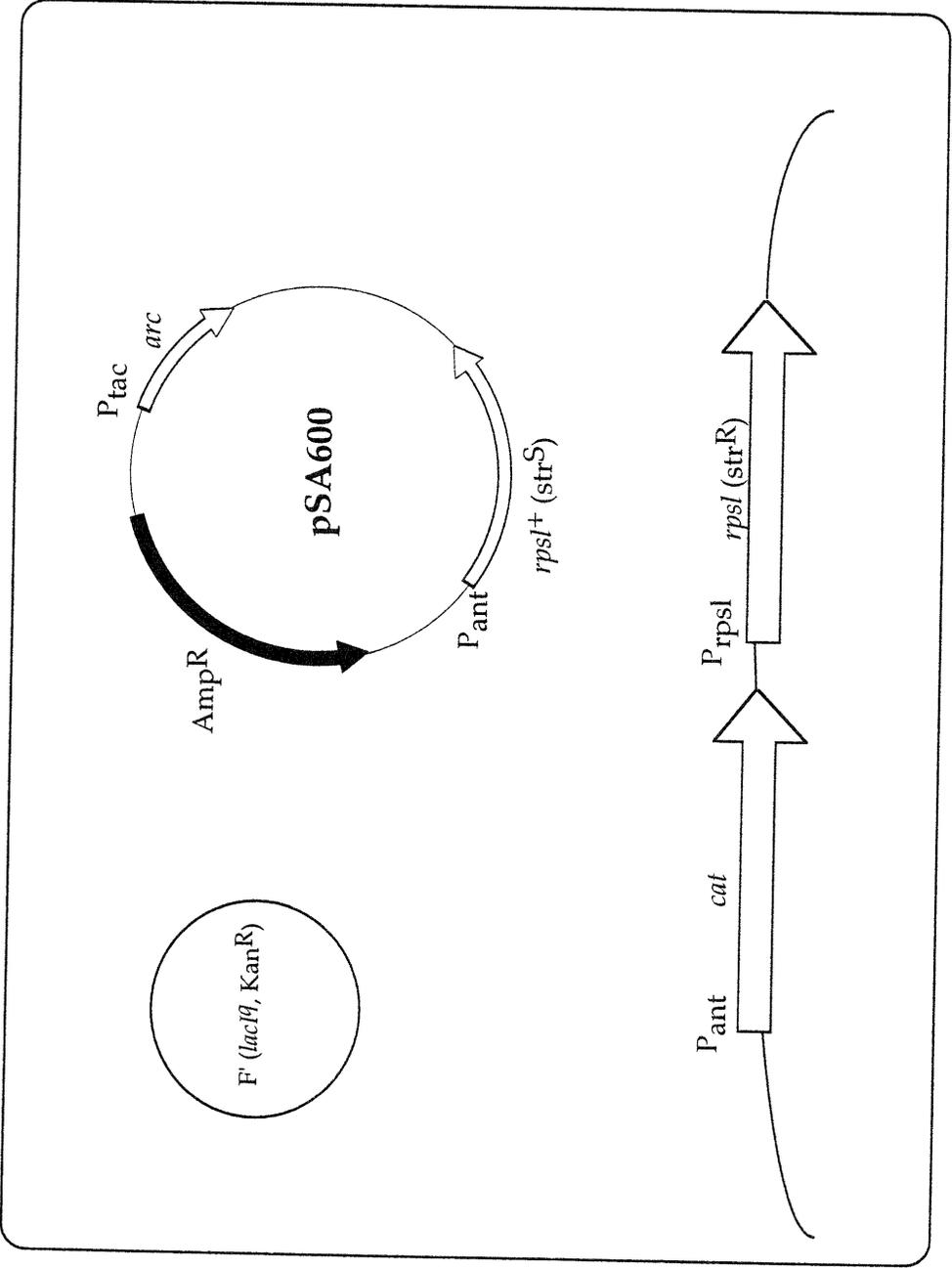Certain versatile transcription factors can either activate or repress transcription depending on the positions of their binding sites with respect to the promoter (1). Here we show that P22 Arc, bound at the same operator position, slows the rate at which RNA polymerase forms open complexes but also accelerates the rate at which the enzyme clears the promoter. These dual activities permit Arc to act as an activator or as a repressor depending upon which step in transcription initiation is rate limiting. In the experiments reported here, the rate-limiting step is determined by the -8 position in the promoter.

Transcription initiation by *E. coli* RNA polymerase is a multistep process involving formation of a competitor-sensitive closed-promoter complex, isomerization to a competitor-resistant open-promoter complex, initiation of RNA synthesis, abortive recycling, and promoter clearance and commitment to elongation after synthesis of approximately eight nucleotides (2). Any of these steps could potentially be accelerated or slowed by transcription factors. In addition, as we show here, variables which change the rate-limiting step can determine whether regulation by a transcription factor is positive or negative.

Using the model system diagrammed in Fig. 1A, we found that a P22 Arc variant (Arc-SL35) could repress transcription *in vitro* from one P_{ant} promoter variant and activate transcription from another (Fig. 1B). The two promoters each contain a single binding site for an Arc dimer. They differ only at one position in the -10 region and are named by whether this position is a consensus (C) or a non-consensus (NC) base for σ^{70} promoters. The NC promoter is repressed and the C promoter is activated by Arc-SL35. The Arc-SL35 variant was used because it prevents non-specific of binding multiple Arc dimers (3). The extent of transcriptional repression or activation is modest (about two-fold in each case) but is very reproducible. The promoter mutation is outside of the operator binding site and does not affect the binding of Arc-SL35 to the template as assayed by the protein concentrations required for half-maximal binding (not shown) or by the pattern of the DNAse I footprint (Fig. 1C). In addition, the DNAse I footprint of the RNA polymerase open complex is identical on the two promoters (Fig. 1C). Since we did not observe any obvious differences in the DNA interactions of Arc-SL35 or RNA polymerase, we reasoned that the ability of Arc-SL35 to repress the NC promoter and activate the C promoter might indicate that the protein is capable of regulating different steps in the initiation process.

92

Arc-SL35 slows formation of competitor-resistant open-promoter complexes on both the NC and the C promoters about 4-fold (Fig. 2; Table 1). The decreased rate of open-complex formation on the NC promoter is consistent with the repression by Arc-SL35 observed in transcription assays with this promoter. In contrast, the slowing of open-complex formation on the C promoter would also be expected to repress transcription, whereas activation is observed. Hence, in the presence of Arc-SL35, formation of open complexes on the C promoter must not be rate limiting, and the protein must be capable of activating this promoter by influencing another step in transcription initiation.

Footprinting and run-off transcription experiments demonstrate that Arc has a previously undetected ability to accelerate promoter clearance (Fig. 3). Arc-SL35 increases the rate of promoter clearance from both the NC and the C promoters by a factor of two to three-fold in both assays (Table 1). In these experiments, open complexes were formed, heparin was added as a competitor to prevent reinitiation, NTPs were added to initiate transcription, and the rate of promoter clearance was determined by measuring either the disappearance of the open-complex DNAse I footprint or the accumulation of completed transcripts as a function of time. An example of a footprinting experiment for the C promoter is shown in Fig. 3A. In this experiment, Arc-SL35 was preincubated with the DNA before the addition of RNA polymerase. The Arc-SL35 footprint can be seen in the absence of RNA polymerase (lane 2) and following promoter clearance (lanes 10-12). At time 0 (lanes 8 and 9), the RNA polymerase footprint completely overlaps that of Arc, making it difficult to know whether both proteins are bound simultaneously to the promoter (see below). Controls show that premixing heparin and template DNA prevents initiation by RNA polymerase (lanes 13-14), and that NTP-independent dissociation of RNA polymerase from the promoter is not significant during the time course of the reaction (lanes 3, 4 & 8, 9). In the absence of Arc, the RNA polymerase open-complex footprint is still visible 6 min after initiation (lane 7). In the presence of Arc-SL35, the open-complex footprint is significantly diminished 1 min after initiation (lane 10).

The ability of Arc-SL35 to stimulate promoter clearance is shared by the wild-type protein but not by a mutant (Arc-RA13) which cannot bind operator DNA (4). Figure 3B shows run-off transcription clearance assays at a single time-point with these different proteins. Unlike the experiment discussed above, the Arc variants were added after the formation of open complexes. Wild-type Arc and Arc-SL35 accelerate clearance on the C and the NC promoters, showing that the order of addition of Arc to the reactions is not

critical. In contrast, the binding-defective RA13 mutant has no effect on promoter clearance on either promoter, suggesting that Arc must be bound to the operator to accelerate promoter clearance. The same conclusion is supported by the finding that Arc-SL35 does not affect clearance from a promoter containing no operator sites (5).

How does Arc stimulate promoter clearance by RNA polymerase? The simplest model is that Arc binds to the operator in close apposition to RNA polymerase bound to the promoter (6) and affects promoter clearance either through direct contacts or indirect effects mediated via DNA structure. For example, by introducing unfavorable electrostatic or steric contacts, Arc might weaken the binding of RNA polymerase to the DNA and thus lower the transition-state energy required for promoter clearance. This model is consistent with the experimental results presented here and is also supported by the finding that ^{35}S-labeled Arc co-migrates with open complexes in DNA mobility shift experiments (5). Moreover, abortive initiation experiments performed by Liao and McClure indicate that Arc slows open-complex formation by retarding isomerization rather than blocking closed-complex formation (7). If Arc binds close to bound RNA polymerase, then some of the same interactions could be responsible for slowing isomerization and for stimulating promoter clearance.

The rates of open-complex formation and promoter clearance shown in Table 1 suggest the following explanations for the observed repression of the NC promoter and activation of the C promoter by Arc-SL35 in $vitro$. Transcription initiation from the NC promoter is limited by the rate of promoter clearance in the absence of Arc-SL35 ($\sim$0.27 min^{-1}) and by the rate of open-complex formation in the presence of Arc-SL35 ($\sim$0.13 min^{-1}). This reduction in the rate of approximately two-fold is similar to the magnitude of repression of the NC promoter by Arc-SL35 (Figure 1B). Transcription from the C promoter is limited by the rate of promoter clearance in both the absence ($\sim$0.08 min^{-1}) and presence of Arc-SL35 ($\sim$0.25 min^{-1}). This increase in the rate is consistent with activation of the C promoter by Arc-SL35 (Figure 1B).

The opposite regulatory effects of Arc-SL35 on the NC and the C promoters under the conditions used in $vitro$ fortuitously revealed the dual activities of Arc. However, in $vivo$, both promoters were found to be repressed by Arc-SL35 (8). This would be expected if clearance from the C promoter is not rate-limiting in the cell in the presence of Arc-SL35. Changes in the rate-limiting step in $vivo$ could be caused by many different

conditions, including the concentration of free RNA polymerase, the level of DNA supercoiling or different ionic strength.

Are the dual activities of Arc on different steps in transcription initiation important for its normal role *in vivo* where it represses transcription from the P_{ant} promoter during late lytic growth of bacteriophage P22 (9)? The wild-type P_{ant} promoter contains two adjacent binding sites for Arc dimers, one located in the same -35 proximal position shown in Fig. 1A, and one located one turn of the DNA helix closer to the -10 region. Arc binds sequentially and cooperatively to these two sites (10), and occupancy of both sites is required for maximal repression (5). However, the -10 proximal *arc* site, unlike the -35 proximal site, overlaps base-specific and phosphate contacts made by RNA polymerase (11), suggesting that RNA polymerase may block binding of Arc to that site. Because Arc must establish repression amidst active transcription from P_{ant}, binding of Arc to the -35 proximal site could help clear the P_{ant} promoter while slowing formation of another open complex. This would free the -10 proximal site to allow cooperative binding of a second Arc dimer to the wild-type operator, thereby blocking further transcription from P_{ant}. This model suggests that the ability of Arc to stimulate promoter clearance may contribute to its efficiency as a negative regulator.

The regulatory protein MerR, like Arc, binds to an operator positioned between the -35 and -10 promoter elements. In the absence of mercury, MerR represses transcription by retarding isomerization to the open complex. In the presence of mercury, MerR becomes an activator and accelerates open-complex formation by a mechanism that apparently involves a conformational change in the protein-DNA complex (12). Although the steps at which Arc and MerR affect transcription are not identical, the similarities in the positions of the operators and the ability to both repress and activate transcription when bound to a single site suggest that proteins which bind between the -35 and -10 elements may be especially well suited to have dual functions in regulating transcription initiation.

In the cell, many variables might affect the rate-limiting step of a promoter, including temperature, DNA superhelicity, ionic strength, nutrient conditions, the concentration of free RNA polymerase, mutations, and the binding of transcription factors (13). Based upon the results presented here, any variable that changes the rate-limiting step in transcription of a promoter could potentially transform a repressor into an activator or *vice versa*.

Notes and References

1. S. T. Woody, R. S.-C. Fong, G. N. Gussin, *J. Mol. Biol.* **229**, 37 (1993); J. Collado-Vides, B. Magasanik, J. D. Gralla, *Microbiol. Rev.* **55**, 371 (1991).

2. W. R. McClure, *Annu. Rev. Biochem.* **54**, 171 (1985); D. C. Straney and D. M. Crothers, *Cell* **43**, 449 (1985).

3. T. L. Smith and R. T. Sauer, *J. Mol. Biol.* **249**, 729 (1995).

4. B. M. Brown, M. E. Milla, T. L. Smith, R. T. Sauer, *Nature Struct. Biol.* **1**, 164 (1994).

5. T. L. Smith, unpublished results.

6. The RNA polymerase footprint completely overlapped the Arc footprint with all reagents tested (DNAse I, hydroxyl radicals, 1,10-phenanthroline-copper, and dimethyl-sulfate) making it difficult to show directly that the two proteins are bound to the DNA simultaneously.

7. S.-M. Liao, Ph. D. thesis, Carnegie-Mellon University (1988).

8. The NC and C promoters gave similar levels of β-galactosidase *in vivo* when fused to a *lacZ* reporter gene, suggesting that the promoter strength differences between the promoters observed *in vitro* are minimized *in vivo*. Furthermore, both the NC and C promoter-*lacZ* constructs were repressed in cells when paired with a construct expressing Arc-SL35 under the conditions tested.

9. M. M. Susskind and P. Youderian, in *Lambda II*, R. W. Hendrix, J. W. Roberts, F. W. Stahl, R. A. Weisberg, Eds. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1983), pp. 347-363.

10. B. M. Brown and R. T. Sauer, *Biochemistry* **32**, 1354 (1993).

11. U. Siebenlist, R. B. Simpson, W. Gilbert, *Cell* **20**, 269 (1980).

12. A. Z. Ansari, J. E. Bradner, T. V. O'Halloran, *Nature* **374**, 371 (1995); A. O. Summers, *J. Bacteriol.* **174**, 3097 (1992); ; A. Z. Ansari, M. L. Chael, T. V. O'Halloran, *Nature* **355**, 87 (1992); A. Heltzel, I. W. Lee, P. A. Totis, A. O. Summers, *Biochemistry* **29**, 9572 (1990); T. V. O'Halloran, B. Frantz, M. K. Shin, D. M Ralston, J. G. Wright, *Cell* **56**, 119 (1989).

13. T. Ellinger, D. Behnke, H. Bujard, J. D. Gralla, *J. Mol. Biol.* **239**, 455 (1994); J. A. Goodrich and R. Tjian, *Cell* **77**, 145 (1994); J. L. Botsford, J. G. Harman, *Microbiol. Rev.* **56**, 100 (1992); S. Adhya, S. Garges, *J. Biol. Chem.* **265**, 10797 (1990); K. Tsung, R. E. Brissette, M. Inouye, *Proc. Natl. Acad. Sci.* **87**, 5940 (1990); G. J. Pruss and K. Drlica, *Cell* **56**, 521 (1989); S. Leirmo, C. Harrison, D. S. Cayley, R. R. Burgess, M. T. Record, Jr., *Biochemistry* **26**, 2095 (1987).

14. Unless noted, all experiments were performed at 37 °C in a buffer containing 30 mM Hepes-KOH (pH 7.5), 100 mM potassium glutamate, 10 mM $MgCl_2$, 1.5 mM $CaCl_2$, 0.1 mM Na_2EDTA, 100 µg/mL bovine serum albumin, 1 mM dithiothreitol, and 0.02% NP-40, and 0.2 nM template DNA (19). RNA polymerase and Arc-SL35 were used at concentrations of 7.5 nM and 75 nM, respectively, unless indicated. In some experiments, wild-type Arc or Arc-RA13 was used instead of Arc-SL35. The final concentrations of NTPs were 167 µM ATP, GTP, and CTP, and 5 µM UTP (at a specific activity of 1.4 x 10^4 cpm/pmol UTP in run-off transcription assays). RNAse inhibitor (Promega; 1 U per reaction) was included in the transcription and open-complex formation reactions. Arc variants containing a C-terminal His_6 tag, and *E. coli* σ^{70} RNA polymerase holoenzyme were purified as described (20. 21). The products of the transcription reactions were analyzed on 12% polyacrylamide, 8 M Urea, 1X TBE gels. Results of all experiments were quantified using a Molecular Dynamics phosphorimager and ImageQuant software.

15. To measure rates of open-complex formation, RNA polymerase (7.5 nM) was added to DNA (0.2 nM), and the reaction was quenched at different times by the addition of heparin to 100 µg/mL and sucrose to 5%. Bound and free DNA fragments were resolved on 4% polyacrylamide, 0.5X TBE gels.

16. For assays of promoter clearance by footprinting, RNA polymerase open complexes were formed on promoter DNA in the presence or absence of Arc-SL35 (75 nM), the reactions were diluted 50-fold into buffer containing 0.1 µg/mL heparin with or without Arc-SL35 (75 nM), and NTPs were added to initiate the reaction. At different times,

aliquots were removed, added to DNAse I (final concentration of 18 ng/mL) for 60 s, and cleavage was quenched by addition of an equal volume of 2.5 M ammonium acetate, 20 mM Na$_2$EDTA, and 10 µg/mL salmon sperm DNA. Digestion products were analyzed on 6% polyacrylamide, 8 M Urea, 1X TBE gels, and loss of the RNA polymerase footprint was quantified.

17. For assays of promoter clearance by a single round of run-off transcription, preformed open complexes were diluted 1.6-fold into buffer containing heparin (4 µg/mL final concentration) with or without Arc-SL35 (75 nM final concentration). NTPs were then added to initiate transcription. Aliquots were removed at different times and quenched with heparin, KCl, and Na$_2$EDTA at final concentrations of 100 µg/mL, 0.35 M, and 25 mM, respectively. In the single timepoint assays shown in Fig. 3B, the concentrations of the Arc variants were increased to 300 nM.

18. Promoter DNA was incubated with or without Arc-SL35 in the presence of NTPs. Transcription reactions were initiated by the addition of RNA polymerase and were quenched after 2 min as described (17).

19. The NC and C promoters were cloned as EcoRI-BstEII cassettes into the backbone of pSA660 (5), a derivative of pSA600 (20). Linear templates were generated by the polymerase chain reaction using a pair of primers with the 5' ends of the primers located at -87 and at +85 with respect to the start site of transcription. The -87 primer was end-labeled with ^{32}P using T4 polynucleotide kinase. To allow normalization of the templates in the reactions, the same stock of labeled primer was used in PCR reactions to make both the NC and the C fragments. The 173 bp PCR products were purified by gel electrophoresis.

20. M. E. Milla, B. M. Brown, R. T. Sauer, *Protein Sci.* **2**, 2198 (1993).

21. D. A. Hager, D. J. Jin, R. R. Burgess, *Biochemistry* **29**, 7890 (1990).

22. B. E. Raumann, M. A. Rould, C. O. Pabo, R. T. Sauer, *Nature* **367**, 754 (1994)

	intrinsic rates		rates + Arc-SL35	
	open cplx (min⁻¹)	clearance (min⁻¹)	open cplx (min⁻¹)	clearance (min⁻¹)
NC promoter	2.4 (±0.2)	0.27 (±0.07)[a] 0.35 (±0.02)[b]	0.13 (±0.02)	0.84 (±0.29)[a] 0.82 (±0.27)[b]
C promoter	≥16	0.08 (±0.01)[a] 0.10 (±0.02)[b]	3.5 (±0.70)	0.25 (±0.05)[a] 0.22 (±0.02)[b]

Table 1. Apparent rates (average ± sd, n≥3) of open-complex formation and promoter clearance. For each promoter, the rate-limiting step in transcription initiation in the presence or absence of Arc-SL35 is boxed. Rates of open-complex formation were determined using 7.5 nM RNA polymerase (14, 15). Rates of promoter clearance were determined from footprinting experiments[a] (14, 16) or from single round run-off transcription experiments[b] (14, 17).

Figure 1. A) Sequence of the NC variant of the P_{ant} promoter (the C promoter is identical except at position -8). Base-pairs specifically contacted by the Arc dimer in the cocrystal structure (22) are boxed. The diagram shown below indicates roughly how Arc bound to its half-site and RNA polymerase bound in the open complex should be closely apposed. B) Run-off transcription assays (14, 18). Left, the radiolabeled DNA template (19) and 85 base run-off transcript bands are marked. Right, a plot of the average ($\pm$ sd, n$\geq$3) of the ratio of transcription with Arc-SL35 to transcription without Arc-SL35. C) DNAse I footprints of Arc-SL35 and the RNA polymerase bound to the NC and the C promoters (14, 16).

NC

-35   *arc*
half-site   -10

GAATTCGTATTGACAT GATAGAAGCACTCGCCTATATT CTGGTAACC
CTTAAGCATAACTGTA CTATCTTCGTGAGCGGATATAA AGACCATTGG

C

A
T

RNAP -35 Arc -10

Activation

No Change

Repression

$\dfrac{Tx\,(+)}{Tx\,(0)}$

3

2

1

0

Promoter

C

NC

C

NC

+   0   +   0

Arc-SL35

template DNA

run-off transcript
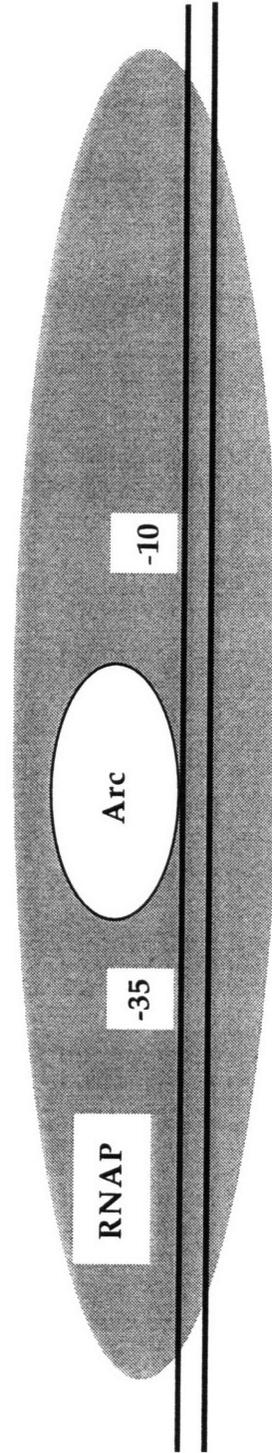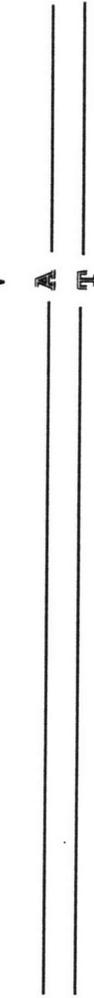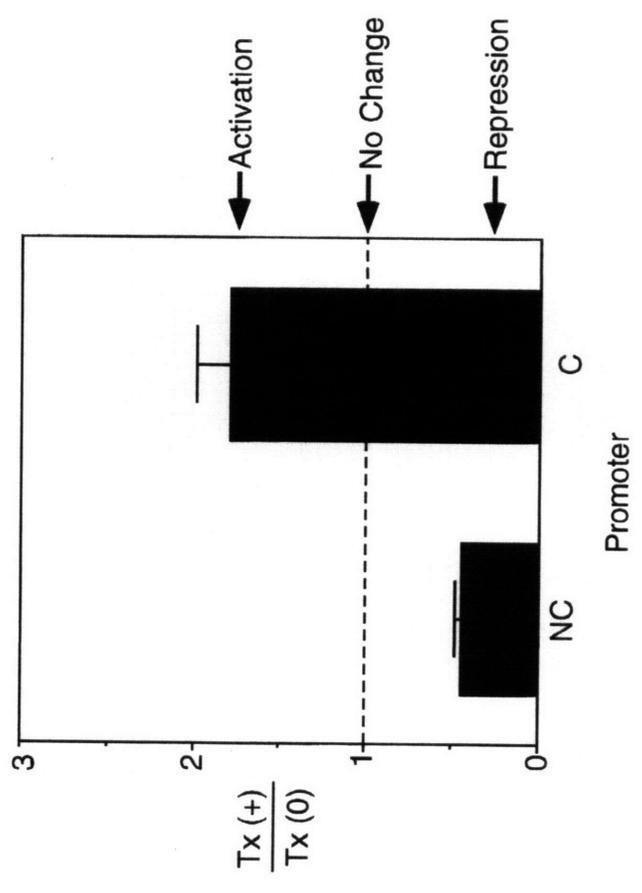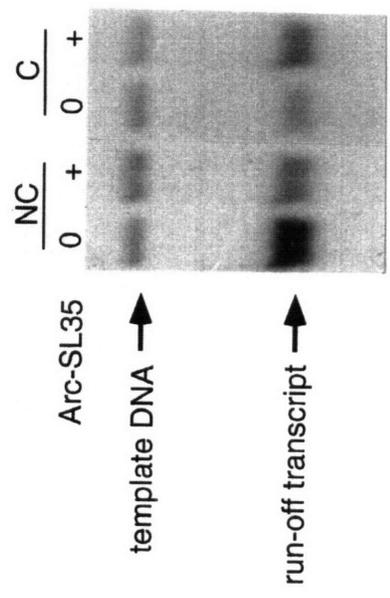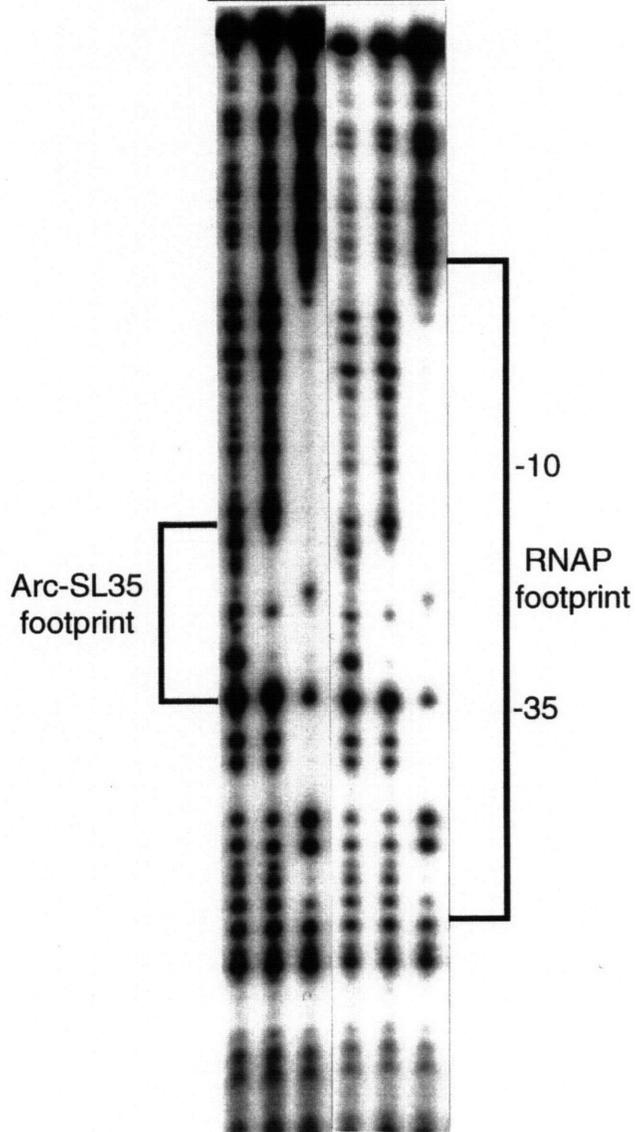
Figure 2. Kinetics of open-complex formation by RNA polymerase with the NC and C promoters (14, 15).

NC

no Arc

+ Arc-SL35

C

no Arc
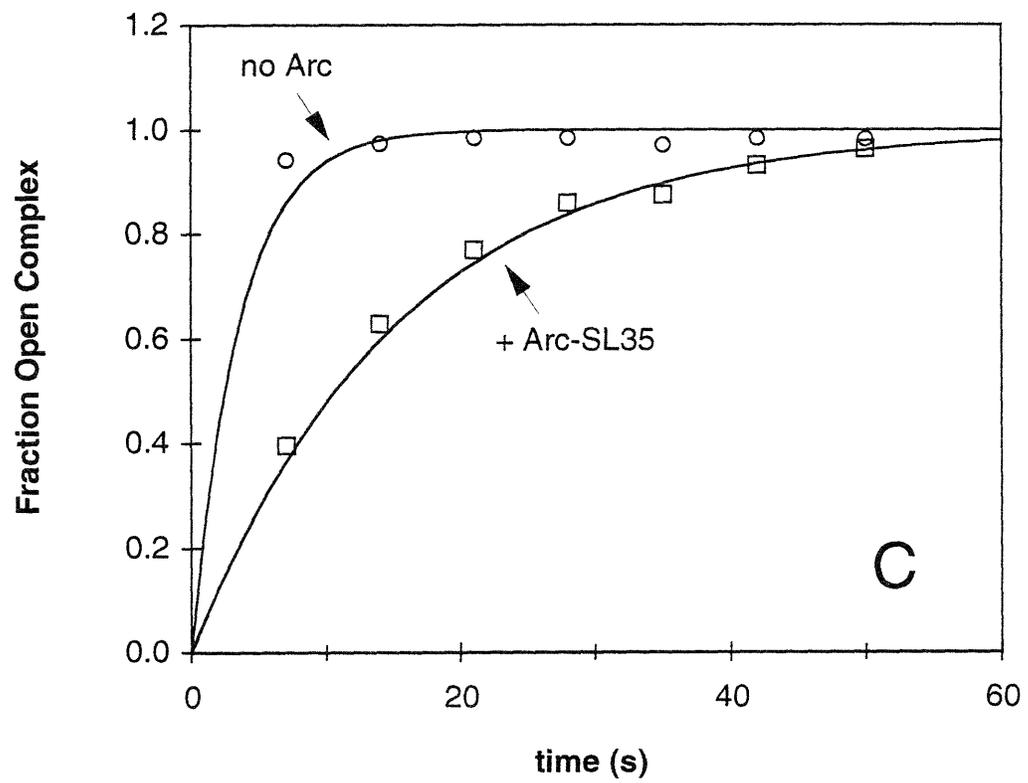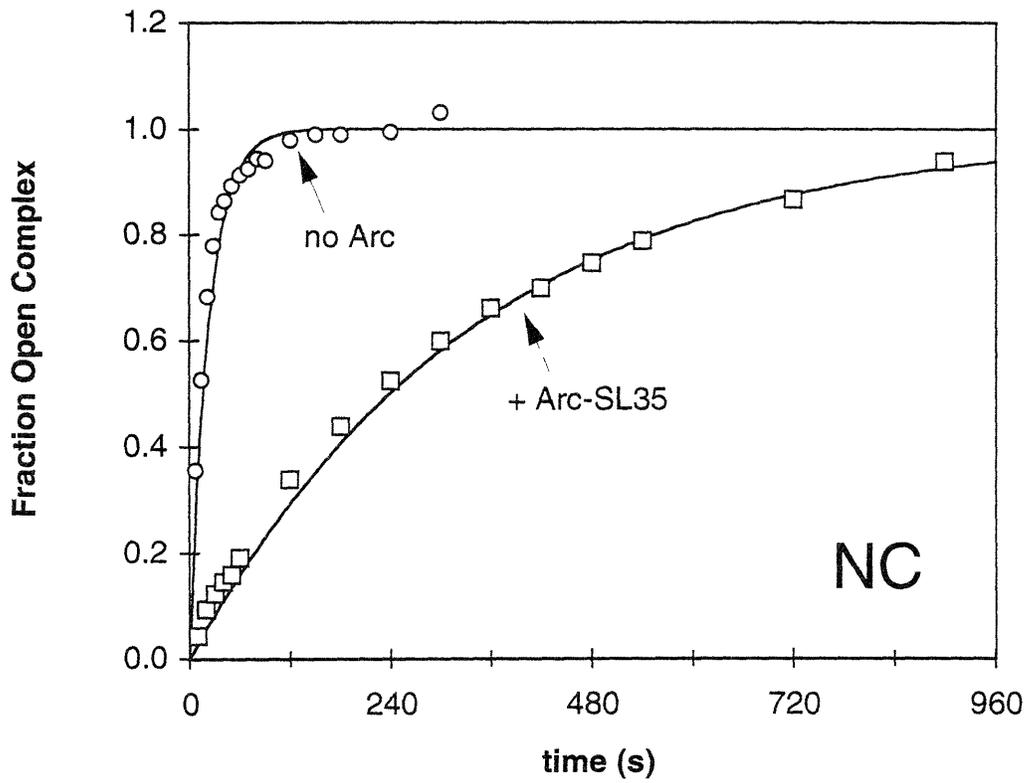
+ Arc-SL35

Figure 3. A) Clearance of RNA polymerase from the C promoter assayed by footprinting (14, 16). Footprints 1, 3, and 6 minutes after the addition of NTPs to preformed open complexes are shown in lanes 5-7 (without Arc-SL35) and lanes 10-12 (with Arc-SL35). Control lanes: (1) No protein with heparin; (2) Arc-SL35 with heparin; (3, 8) open complex without NTPs but with heparin for 10 s; (4, 9) open complex without NTPs but with heparin for 10 min; (13, 14) 1 and 10 min after RNA polymerase was added to premixed heparin and DNA. The clearance rates for this experiment were 0.08 min^{-1} without Arc-SL35 and 0.20 min^{-1} with Arc-SL35. B) Promoter clearance of preformed open complexes assayed by run-off transcription at a single timepoint with Arc-SL35, wild-type Arc, and the operator-binding defective mutant Arc-RA13 (14, 17).

run-off transcript 1 min after addition of NTPs

template DNA

no Arc
Arc-RA13
Arc
Arc-SL35
no Arc

C

no Arc
Arc-RA13
Arc
Arc-SL35
no Arc

NC
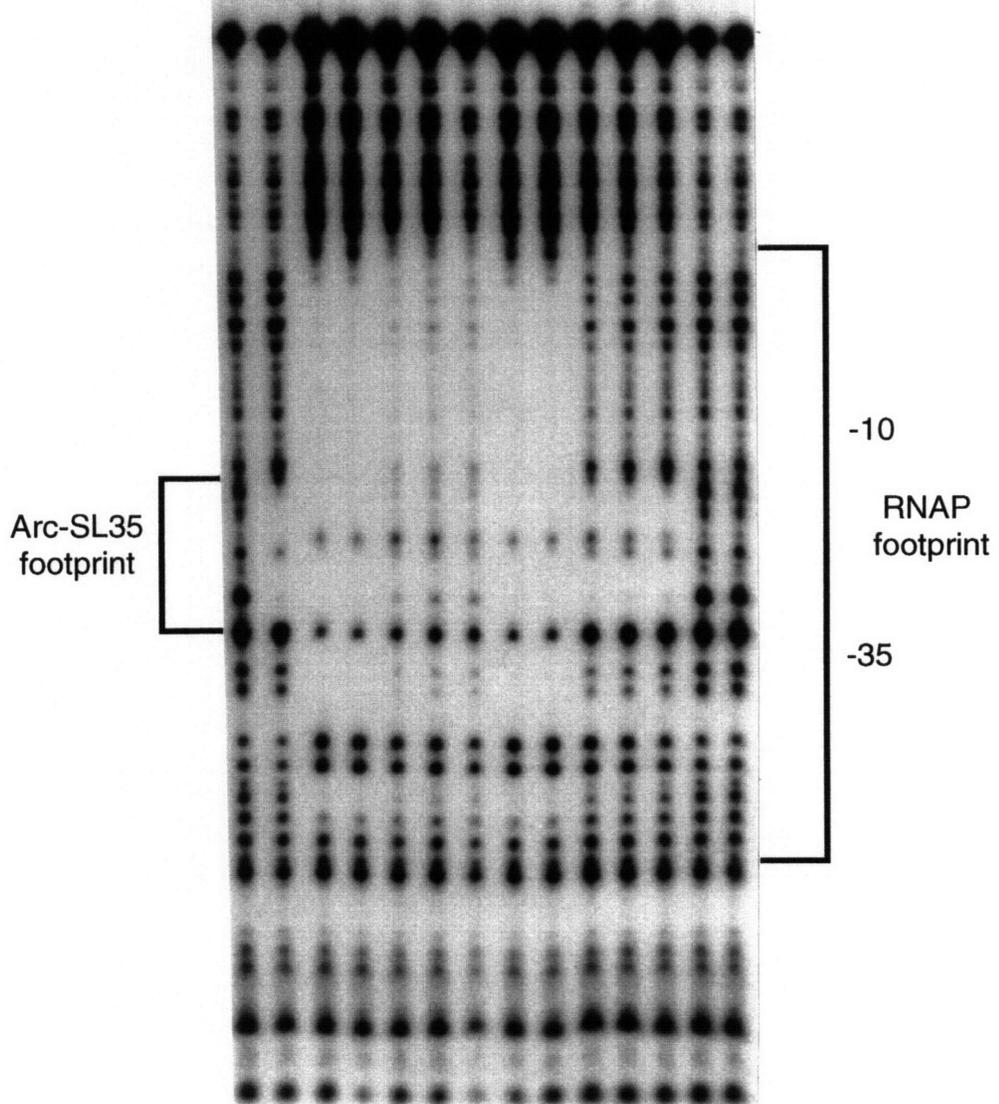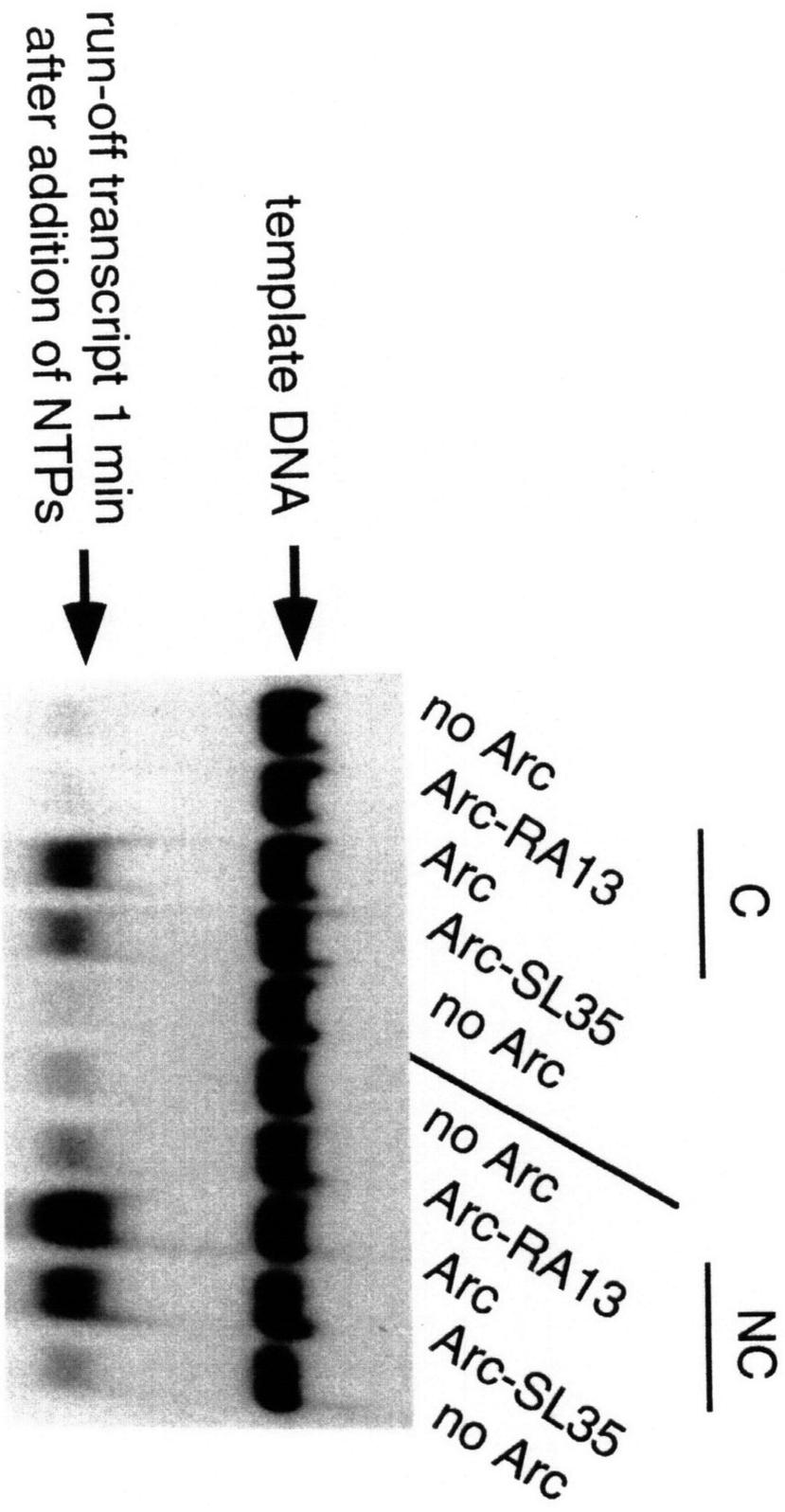
Chapter 4

Role of Operator Subsite Binding in Arc Repressor Function

Introduction

It is relatively common for transcription factors to bind to adjacent DNA regulatory sites near the promoters they control (Carlson & Little, 1993; Collado-Vides *et al.*, 1991; Kim & Little, 1992; Phillips *et al.*, 1989; Ptashne, 1986). Proteins bound at tandem sites of this type have the potential to interact through cooperative protein-protein contacts, which would serve to increase both the affinity and the specificity of DNA binding. In principle, binding to adjacent sites could also permit regulatory proteins to serve distinct functional roles. In this paper, we analyze the functional importance of the binding of Arc dimers to individual subsites in the *arc* operator. During lytic growth of bacteriophage P22, Arc dimers bind to adjacent DNA subsites in the *immunity I* operon (Fig. 1), thereby repressing transcription from the divergent P_{ant} and P_{mnt} promoters (Susskind & Youderian, 1983). The DNA sequences of the two operator subsites and their affinities for Arc are similar but not identical (Brown & Sauer, 1993). Protein-protein interactions between adjacently bound Arc dimers stabilize DNA binding (Brown & Sauer, 1993; Raumann *et al.*, 1994b), and Arc variants with normal subsite binding but diminished cooperativity are poorer repressors *in vivo* than wild-type Arc (Smith & Sauer, 1995). However, cooperativity defective mutants are fully active at high protein concentrations *in vitro,* suggesting that cooperative binding to adjacent subsites is required for optimal function *in vivo* principally to allow saturation of the operator at low Arc concentrations (Smith & Sauer, 1995).

The studies presented here show that Arc dimers bound to single operator subsites can repress transcription of variants of both P_{ant} and P_{mnt}. The efficiency of this repression, however, is subsite and promoter dependent. Occupancy of the subsite proximal to the P_{ant} -35 region and the P_{mnt} -10 region (the left subsite in Fig. 1), results in stronger repression of both promoters than occupancy of the other subsite (the right subsite in Fig. 1). Thus, the left operator subsite appears to play the dominant role in mediating repression, and the right subsite may be needed primarily to allow cooperative stabilization of the dimer bound to the left subsite.

Results

Construction of promoters containing single subsites for Arc binding

The positions of the two subsites of the *arc* operator in relation to the overlapping P_{ant} and P_{mnt} promoters are shown in Fig. 1. Promoter variants with only a single active subsite were constructed by changing base pairs (bp) in the other subsite to eliminate Arc binding (Vershon *et al.*, 1989). In the wild-type context, transcription from the P_{ant} promoter interferes with transcription from the weaker P_{mnt} promoter (Vershon *et al.*, 1987a; Vershon *et al.*, 1987b). To avoid complications caused by this interaction, we also inactivated one promoter or the other with mutations, allowing the effects of Arc binding to single subsites to be independently assayed for the variants of P_{ant} and P_{mnt}. The promoter variants used for these studies are diagrammed in Fig. 2 and the associated sequence changes are listed in Table 1. We refer to the single-subsite variants using the promoter name and a number indicating whether the active subsite is proximal to the -35 region or the -10 region (e.g., $P_{mnt}/35$). For the $P_{ant}/35$ variants, the subsites are further designated A (wild-type sequence) or B (-35 proximal subsite has the sequence of the wild-type -10 proximal site). The latter construct controls for the minor sequence differences between subsites. There are also A and B variants of the $P_{ant}/10$ constructs, but in this case the promoters differ at a position (-30) that is not important for Arc binding but has been shown to affect recognition by RNA polymerase (McClure, 1985). Promoters with two active subsites are designated $P_{ant}/both$ and $P_{mnt}/both$, whereas those with no active subsites inactivated are designated $P_{ant}/none$ and $P_{mnt}/none$.

On each promoter construct, only one RNA polymerase open-complex footprint was seen at the expected position (not shown) and only one major transcript of the expected size was observed in run-off transcription assays (see below). In addition, footprinting with saturating concentrations of the cooperativity defective mutant, Arc-SL35, showed that each construct bound Arc only at the expected subsites (Figure 3). Arc-SL35 was used throughout these studies intsead of wild-type Arc because this mutant did not stabilize nonspecific binding of additional dimers through cooperative protein-protein interactions (Smith & Sauer, 1995).

111

Run-off transcription experiments

Run-off transcription experiments with RNA polymerase in excess of template
DNA were performed in the absence of repressor or in the presence of a concentration of
Arc-SL35 sufficient to saturate the subsite in the footprinting assay. Figure 4A shows an
example of a run-off transcription experiment with the P_{ant}/both promoter variant. In the
absence of Arc, the major bands are the run-off transcript of approximately 85 nucleotides
(nt) and the ^{32}P-labeled DNA template. In addition, an end-to-end transcript was seen
when no promoter was present (not shown) or when the promoter was strongly repressed
(+ Arc-SL35 lane in Fig. 4A). In the presence of Arc-SL35, transcription from P_{ant}/both
was significantly diminished. The results of run-off transcription experiments with each
promoter variant are shown in Fig. 4B.

Some promoter variants containing only a single subsite can be significantly
repressed by the binding of an Arc-SL35 dimer (Fig. 4B). For the P_{ant} promoter variants,
Arc-SL35 mediated repression was about 80-90% efficient under the conditions of this
assay when both operator subsites were present, about 30-50% efficient with an active
subsite proximal to the -35 region, and only 0-20% efficient when the active subsite was
proximal to the -10 region (Fig. 4B). Differences in the overall promoter strengths of these
variants were less than two-fold (Table 1). Thus, the main variables affecting repression
seem to be the number and positions of active subsites. Changing the sequence of the -35
proximal active subsite from wild type (P_{ant}/35A) to that of the normal -10 proximal
subsite (P_{ant}/35B) did not affect the ability of that position to mediate repression, indicating
that the minor sequence and affinity differences between subsites were not important
determinants of repression efficiency. Altering the promoter context by changing position
-30 from a consensus (P_{ant}/10A) to a nonconsensus bp (P_{ant}/10B) resulted in slightly
better repression, indicating that promoter context may affect regulation by Arc (see
Discussion). Fig. 4C shows run-off transcription as a function of Arc-SL35 concentration
for the P_{ant}/both, P_{ant}/35A, P_{ant}/10A, and P_{ant}/none promoters. At very high
concentrations of Arc-SL35, repression of the promoter with the active -35 proximal
subsite approaches that of the promoter with both active subsites. The promoter variant
with the -10 proximal subsite, however, shows almost no repression even at the highest
repressor concentrations tested. Thus, the subsite next to the -35 promoter element of P_{ant}
appears to play a major role in repression while the subsite next to the -10 promoter element
is less important.

Repression efficiency of the P_{mnt} variants by Arc also changed depending on the position of the active subsite (Fig. 4B). In this case, however, the variant with the active subsite proximal to the -10 region (P_{mnt}/10) was repressed to the same extent ($\approx 75\%$) as the variant with two active subsites (P_{mnt}/both). The promoter with the single active subsite proximal to the -35 region (P_{mnt}/35) was repressed only at about 25% efficiency. P_{mnt}/both and P_{mnt}/10 were slightly weaker promoters than the P_{ant} variants described above, while both P_{mnt}/35 and P_{mnt}/none had similar promoter strengths as the P_{ant} variants (Table 1). Thus, repression efficiency is also inversely correlated with promoter strength for these P_{mnt} promoter variants. However, a clear conclusion is that the -10 proximal subsite is capable of repressing a promoter with strength equal to that of the promoter with both subsites present.

DNA mobility shift assays for open-complex formation

As an independent test of the effects of Arc-SL35 binding to different subsites, DNA mobility shift assays were used to monitor the extent of open-complex formation by RNA polymerase in the absence and the presence of Arc-SL35. Fig. 5A shows an example of an assay for open-complex formation with the P_{ant}/both variant at a single 2 min timepoint in which Arc-SL35 effectively prevented RNA polymerase open-complex formation. A summary of similar single-point assays (2 min, 75 nM Arc-SL35) for the extent of open-complex formation with each promoter construct in the presence and absence of Arc-SL35 is shown in Fig. 5B. For the P_{ant} variants in the single-point assays, Arc-SL35 repressed open-complex formation by RNA polymerase most efficiently when both subsites were filled (≈ 90-95%), next most efficiently when the subsite proximal to the -35 region was occupied (≈ 60-75%), and least efficiently when the subsite proximal to the -10 region was bound (≈ 0-30%). Fig. 5C shows an assay of open-complex formation at 2 min in the presence of increasing concentrations of Arc-SL35 for the P_{ant}/both, P_{ant}/35A, P_{ant}/10A, and P_{ant}/none promoters. Again, it is clear that an Arc-SL35 dimer bound just to the -35 proximal subsite is an efficient repressor whereas a dimer bound just to the -10 proximal subsite is a very poor repressor.

For the P_{mnt} variants in single-point assays, repression of open-complex formation by Arc was most efficient when both subsites were present ($\approx 95\%$), slightly less efficient when only the -10 proximal subsite was functional ($\approx 80\%$), and only marginally efficient when just the -35 proximal subsite was active ($\approx 25\%$). Similar results for the P_{ant} and

113

P_{mnt} variants were obtained when open-complex formation was assayed at an earlier timepoint (20 s, 75 nM Arc-SL35) at which open complex formation was less complete (not shown). Repression ratios calculated from the open-complex formation assays correlate well ($r = 0.967$) with those calculated from the run-off transcription assays, suggesting that repression by Arc-SL35 dimers bound to one or two subsites largely reflects its activity in slowing the rate of open-complex formation by RNA polymerase.

Assays for dimer repression in vivo

Plasmids containing the P_{ant}/both, P_{ant}/35A, P_{ant}/10A, and P_{ant}/none promoters fused to the *lacZ* gene were constructed to allow assays of repression *in vivo*. As shown in Table 2, Arc repressed the P_{ant}/both-*lacZ* fusion approximately 75% and the P_{ant}/35A-*lacZ* fusion approximately 50%. No significant repression was observed for the P_{ant}/10A-*lacZ* and P_{ant}/none-*lacZ* constructs. Controls showed that Arc-RA13, a mutant defective in operator binding (Brown *et al.*, 1994), failed to repress any of the constructs. These repression assays *in vivo* are generally consistent with the results obtained *in vitro* with the same promoter variants. The promoter strengths of the P_{ant} variants were similar both *in vitro* and *in vivo*, and an Arc dimer bound to the P_{ant} -35 proximal subsite was more efficient at repressing transcription *in vitro* and *in vivo* than a dimer bound to the -10 proximal subsite.

Discussion

Arc is a member of the ribbon-helix-helix family of DNA binding proteins. Arc and the other well-characterized members of this family, the Mnt and MetJ repressors, use protein-protein interactions between DNA-bound dimers to stabilize binding to tandemly arrayed operator subsites (Brown & Sauer, 1993; He *et al.*, 1992; Waldburger & Sauer, 1995). Are tandem binding sites required for regulation by proteins in this family? Here, we have used a set of designed promoter variants to test whether the reiterated subsites in the Arc system are required for repression. We found that an Arc dimer bound to a single subsite could repress transcription of certain P_{ant} and P_{mnt} variants *in vitro* and *in vivo*. Thus, tandem binding to multiple *arc* subsites is not essential for regulation.

Arc dimers bound to the left or right operator subsites have significantly different repressor activities. In particular, for the P_{ant} variants, a dimer bound in the P_{ant} -35 proximal position slowed open-complex formation and repressed transcription reasonably efficiently, whereas a dimer bound to the P_{ant} -10 proximal site was much less effective. In the case of the P_{mnt} variants, a dimer bound proximal to the -10 region was more effective than a dimer bound near the -35 region. In the wild-type context, the subsite proximal to the -10 region of P_{mnt} is the same as the subsite proximal to the -35 region of P_{ant} (Fig. 1). Hence, the left *arc* subsite shown in Figure 1 appears to be the primary mediator of Arc repression. Since cooperative DNA binding of two Arc dimers is important for normal repression efficiency *in vivo* (Smith & Sauer, 1995), occupancy of the right subsite may be required mainly to facilitate cooperative stabilization of the dimer bound to the left subsite. However, since transcriptional repression of P_{ant} is slightly more efficient *in vitro* and *in vivo* when both subsites are filled than when just the left subsite is filled by Arc-SL35, (Fig. 4C and Table 2), noncooperative occupancy of the right subsite may also have a small additive or synergistic effect on repression.

Since the *arc* operator is positioned between the -35 and -10 promoter elements, it is possible that changing the *arc* subsite sequences could potentially alter transcription initiation by RNA polymerase in a way that would complicate the results. A change from a consensus (P_{ant}/10A) to a nonconsensus (P_{ant}/10B) bp at position -30, which is known to be important for initiation by RNA polymerase, did indeed affect repression by Arc (Fig. 4B and Fig. 5B). Therefore, we cannot completely rule out a similar effect of the operator changes although the mutations are outside of the -35 and -10 regions. However, similar results were obtained *in vitro* (Table 1, Fig 4B, and Fig. 5B) and *in vivo* (Table 2) with the P_{ant} variants, and the titration experiments with the Arc-SL35 *in vitro* (Fig. 4C and Fig. 5C) further indicate that the different repression efficiencies seen with the P_{ant}/10A and the P_{ant}/35A promoters (the variants with the fewest changes from the P_{ant}/both promoter) exists even at high concentrations of Arc. Based on these results, it seems likely that the sequence changes made in the operator do not dramatically affect transcription initiation at the P_{ant} variants under the conditions tested *in vitro* or under the possibly different conditions *in vivo*. Thus, the different repression efficiences of the P_{ant} -35 and the -10 proximal subsites is likely to result from the diffferent positions of those subsites with respect to the P_{ant} promoter. The P_{mnt} variants did show larger differences in promoter strength *in vitro* (Table 1) suggesting that initiation by RNA polymerase at a promoter similar to P_{mnt} may be altered by mutations in one of the *arc* subsites. Nonetheless, P_{mnt}/both and P_{mnt}/10 have similar strengths (Table 1), and in these

constructs, the P_{mnt} -10 proximal subsite repressed transcription as well as both subsites (Fig. 4B and Fig. 5B). Thus, the left subsite in Fig. 1 is likely to be sufficient for repression of the wild-type P_{mnt} promoter and probably also plays the primary role in repression of the wild-type P_{ant} promoter.

Transcription initiation by *E. coli* RNA polymerase is a multistep process involving binding of the enzyme to the promoter in an unstable closed complex, isomerization to a stably bound open complex, initiation of transcript formation, and clearance from the promoter (McClure, 1985). The experiments presented in Fig. 4 and Fig. 5 show that repression of a variety of promoter constructs by Arc-SL35 is strongly correlated with its ability to decrease the rate of RNA polymerase open-complex formation on these promoters. Experiments performed by Liao and McClure suggest that wild-type Arc represses the wild-type P_{ant} promoter by slowing the rate at which closed-promoter complexes isomerize to form open complexes (Liao, 1988). We have recently found that Arc can also increase the rate of RNA polymerase clearance from certain P_{ant} promoter variants (Smith & Sauer, 1996). For Arc to slow isomerization or to accelerate promoter clearance likely requires that it bind to the promoter simultaneously with RNA polymerase. The multiple activities of Arc and potential concomitant binding of Arc and RNA polymerase suggest the possibility of interactions between Arc and RNA polymerase, at least at the P_{ant} promoter. Arc dimers bound to the left and right operator subsites might have unequal repression efficiencies because only one Arc subunit is properly positioned to interact with RNA polymerase or because a dimer bound to only one subsite affects access of RNA polymerase to an important region of the promoter DNA.

The assignment of a primary regulatory role to a single subunit of a multiprotein complex has also been proposed for other systems. Bacteriophage λ cI dimers bound to the three operators O_R1, O_R2, and O_R3 have distinct functional roles in regulating two divergent promoters, P_R and P_{RM}. The dimers at O_R1 and O_R2 repress P_R, a dimer at O_R2 activates P_{RM}, and a dimer at O_R3 represses P_{RM} (Meyer *et al.*, 1980; Ptashne, 1986). The location of the operator sites with respect to the promoters is a key determinant of regulation in this system. Further specification of the functional roles of the two monomers in the λ cI dimer bound to O_R2 can also be detected. Activation of the P_{RM} promoter by a dimer bound to O_R2 operator depends on protein-protein interactions between cI and RNA polymerase, and only the promoter proximal monomer of the λ cI dimer is thought to mediate the activation (Hochschild *et al.*, 1983; Kuldell & Hochschild, 1994; Li *et al.*, 1994; Ptashne, 1986). Regulation of promoters dependent on the cAMP receptor protein (CRP) also results from interactions between CRP and RNA polymerase

116

(Bell *et al.*, 1990; Niu *et al.*, 1994; Zhou *et al.*, 1994a; Zhou *et al.*, 1993b). At some promoters, CRP binds upstream of the -35 promoter region, whereas at other promoters, the CRP site overlaps the -35 region. In the former case the promoter-proximal monomer of the CRP dimer is responsible for activation whereas in the latter case, the promoter-distal monomer activates transcription (Zhou *et al.*, 1993a; Zhou *et al.*, 1994b). Thus, which CRP monomer is functional in activation depends on the position of the binding site with respect to the promoter. Arabinose-induced activation by AraC involves binding of AraC to two subsites composed of direct repeats separated by less than a helical turn of DNA. The promoter-proximal subsite slightly overlaps the -35 region of the promoter, and binding of a monomer of AraC to this subsite is absolutely required for activation. Binding of an AraC monomer to the promoter- distal subsite plays a less significant role in activation. Switching the orientation of the direct repeat subsites results in no activation by AraC. Therefore, regulation by AraC mainly requires binding in a particular orientation to one subsite that overlaps the -35 region (Reeder & Schlief, 1993). Arc is an interesting addition to the set of regulators of this type because the same *arc* DNA subsite may be primarily responsible for regulation of two divergent promoters even though it is in a different position with respect to each promoter.

Materials and Methods

Cloning of Promoter Fragments

Oligonucleotides were synthesized using an Applied Biosystems 381A DNA synthesizer and were annealed and ligated to the backbones of the appropriate plasmids for cloning. Each of the promoter variants was cloned as an EcoRI-BstEII cassette into the backbone of pSA660, a derivative of pSA600 (Milla *et al.*, 1993) in which the EcoRI site in the *arc* gene was eliminated by cassette mutagenesis (without changing the amino acid sequence), and an EcoRI site was added next to the P_{ant} promoter to allow cloning of promoter fragments. The sequence of the region surrounding the cloning sites does not derive from the *immunity I* operon and is shown for the P_{ant} cassette in Fig. 6.

Four of the P_{ant} promoter variants (P_{ant}, P_{ant} /35A, P_{ant}/10A, and P_{ant}/none) were also ligated as EcoRI-BstEII cassettes to a promoterless trpA'-lacZ fusion in plasmid pTS300. The pTS300 plasmid was constructed in the following manner. The plasmid pACYC184 was digested with HindIII and EagI and the large backbone fragment was ligated to the cassette

117

```
5'-AGCTTAGCCCGCCTAATGAGCGGGCTTTTTTTTGGCCGCCTAGGCCC-3'
    3'-ATCGGGCGGATTACTCGGCCGAAAAAAAACCGGCGGATCCGGGCCGG-5'
```

(which includes a trpA transcriptional terminator and an SfiI restriction site) to generate pTS100. The SfiI fragment from the plasmid miniTn5lacZ1 (De Lorenzo *et al.*, 1990) which contains the same promoterless trpA'-lacZ fusion as pRZ5605 (Mandecki & Reznikoff, 1982) was cloned into the SfiI site of pTS100 to generate pTS200. Plasmid pTS300 was generated by cloning the cassette

```
5'-AATTCGTATTGACATGATAGAAGCACTCTACTATATTCTGGTAACCGACGCCATGGCACCCC-3'
    3'-GCATAACTGTACTATCTTCGTGAGATGATATAAGACCATTGGCTGCGGTACCGTGGGG-5'
```

between the EcoRI and SmaI sites upstream of the trpA'-lacZ fusion. Introduction of this cassette placed the *lacZ* gene under the control of the wild-type P_{ant} promoter variant and resulted in the proper EcoRI and BstEII sites needed to clone the additonal promoter variants. The other promoter variant-*lacZ* constructs were made by cloning the appropriate cassette between these EcoRI and the BstEII sites. These plasmids are compatible with the pBR322 derived plasmid pSA600 which carries the *arc* gene (Milla *et al.*, 1993), allowing measurements of the effects of regulation by Arc *in vivo*.

Template Preparations

Linear templates for transcription and footprinting reactions were generated by the polymerase chain reaction (PCR) using a pair of primers with 5' ends located at - 87 and +85 with respect to the start site of transcription. The -87 primer was end-labeled with γ^{32}P-ATP using T4 polynucleotide kinase. PCR reactions were performed with the labeled -87 primer, the unlabeled +85 primer, and the linearized plasmid template DNA. The 173 bp PCR product was purified on a 10% polyacrylamide 0.5X TBE gel.

118

Protein Purification

The cooperativity defective variant Arc-SL35 was used in all experiments. Unlike wild-type Arc, this variant does not facilitate nonspecific binding of additional dimers next to specifically bound dimers (Smith & Sauer, 1995). For purification purposes, a C-terminal tail of six histidines (st6) was also added to the Arc-SL35 variant. This tail does not affect the stability or DNA binding properties of Arc (Milla *et al.*, 1993). The Arc-SL35(st6) protein was purified as described previously using nickel chelate chromatography (Qiagen) followed by SP-Sephadex chromatography (Milla *et al.*, 1993). RNA polymerase σ^{70} holoenzyme was purified as described (Hager *et al.*, 1990; Smith & Sauer, 1995).

DNAse I Footprinting Reactions

DNA (0.2 nM) was incubated with or without Arc-SL35 (75 nM) at 37 °C in the same buffer used for transcription assays (see below). DNase I was added to a final concentration of 18 ng/ml for 1 min. Reactions were quenched by the addition of an equal volume of 2.5 M ammonium acetate, 20 mM Na_2EDTA, and 10 μg/ml sonicated salmon sperm DNA. The reactions were extracted with phenol/chloroform, precipitated with ethanol using glycogen as a carrier, washed with 70% ethanol, dried, and resuspended in 5 μl of 85% (v/v) formamide, 10 mM NaOH, 1 mM EDTA, 0.1% (v/v) bromophenol blue, and 0.1% (v/v) xylene cyanole. Products of the reactions were resolved on 6% polyacrylamide, 8.3 M urea gels run in 1X TBE (90 mM Tris/Borate (pH 8.0), 2 mM Na_2EDTA).

Run-off Transcription Assays

DNA (0.2 nM) was incubated at 37 °C with or without Arc-SL35 (75 nM in the single-point assays) in a volume of 25 μl of buffer containing 30 mM Hepes-KOH (pH 7.5), 100 mM potassium glutamate, 10 mM $MgCl_2$, 1.5 mM $CaCl_2$, 0.1 mM Na_2EDTA, 100 μg/ml bovine serum albumin, 1 mM dithiothreitol, 0.02% NP-40, and nucleoside triphosphates (NTPs; 200 μM each GTP, CTP, ATP, and 6 μM UTP at a specific activity of 1.4×10^4 cpm/pmol). 1 unit of RNase inhibitor (Promega) was also included in each reaction. Reactions were initiated by the addition of RNA polymerase σ^{70} holoenzyme in a

volume of 5 μl to a final concentration of 7.5 nM. After 10 minutes, 5 μl of heparin (0.7 mg/ml) was added. After 10 additional minutes, 2 μl of 0.5 M Na_2EDTA and 40 μl of 8 M Urea, 0.1 % xylene cyanole, 0.1% bromophenol blue, and 50% glycerol were added and the reactions were heated to 90 °C before loading the gel. The products of the reactions were analyzed on 12% polyacrylamide 8.3 M urea gels run in 1X TBE. End-labeled DNA markers were used for approximate size standards. The major products of the reactions were transcripts of approximately 85 nucleotides in size. Appearance of these products was dependent on the addition of RNA polymerase, NTPs, and a promoter sequence in the appropriate position of the template. The template and transcript bands were quantitated using a Phosphorimager and the ImageQuant program (Molecular Dynamics), and the transcript bands were normalized to the amount of template in each lane.

DNA Mobility Shifts Experiments for Open-Complex Formation

DNA (0.2 nM) was incubated at 37 °C with or without Arc-SL35 (75 nM in the single-point assays) in 25 μl of the same buffer used for the transcription assays plus 1 unit of RNase inhibitor. RNA polymerase was added in a volume of 5 μl to a final concentration of 7.5 nM. After 2 min, formation of additional open complexes was stopped by the addition of 5 μl of a solution containing heparin (0.7 mg/ml) and 36% sucrose. In control reactions, this concentration of heparin prevented any association of free RNA polymerase with the DNA (not shown).

β-galactosidase Assays for Repression In Vivo

The pTS300 promoter variant plasmids (P_{ant}/both, P_{ant}/35A, P_{ant}/10A, and P_{ant}/none) were transformed into *E. coli* strain XL-1 Blue either alone or with a pSA600 plasmid encoding Arc-SL35 (Smith & Sauer, 1995) or Arc-RA13 (Brown *et al.*, 1994). The pTS300 plasmid was selected with 100 μg/ml chloramphenicol and the pSA600 plasmid was selected with 150 μg/ml ampicillin. Freshly transformed cells were grown overnight in LB broth with the appropriate antibiotics and were subcultured the next day and allowed to grow to an A_{600} of approximately 0.4-0.8. β-galactosidase assays were then performed as described (Miller, 1972).

Acknowledgements

References

Bell, A., Gaston, K., Williams, R., Chapman, K., Kolb, A., Buc, H., Minchin, S., Williams, J. & Busby, S. (1990). Mutations that alter the ability of the Escherichia coli cyclic AMP receptor protein to activate transcription. *Nucleic Acids Res.*, **18**, 7243-7250.

Brown, B.M., Milla, M.E., Smith, T.L. & Sauer, R.T. (1994). Scanning mutagenesis of the Arc repressor as a functional probe of operator recognition. *Nature Struct. Biol.*, **1**, 164-168.

Brown, B.M. & Sauer, R.T. (1993). Assembly of the Arc repressor-operator complex: cooperative interactions between DNA-bound dimers. *Biochemistry*, **32**, 1354-1363.

Carlson, N.G. & Little, J.W. (1993). Highly cooperative DNA binding by the coliphage HK022 repressor. *J. Mol. Biol.*, **230**, 1108-1130.

Collado-Vides, J., Magasanik, B. & Gralla, J.D. (1991). Control site location and tanscriptional regulation in *Escherichia coli. Microbiol. Rev.*, **55**, 371-394.

De Lorenzo, V., Herrero, M., Jakubzik, U. & Timmis, K.N. (1990). Mini-Tn5 transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in gram-negative eubacteria. *J. Bacteriol.*, **172**, 6568-6572.

Hager, D.A., Jin, D.J. & Burgess, R.R. (1990). Use of Mono Q high-resolution ion-exchange chromatogaphy to obtain highly pure and active *Echerichia coli* RNA polymerase. *Biochemistry*, 7890-7894.

He, Y.-Y., McNally, T., Manfield, I., Navratil, O., Old, I., Phillips, S.E.V., Saint-Girons, I. & Stockley, P.G. (1992). Probing *met* repressor-operator recognition in solution. *Nature*, **359**, 431-433.

Hochschild, A., Irwin, N. & Ptashne, M. (1983). Repressor structure and the mechanism of positive control. *Cell*, **32**, 319-325.

Kim, B. & Little, J.W. (1992). Dimerization of a specific DNA-binding protein on the DNA. *Science*, **255**, 203-206.

Kuldell, N. & Hochschild, A. (1994). Amino acid substitutions in the -35 recognition motif of σ^{70} that result in defects in phage λ repressor-stimulated transcription. *J. Bacteriol.*, **176**, 2991-2998.

Li, M., Moyle, H. & Susskind, M.M. (1994). Target of the transcriptional activation function of phage λ cI protein. *Science*, **263**, 75-77.

Liao, S.-M. (1988) Ph.D. Thesis.. Carnegie-Mellon University.

Mandecki, W. & Reznikoff, W.S. (1982). A lac promoter with a changed distance between -10 and -35 regions. *Nucleic Acids Res.*, **10**, 903-912.

McClure, W.R. (1985). Mechanism and control of transcription initiation in prokaryotes. *Annu. Rev. Biochem.*, **54**, 171-204.

Meyer, B.J., Maurer, R. & Ptashne, M. (1980). Gene regulation at the right operator (OR) of bacteriophage 1. II. OR1, OR2, and OR3: their roles in mediating th effects of repressor and cro. *J. Mol. Biol.*, **139**, 163-194.

Milla, M.E., Brown, B.M. & Sauer, R.T. (1993). P22 Arc repressor: enhanced expression of unstable mutants by addition of polar C-terminal sequences. *Protein Science*, **2**, 2198-2205.

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

Niu, W., Zhou, Y., Dong, Q., Ebright, Y. & Ebright, R.H. (1994). Characterization of the activating region of Escherichia coli catabolite gene activator protein (CAP). I. Saturation and alanine-scanning mutagenesis. *J. Mol. Biol.*, **243**, 595-602.

Phillips, S.E.V., Manfield, I., Parsons, I., Davidson, B.E., Rafferty, J.B., Somers, W.S., Margarita, D., Cohen, G.N., Saint-Girons, I. & Stockley, P.G. (1989). Cooperative tandem binding of *met* repressor of *Escherichia coli. Nature*, **341**, 711-715.

Ptashne, M. (1986). *A Genetic Switch.* Cell Press, Cambridge, MA.

Raumann, B.E., Rould, M.A., Pabo, C.O. & Sauer, R.T. (1994b). DNA recognition by β-sheets in the arc repressor-operator crystal structure. *Nature*, **367**, 754-757.

Reeder, T. & Schlief, R. (1993). AraC protein can activate transcription from only one position and when pointed in only one direction. *J. Mol. Biol.*, **231**, 205-218.

Smith, T.L. & Sauer, R.T. . (1995). P22 Arc Repressor: Role of cooperativity in repression and binding to operators with altered half-site spacing. *J. Mol. Biol.*, **249**, 729-742.

Smith, T.L. & Sauer, R.T. (1996). Dual regulation of distinct steps in transcription explains a novel repressor to activator switch. *submitted.*

Susskind, M.M. & Youderian, P., ed. *Bacteriophage P22 antirepressor and its control.* Lambda II, ed. Hendirx, R.W., *et al.* 1983, Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY.

Vershon, A.K., Kelley, R.D. & Sauer, R.T. (1989). Sequence-specific binding of Arc repressor to DNA. *J. Biol. Chem.*, **264**, 3267-3273.

Vershon, A.K., Liao, S.-M., McClure, W.R. & Sauer, R.T. (1987a). Bacteriophage P22 Mnt repressor: DNA binding and effects on transcription in vitro. *J. Mol. Biol.*, **195**, 311-322.

Vershon, A.K., Liao, S.-M., McClure, W.R. & Sauer, R.T. (1987b). Interaction of the bacteriophage P22 Arc repressor with operator DNA. *J. Mol. Biol.*, **195**, 323-331.

Waldburger, C.D. & Sauer, R.T. (1995). Domains of Mnt Repressor: Roles in Tetramer Formation, Protein Stability, and Operator DNA Binding. *Biochemistry*, **34**, 13109-13116.

Zhou, Y., Busby, S. & Ebright, R.H. (1993a). Identification of the functional subunit of a dimeric transcription activator protein by use of oriented heterodimers. *Cell*, **73**, 375-379.

Zhou, Y., Merkel, T.J. & Ebright, R.H. (1994a). Characterization of the activating region of *Escherichia coli* catabolite gene activator protein (CAP). II. Role at class I and class II CAP-dependent promoters. *J. Mol. Biol.*, **243**, 603-610.

Zhou, Y., Pendergrast, P.S., Bell, A., Williams, R., Busby, S. & Ebright, R.H. (1994b). The functional subunit of a dimeric transcription activator protein depends on promoter architecture. *EMBO J.*, **13**, 4549-4557.

Zhou, Y., Zhang, X. & Ebright, R.H. (1993b). Identification of the activating region of catabolite gene activator protein (CAP): Isolation and characterization of mutants of CAP specifically defective in transcription activation. *Proc. Natl. Acad. Sci.*, **90**, 6081-6085.

Promoter Variant	Mutations	Promoter Strength
P_{ant}/both	—	1.00
P_{ant}/35A	GTGC at -17 to -14	0.70 ± 0.04
P_{ant}/35B[a]	AG at -28 to -27, GT at -22 to -21, GTGC at -17 to -14	0.85 ± 0.14
P_{ant}/10A	AGGCACGT at -28 to -21	1.50 ± 0.05
P_{ant}/10B[b]	G at -30, AGGCACGT at -28 to -21	1.53 ± 0.07
P_{ant}/none	AGGCACGT at -28 to -21, GTGC at -22 to -21	1.29 ± 0.01
P_{mnt}/both	—	0.26 ± 0.09
P_{mnt}/35	GT at -14 to -13	1.13 ± 0.26
P_{mnt}/10	AC at -21 to -20	0.43 ± 0.04
P_{mnt}/none	AC at -21 to -20, GT at -14 to -13	1.74 ± 0.40

Table 1. Mutations in the promoter variants shown in Fig. 2 and promoter strengths *in vitro* (relative to the P_{ant}/both variant) in the absence of Arc (shown are the averages ± standard deviation, n ≥ 3). The sequences of the reference promoter variants, P_{ant}/both and P_{mnt}/both, are shown in Fig. 2; sequence substitutions refer to the top strands shown in Fig. 2. [a]In the P_{ant}/35B variant, the sequence of the wild-type P_{ant} -35 proximal *arc* subsite is replaced by the sequence of the -10 proximal subsite. [b]In the P_{ant}/10B variant, there is a change from a consensus to a nonconsensus bp in the -35 promoter region relative to the P_{ant}/10A variant.

Promoter Variant	Protein	β-galactosidase (Miller Units)
P$_{ant}$/both	no Arc	$3.0\ (\pm0.5) \times 10^4$
	Arc-RA13	$3.0\ (\pm0.8) \times 10^4$
	Arc-SL35	$0.7\ (\pm0.1) \times 10^4$
P$_{ant}$/35A	no Arc	$2.6\ (\pm0.5) \times 10^4$
	Arc-RA13	$2.7\ (\pm0.6) \times 10^4$
	Arc-SL35	$1.2\ (\pm0.2) \times 10^4$
P$_{ant}$/10A	no Arc	$2.6\ (\pm0.4) \times 10^4$
	Arc-RA13	$3.1\ (\pm0.8) \times 10^4$
	Arc-SL35	$2.1\ (\pm0.3) \times 10^4$
P$_{ant}$/none	no Arc	$2.6\ (\pm0.5) \times 10^4$
	Arc-RA13	$2.6\ (\pm0.2) \times 10^4$
	Arc-SL35	$2.4\ (\pm0.3) \times 10^4$

Table 2. β-galactosidase expression from fusions of certain P$_{ant}$ promoter variants to *lacZ* assayed in the absence or presence of Arc-SL35 or the inactive Arc-RA13 mutant. Values shown are averages ($\pm$ standard deviations, $n \geq 3$).

Figure 1. Wild-type immunity I region of bacteriophage P22. The -35 and -10 regions of the P_{ant} and P_{mnt} promoters are indicated. The six bp sequences to which Arc dimers make base-specific contacts in the Arc-operator cocrystal structure are boxed (Raumann *et al.*, 1994). Arc dimers, shown as ovals, interact to form a cooperatively stabilized tetramer on the DNA.

$P_{ant} \longrightarrow$

-35   -10

TAGTGTATTGACAT GATAGA AGCAC TCTACT ATATTCTCAATAGG
ATCACATAACTGTA CTATCT TCGTG AGATGA TATAAGAGTTATCC

        -10   -35

arc binding sites
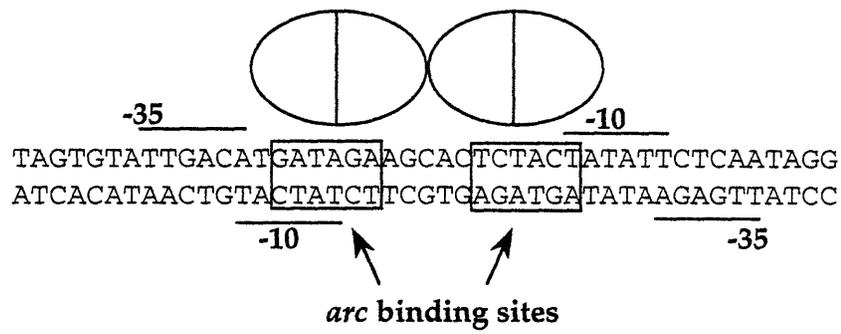
$\longleftarrow P_{mnt}$

Figure 2. Promoter variants. The *arc* operator subsites are shown as boxes and the Arc dimers as ovals. Subsites inactivated by mutations are marked by crosses. The sequences of the reference P_{ant}/both and P_{mnt}/both promoters are shown. Sequence changes in the remaining variants are listed in Table 1.

P$_{ant}$ Promoter Variants

-35   -10
GAATTCGTATTGACATGATAGAAGCACTCTACTCTATATTCTGGTAAC
CTTAAGCATAACTGTACTATCTTCGTGAGATGATATAAGACCATTG

P$_{ant}$/both

P$_{ant}$/35A or 35B

P$_{ant}$/10A or10B

P$_{ant}$/none

P$_{mnt}$ Promoter Variants

-35   -10
GAATTCGTATTGAGAATATAGTAGAGTGCTTCTATCTATCAGGGTAACC
CTTAAGCATAACTCTTATATCTCACGAAGAGATAGTAGTCCCATTGG

P$_{mnt}$/both

P$_{mnt}$/35

P$_{mnt}$/10

P$_{mnt}$/none

130

Figure 3. DNase I footprinting assays of the binding of Arc-SL35 to promoter variants. Protection of only the active subsites is observed. Arc-SL35 footprints on the $P_{ant}/35B$ and the $P_{ant}/10B$ variants (not shown) were similar to those on $P_{ant}/35A$ and $P_{ant}/10A$, respectively.

P_{mnt} Promoter Variants

P_{ant} Promoter Variants

Figure 4. A) Run-off transcription ($\pm$ Arc-SL35) from the P_{ant}/both promoter assayed by gel electrophoresis. The positions of the run-off transcript, the template DNA, and the end-to-end transcript on the gel are indicated. B) Repression of the P_{ant} and P_{mnt} promoter variants by Arc-SL35. Shown are the averages ($\pm$ standard deviation, $n \geq 4$) of the amount of transcript present with Arc-SL35 divided by the amount of transcript present without Arc-SL35. A value of 1.0 indicates no repression; a value of 0 indicates complete repression. C) Run-off transcription from the P_{ant}/both, P_{ant}/35A, P_{ant}/10A, and P_{ant}/none promoters as a function of increasing Arc-SL35 concentration. Shown is the amount of transcript present with Arc-SL35 divided by the amount of transcript present without Arc-SL35.

Arc-SL35

end-to-end transcript
template DNA

run-off transcript
from P_{ant}/both

$$\frac{Tx + Arc\text{-}SL35}{Tx\ No\ Arc}$$

No Repression

Promoter Variant

Pmnt/none
Pmnt/10
Pmnt/35
Pmnt/both
Pant/none
Pant/10B
Pant/10A
Pant/35B
Pant/35A
Pant/both

1.5
1.0
0.5
0

$$\frac{\text{Tx (+ Arc-SL35)}}{\text{Tx (no Arc)}}$$

Arc-SL35 (nM)
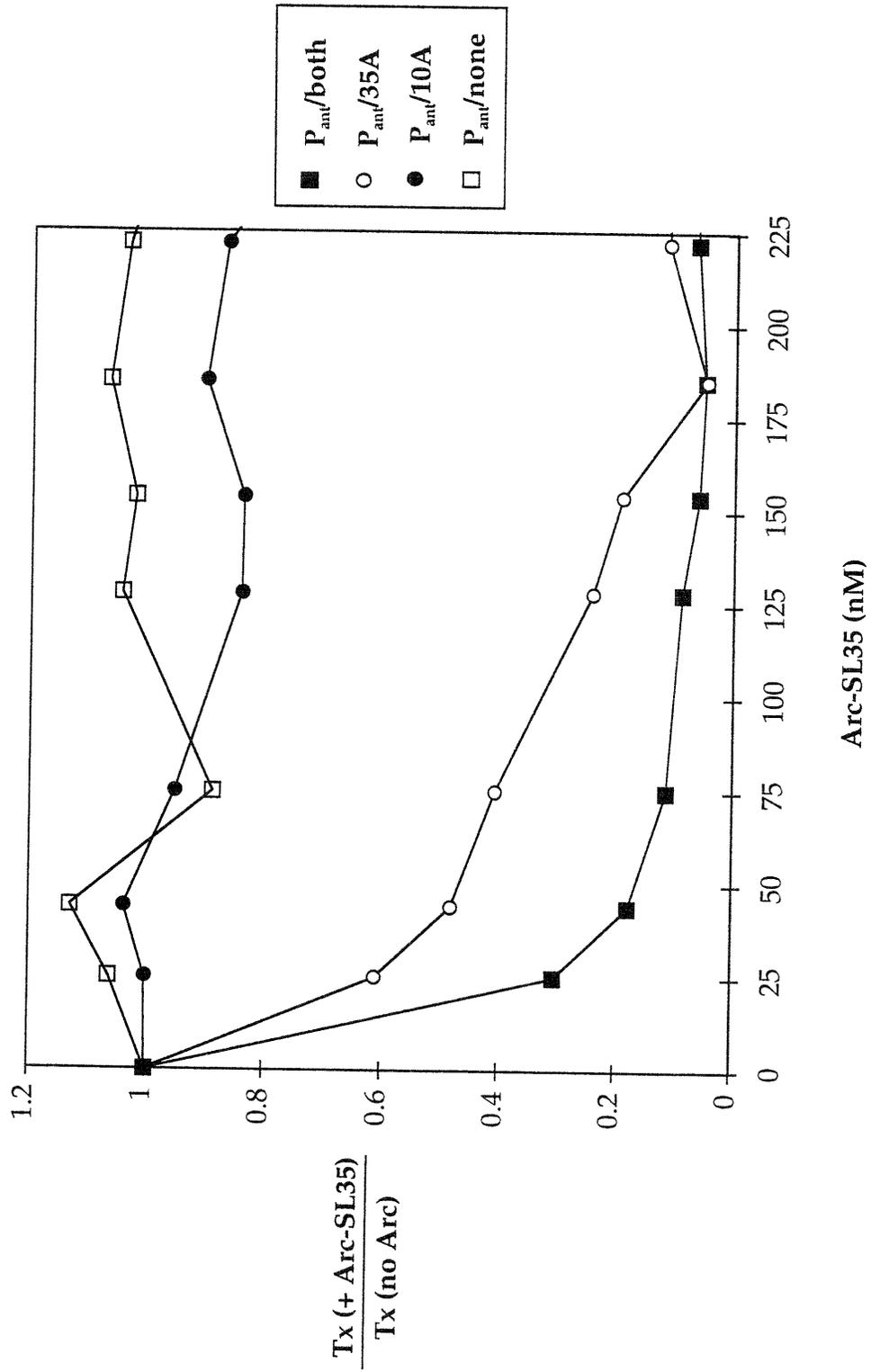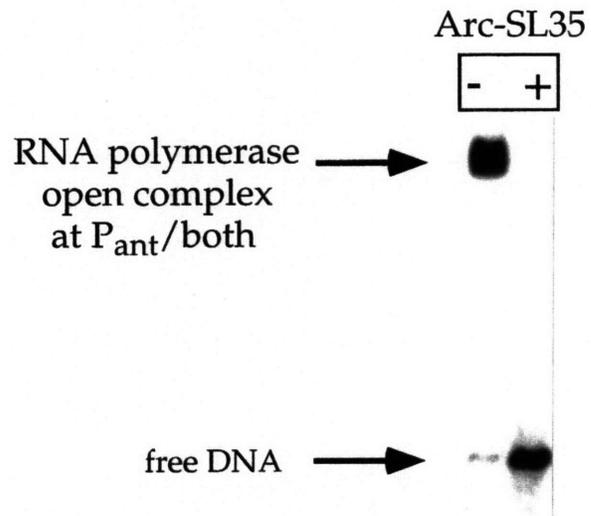
P_{ant}/both
P_{ant}/35A
P_{ant}/10A
P_{ant}/none
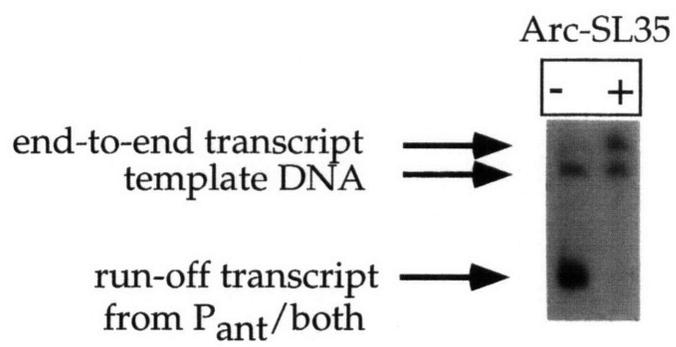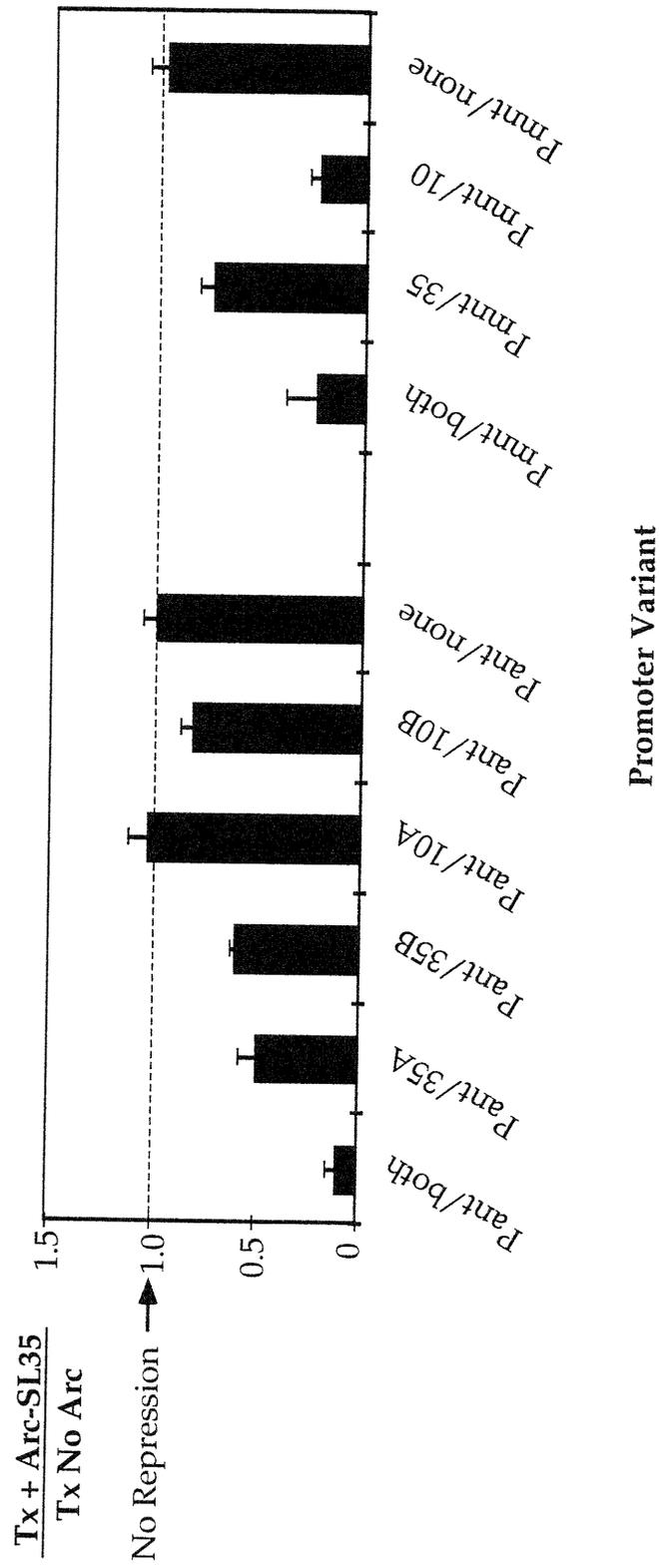
Figure 5. A) Formation of open complexes assayed by gel electrophoresis ($\pm$ Arc-SL35) 2 min after mixing the P_{ant}/both variant and RNA polymerase. The positions of the free DNA and the RNA polymerase open complex on the gel are indicated. B) Extent of open-complex formation 2 min after addition of RNA polymerase to each promoter variant in the presence or absence of Arc-SL35. Shown are the averages ($\pm$ standard deviation, $n \geq 3$) of the fraction of DNA in the open-complex band. C) Extent of open-complex formation as a function of Arc-SL35 on the P_{ant}/both, P_{ant}/35A, P_{ant}/10A, and P_{ant}/none promoters. Shown is the fraction of DNA in open complexes in the presence of Arc-SL35 divided by the fraction DNA in open complexes in the absence of Arc-SL35.

Arc-SL35

| − | + |

RNA polymerase ⟶
open complex
at P_{ant}/both

free DNA ⟶

140
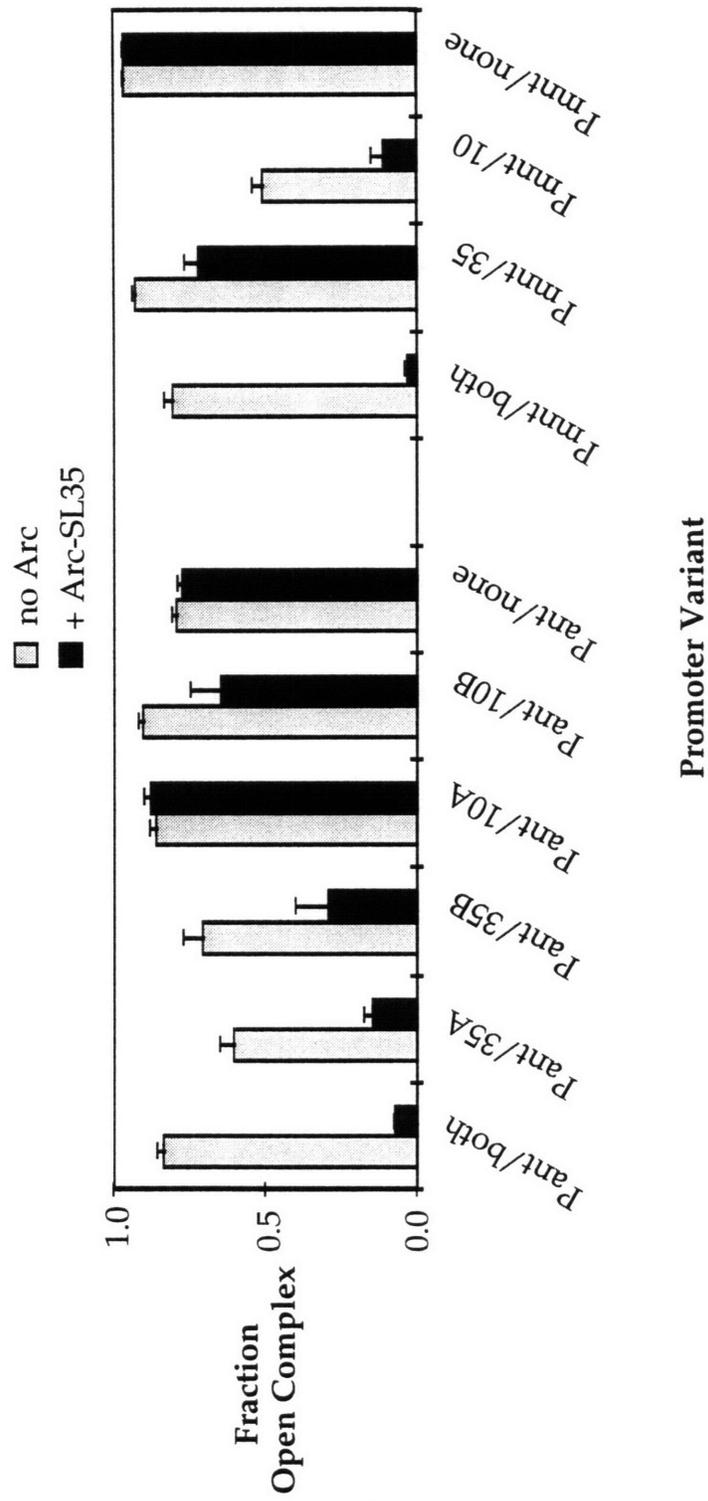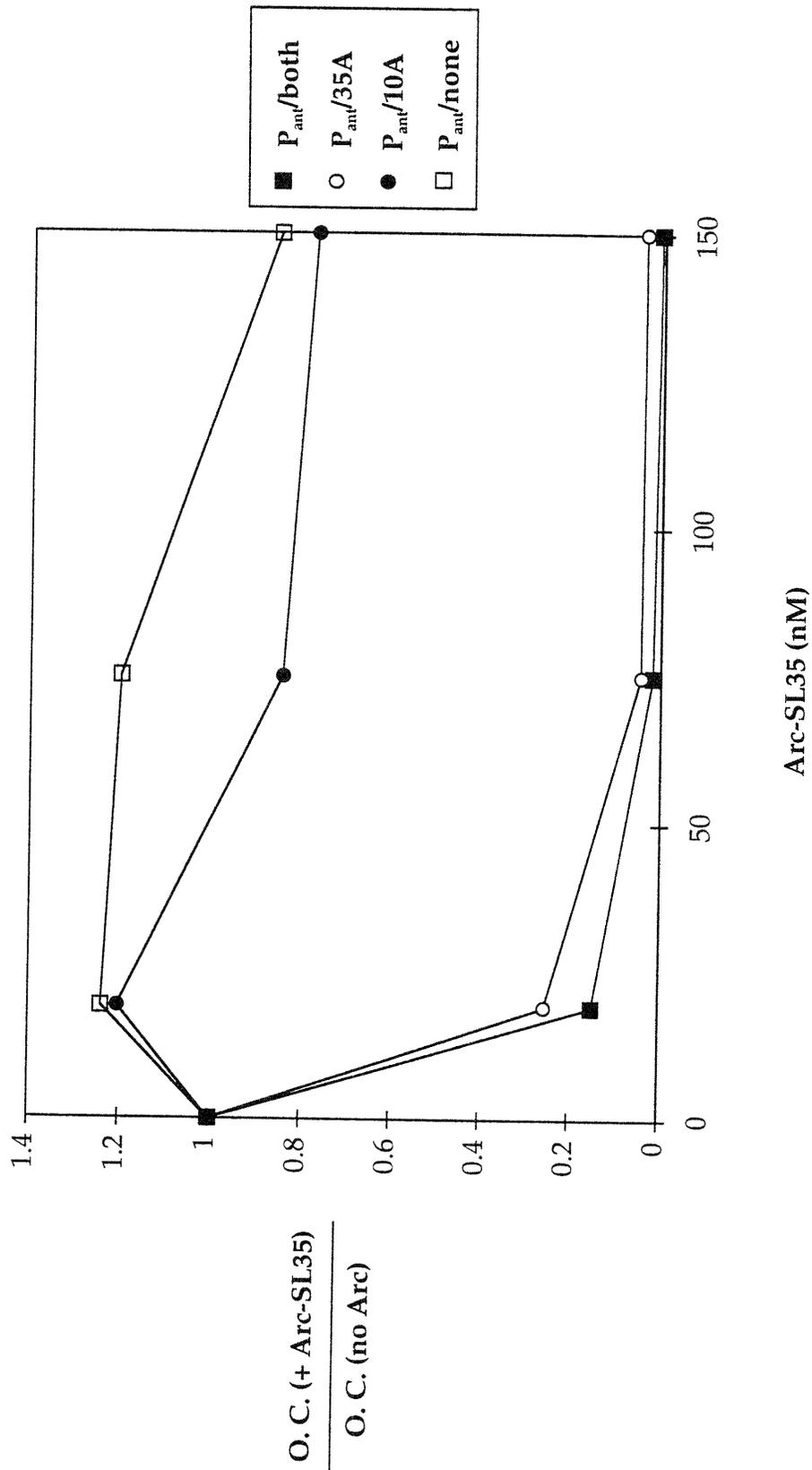
Figure 6. Flanking sequence of the P_{ant}/both variant showing the cloning sites.

arc binding sites

EcoRI P_{ant} -35 P_{ant} -10   BstEII

```
GCATCTGTGCGGTATTTCACACCGCCATAGAATTCGTATTGACATGATAGAAGCACTCTACTTATATTTCTGGGTAACCGTCGACAGGAGCTATTA
CGTAGACACGCCATAAAGTGTGGCGGTATCTTAAGCATAACTGTACTATCTTCGTGAGATGAATAAGACCCATTGGCAGCGTCCTCGATAAT
```

Chapter 5

Summary and Directions for Future Research

Overview of Thesis

The experiments presented in this thesis were designed to investigate cooperative DNA binding and transcriptional regulation by Arc repressor. By focusing on mutant proteins and operators that cannot support cooperative DNA binding of Arc, I showed that Arc cooperativity is inflexible with respect to half-site spacing and is required for repression, most likely to saturate the operator at low concentrations of Arc and over a small range of concentrations. I have also demonstrated that one Arc dimer bound to a single site can regulate transcription both negatively or positively. This depends on the promoter context because Arc can regulate two steps in transcription initiation. Lastly, I presented results which suggest that the two Arc dimer binding sites in the wild-type operator may not be equivalent in repression of P_{ant} and P_{mnt} and that the site proximal to the P_{ant} -35 region may play a more important role in repression of both promoters.

In this chapter, I compare some of the results obtained with Arc to the examples described in Chapter 1 and suggest directions for future research on cooperative DNA binding and transcriptional regulation by Arc. The preliminary results presented in this chapter are meant to indicate the types of experiments that could be conducted to gain further knowledge about this system and are not meant to be comprehensive.

Chapter 2 Summary and Future Directions

The results presented in Chapter 2 demonstrate that cooperativity is important for repression by Arc, as it is for regulation by λ repressor. However, cooperativity is not important structurally for maintaining a particular DNA structure or for creating a specific protein-protein interface recognized by RNA polymerase. The possibility of a structural contribution of cooperativity was considered because the operator DNA is bent and underwound in the Arc-operator cocrystal structure (see Figure 1). It was reasonable to suggest that the DNA distortion seen in crystal structure is mediated by the cooperative contacts. Since the Arc operator is positioned between the -35 and the -10 elements of the P_{ant} promoter, the bending and underwinding resulting from Arc binding could conceivably place those elements out of alignment, thus repressing transcription. This model is similar to the DNA distortion model proposed for activation of transcription by MerR which is dependent on a distortion that favorably realigns the promoter elements. The fact that the cooperativity defective Arc mutants described in Chapter 2 have 1, 10-

144

phenanthroline-copper footprints identical to that of the wild-type protein suggests that the distortions are not dependent on cooperativity because that reagent is sensitive to the geometry of the minor groove. In addition to having similar footprints, the cooperativity defective Arc mutants also repress transcription as well as wild-type Arc when the operator is saturated *in vitro*. From these experiments, we learned that that the cooperative DNA contacts do not appear to be responsible for the DNA distortions in the operator and that cooperativity is not required in repression for any other structural purpose.

It is difficult to know with certainty whether the DNA distortion is responsible for repression because Arc dimer binding may be all that is needed to distort each half site. This model is supported by the finding that certain mutations in Arc that disrupt direct base-specific contacts result in slightly altered 1, 10-phenanthroline-copper footprints (see Figure 2), indicative of altered DNA distortions. Distortion of the DNA caused by Arc binding might be important for repression because it unfavorably realigns the -35 and -10 promoter elements. To address this possibility, promoters were constructed in which the -35 and -10 elements of a P_{ant} promoter variant were separated or brought together by 1 or 2 bp without altering the important *arc* operator sequences, and regulation by Arc was measured *in vitro*. The rationale behind these experiments was that if the realignment model is correct, then one of these promoters may no longer be repressed or may be activated by Arc binding. However, each of these promoters was repressed by Arc (data not shown). This negative result is somewhat difficult to interpret because the mechanisms of repression of the promoters may differ. Nevertheless, the fact that a single Arc dimer bound to a half site can repress transcription fairly well when the site is saturated (Chapters 3 and 4; also see below) argues against the realignment model because the operator distortions in both half sites (as indicated by footprinting) are not seen when a single dimer binds (data not shown). Furthermore, the results in Chapter 4 (see below) suggest that the two Arc dimer positions in the operator are not equivalent in repression although each is capable of causing some distortion in the DNA.

Chapter 2 also demonstrates that the cooperative DNA binding of Arc, unlike that of λ repressor, cannot accommodate changes in half-site spacing. This is not unexpected because the DNA binding residues and the cooperativity residues of Arc are on the same small domain whereas the C-terminal cooperativity domain of λ repressor is flexibly tethered to a separate DNA binding domain. It may be possible to design or select Arc variants that use a different surface for cooperativity and thus bind cooperatively to operators with different half-site spacings. This is suggested because the crystal strucure

of another ribbon-helix-helix protein, MetJ shows that this protein uses a different face of the ribbon-helix-helix motif for cooperative DNA binding of two dimers to an operator (Phillips *et al.*, 1989; Somers & Phillips, 1992). Each MetJ dimer DNA-binding domain in this cocrystal structure interacts with a half site in a manner similar to an Arc dimer bound to one of the two half sites in its operator. The MetJ operator also consists of two half sites of the same size as the *arc* half sites. However, the spacing between the MetJ half sites is 3 bp less than the spacing between the Arc half sites (see Figure 3). This spacing orients the dimers on the DNA such that two MetJ A helices are apposed, allowing the two dimers to interact cooperatively (see Figure 4). The results in Chapter 2 demonstrate that two Arc dimers can bind noncooperatively to a similar operator, termed the -3 operator, in which 3 bp have been deleted between the half sites. DNAse I footprinting confirms that the Arc tetramer on this operator is positioned as expected (see Figure 5).

In preliminary experiments, I attempted to randomize the solvent exposed residues of helix A in Arc and to select for mutants that could bind cooperatively to the -3 operator. In these experiments, selection constructs were made in which the wild-type *arc* operator or the -3 operator was positioned after the +1 start site of transcription of the streptomycin sensitive gene *rpsl* (see Figure 6). The *rpsl* gene encodes a ribosomal subunit. When *rpsl* is overexpressed on a plasmid, it confers a dominant streptomycin sensitivity phenotype to cells containing a streptomycin resistant allele of *rpsl* in the chromosome (see Chapter 2 for a full description of this assay). The wild-type operator resulted in repression by wild-type Arc (a streptomycin resistant phenotype) whereas the -3 operator did not (a streptomycin sensitive phenotype). This type of selection system could potentially be used to select for altered cooperativity mutants. Such mutants would be expected to bind more tightly to the -3 operator and thus to repress transcription of the streptomycin sensitive *rpsl* gene. However, the constructs described above cause *E. coli* to grow very slowly. It will be necessary to optimize these constructs for better growth conditions, perhaps by reducing the strength of the promoter driving the *rpsl* gene on the plasmid.

Chapter 3 Summary and Future Directions

The results in Chapter 3 demonstrate that a single Arc dimer bound to the P_{ant} -35 proximal half site can both repress and activate transcription, depending on the rate-limiting step of the P_{ant} promoter variant in the presence of Arc. This occurs because Arc can act at two steps in transcription initiation when bound to this site, slowing open-complex

146

formation and accelerating promoter clearance. Can the other half-site proximal to the P_{ant} -10 region affect these two steps? Chapter 4 (see below) demonstrates that the P_{ant} -10 proximal site is not as efficient as either the wild-type operator or the P_{ant} -35 proximal site in preventing RNA polymerase open-complex formation. To address whether the the wild-type operator and/or the P_{ant} -10 proximal *arc* half site alone can affect promoter clearance, two constructs were made that are predicted to have slower rates of clearance because of the change to a consensus bp (T$\rightarrow$A) at positon -8 (indicated by a * in Figure 7). One is similar to the P_{ant}/both variant and the other is similar to P_{ant}/10A described in Chapter 4. Arc can accelerate clearance from the P_{ant}/both* promoter but not from the P_{ant}/10A* promoter as shown by an assay that measures the accumulation of transcripts from preformed open complexes (see Figure 7). Arc did not enhance clearance from the P_{ant}/10A* variant even at a concentration of 1 μM, whether or not Arc was added prior to open complex formation (not shown).

The inability of Arc to accelerate promoter clearance from the P_{ant}/10A* half-site construct can be explained in two ways. Either Arc cannot bind to the -10 proximal site when RNA polymerase is bound in the open complex at P_{ant} or an Arc dimer bound to this site cannot accelerate clearance because it is not properly positioned. The former steric occlusion model is supported by a comparison of the base and phosphate contacts known to be made by Arc and by RNA polymerase bound to other promoters (see Figure 8). The P_{ant} -10 proximal *arc* half-site overlaps a region of extensive contact by RNA polymerase, suggesting that an Arc dimer may not be able to occupy that site simultaneously with RNA polymerase in the open complex. The hypothesis that Arc cannot bind with the RNA polymerase open complex could be tested by radiolabeling Arc with ^{35}S and testing whether Arc migrates with open complexes on this promoter construct. Alternatively, crosslinking the Arc-DNA complexes or chemically modifying Arc to cleave DNA when it binds could be employed to address this question.

The observation that the P_{ant} -35 proximal *arc* site but not the P_{ant} -10 proximal site can mediate acceleration of promoter clearance by Arc is similar to results described in Chapter 4 (see below) which suggest that the two *arc* half-sites also may not be equivalent in repression and that the P_{ant} -35 proximal half-site may be more important for negative regulation of both P_{ant} and P_{mnt}. Thus, there may be functional differences in regulation by the two dimers bound to the wild-type operator. These differences may exist because of the positions of the binding sites with respect to the -35 and -10 promoter elements. Since the Arc binding sites are positioned differently with respect to the two sets of promoter

147

elements of the divergent P_{ant} and P_{mnt} promoters, it would be interesting to determine if Arc can accelerate promoter clearance from the P_{mnt} promoter or its half-site operator variants.

As shown in Chapter 3, Arc can both slow open-complex formation and accelerate promoter clearance. λ repressor and CRP can act at different steps in transcription, but this seems to be dependent on the position of the binding site with respect to the promoter in both cases. In contrast, both Arc and MerR can act at different steps in transcription when bound to one position. Arc binds between the P_{ant} -10 and -35 sites in a position similar to that occupied by MerR at the $P_{TP(C)AD}$ promoter. Perhaps this configuration is ideal for multiple functions because it is close to RNA polymerase and to the important promoter DNA elements. λ repressor, CRP, and MerR all seem to interact with RNA polymerase to regulate multiple functions, and it may be that Arc also has specific interactions with RNA polymerase that prevent isomerization and/or accelerate clearance. Screening for Arc mutants that can still bind DNA but cannot repress might identify residues important for repression. I have screened some Arc mutants (Milla et al., 1994) with changes on the surface of helix A of Arc to analyze whether they can acclerate clearance from the P_{ant}/both* promoter (see Figure 7 for a description of P_{ant}/both*). The Arc variants were added to preformed open complexes, and the accumulation of transcripts after 1 minute was measured. These mutants were tested because the surface of helix A is in a position that could contact RNA polymerase. This preliminary analysis suggested that the Arc-RA23 mutant could be defective in clearance acceleration (see Figure 9). This mutant binds to DNA at the concentrations used for the clearance assays (not shown), but it is known to have an increased dissociation rate from wild-type operator DNA (Brown et al., 1994). It will need to be determined if other Arc mutants with increased dissociation rates show a similar inability to accelerate clearance before a direct contact between Arg23 and RNA polymerase can be proposed.

The clearance acceleration results presented in Chapter 3 indicate that Arc bound to the P_{ant} -35 proximal half site can occupy the DNA simultaneously with RNA polymerase in the open complex at P_{ant}, but the P_{ant} -10 proximal site may not be occupied when RNA polymerase is bound (Figure 8). Following logically from these results with the open complex is that a dimer at the -35 proximal position (but perhaps not at the -10 position) might occupy the DNA when RNA polymerase is in the closed complex and have a repression effect on later steps in open-complex formation. Simultaneous binding of Arc and RNA polymerase was also suggested by the results of abortive initiation experiments

148

performed by Liao and McClure which indicated that Arc represses P_{ant} by affecting the isomerization to the open complex and not by affecting binding of the closed complex (Liao, 1988). Throughout the course of my thesis research, I performed many DNA mobility shift, crosslinking, and footprinting experiments attempting to confirm simultaneous binding of Arc and RNA polymerase in the closed complex on the wild-type P_{ant} promoter directly, but all results were negative. It may be difficult to trap such a complex if it consists only of an Arc dimer and the RNA polymerase closed complex because the half-life of an Arc dimer bound to DNA is rather short and the dissociation of the closed complex is also expected to rapid. In contrast to the Arc situation, it can be directly demonstrated that MerR can bind simultaneously with the RNA polymerase closed complex. Simultaneous binding is demonstrable in this system most likely because MerR binds tightly to its operator and actually stabilizes the closed complex at $P_{TP(C)AD}$.

Chapter 4 Summary and Future Directions

The results presented in Chapter 4 suggest that the two dimer binding sites in the wild-type *arc* operator may not be equivalent in repression of P_{ant} and P_{mnt}. The P_{ant} -35 proximal site (which overlaps the P_{mnt} -10 region because the two promoters are divergent) seems to play a greater role in repression (see Figure 10). The nonequivalence of the sites is also supported by the preliminary results presented above which indicate that the P_{ant} -35 *arc* site but not the P_{ant} -10 site can mediate acceleration of promoter clearance from several P_{ant} promoter variants. Thus, in both the repressing and activating activities of Arc, the P_{ant} -35 proximal site may play a greater role. In λ repressor activation of P_{RM} and in CRP regulation at class I and class II promoters, only one subunit of the DNA bound dimer is primarily responsible for activation even though the subunit cannot bind tightly to DNA on its own. This is similar to the situation described here, except that in the Arc system, the responsible subunit (an Arc dimer) is capable of binding to DNA with a reasonable affinity. A dimer bound to the P_{ant} -35 proximal site may be properly positioned to affect steps in transcription initiation (at least at P_{ant}) both positively and negatively. The second dimer binding site proximal to the P_{ant} -10 region may contribute to regulation mainly by enabling cooperative stabilization of the dimer proximal to the P_{ant} -35 region.

149

Summary

A large amount of information about Arc was available when I began an investigation of its regulatory role, and yet the Arc system has revealed some surprises and some interesting variations on common themes of transcriptional regulation. Most notably, the small Arc protein has the unexpected ability to regulate two steps in transcription initiation, and the two dimers of the Arc tetramer, which stabilize each other on the DNA through cooperative contacts, may play unequal roles in regulation, perhaps because only one dimer overlaps an important DNA region or because only one is positioned properly to interact with RNA polymerase. Further research on the Arc system may reveal that certain Arc residues do interact with RNA polymerase to either slow open-complex formation or to enhance promoter clearance, possibly leading to an understanding of the mechanisms by which a small regulatory protein can affect RNA polymerase at two separate steps in transcription initiation.

References

The citations in this chapter refer to the references in Chapter 1.

Figure 1. Arc tetramer-operator cocrystal structure (Raumann *et al.*, 1994). The two Arc dimers are shown in ribbon form, and a trace of the DNA backbone is shown below. Note the distortion of the DNA, especially the narrowness of the minor groove below the cooperative interface of the two dimers. This figure was prepared using the program MOLSCRIPT (Kraulis, 1991).
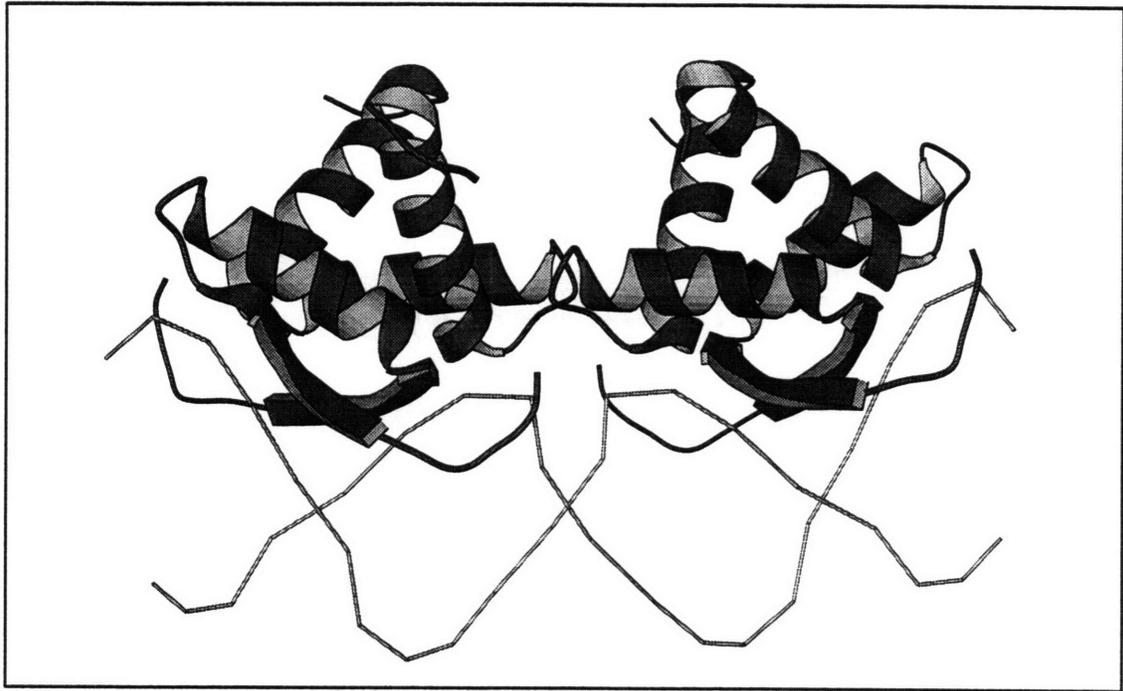
Figure 2. Left, summaries of 1, 10-phenanthroline-copper footprints of 19 Arc DNA-contact, alanine-scan mutants on the wild-type *arc* operator (see Brown *et al.*, 1994 for full descriptions). The positions of protections and enhancements are indicated. Larger arrows indicate enhanced cleavage, solid circles mark unprotected positions, and the remaining bases within the brackets are protected from cleavage. Right, example traces of 1, 10-phenanthroline-copper footprints. The dotted lines mark the operator positions 1, 11, and 21 seen in the footprint summaries. 1, 10-phenanthroline-copper is sensitive to the accessibility of the minor groove and can indicate regions of distorted DNA. All showed the same footprint pattern except for three direct DNA contact mutants, MA4, QA9 and NA11, and one indirect contact mutant, NA34. NA34 bridges the β-sheet direct DNA contacts and the phosphate contacts made by another portion of the protein. The enhancements and protections of the 1, 10-phenanthroline footprints are likely to be indicative of the bending and/or unwinding of the operator. Therefore, these results indicate that mutation of residues that disrupt DNA contacts affect the distortion of the DNA. This figure was published in Brown *et al.* (1994).

Wild type

QA9

DNA alone

Wild type and
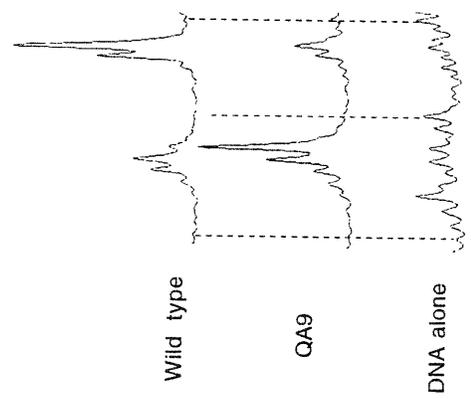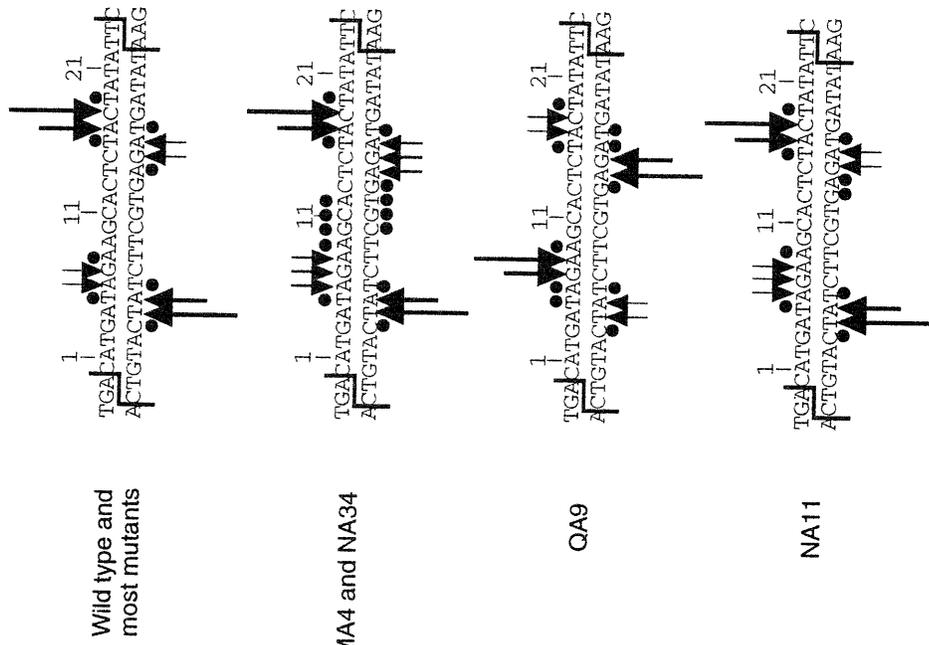most mutants

MA4 and NA34

QA9

NA11

155

Figure 3. Comparison of *arc* and *metJ* operators. Both Arc and MetJ bind as tetramers to these operators. Each MetJ and Arc dimer specifically recognizes bases within the six bp sequences boxed in the operators. The two half sites are separated by 5 bp in the *arc* operator and by 2 bp in the *metJ* operator.

Comparison of *arc* and *metJ*
Half-Site Spacings

arc AT|GATAGA|AGCAC|TCTACT|AT
     TA|CTATCT|TCGTG|AGATGA|TA

metJ CTA|GACGTC|TA|GACGTC|TAG
      GAT|CTGCAG|AT|CTGCAG|ATC

Figure 4. Comparison of cooperative interfaces of Arc and MetJ in the cocrystal structures (Raumann et al., 1994; Somers & Phillips, 1992). These views are down the two-fold axes of the complexes; the DNA is behind the two dimers. Both Arc and MetJ use the ribbon-helix-helix motif to bind to DNA half sites. The half-site spacing of the metJ operator is 3 bp less than that of the arc operator, resulting in a different positioning of the two dimers and different modes of cooperative DNA binding. Arc uses the loops between helices A and B for cooperative contacts whereas MetJ uses the surface residues of helix A. This figure was prepared using the program MOLSCRIPT (Kraulis, 1991).
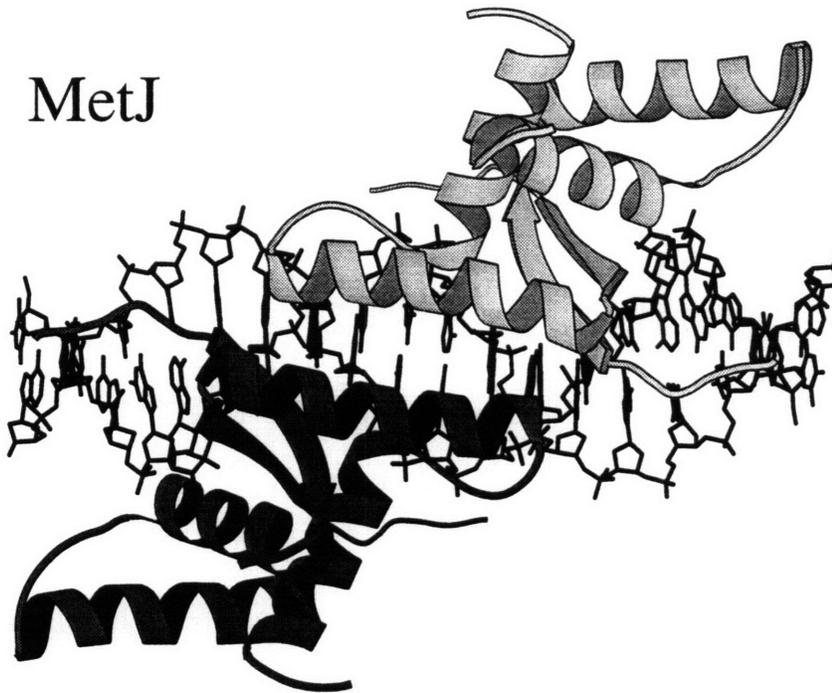
Arc

MetJ

159

Figure 5. Footprints of Arc-SL35 on the wild-type operator, the -3 operator, and a half-site operator. Left panel, the wild-type operator and the -3 operator. Right panel, the wild-type operator and the half-site operator. Both wild-type operator constructs have the operator in the same orientation. Note that the footprint of the -3 operator is more extensive than the footprint of the half-site operator when each is compared the the neighboring wild-type footprint, supporting the hypothesis that two dimers are bound to the -3 operator as was suggested by the DNA mobility shift results with the -3 operator shown in Chapter 2.

Footprinting of Wild-Type, -3, and Half-Site Operators

1. No Protein
2. SL35 Arc
3. SL35 Arc
4. No Protein

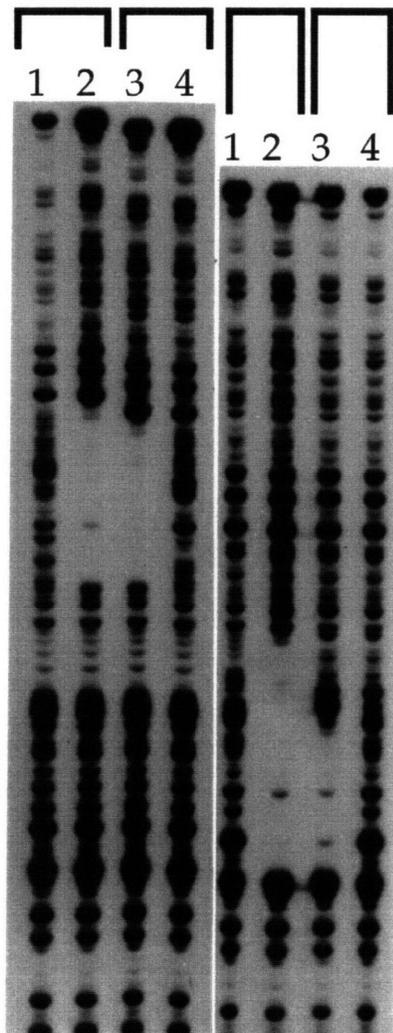WT   -3   WT Half

1   2   3   4

1 2 3 4

Figure 6. Selection construct controls for altered cooperativity Arc mutants. Plasmids bearing the *arc* gene under control of the P_{tac} promoter also contain one of the reporter constructs for Arc repression shown above. The wild-type *arc* operator positioned after the +1 start site of transcription can prevent transcription of the streptomycin-sensitive gene *rpsl* whereas the -3 operator cannot. Presumably, this lack of repression occurs because of the inability of the -3 operator to mediate cooperative Arc binding. A version of the -3 selection construct could be used to select for Arc mutants that can bind cooperatively to the -3 operator.

Selection Construct Controls for Altered Cooperativity Mutants

Wild-Type *arc* Operator Mediates WT-Arc Repression *In Vivo*

```
-35                        -10              wild-type operator
TTGACATAGGCACGTCACGTGCCTATATT_____     ATGATAGAAGCACTCTACTAT
AACTGTATCCGTGCAGTGCACGGATATAA                TACTATCTTCGTGAGATGATA      rpsl
                                        +1
```

-3 *arc* Operator Does Not Mediate WT-Arc Repression *In Vivo*

```
-35                        -10              -3 operator
TTGACATAGGCACGTCACGTGCCTATATT_____     ACGATAGAACTCTACTAT
AACTGTATCCGTGCAGTGCACGGATATAA                TGCTATCTTGAGATGATA      rpsl
                                        +1
```

Figure 7. Promoter clearance acceleration assays for variants of the P_{ant} promoter (see Chapter 3 for a full description of this assay). P_{ant}/both* and P_{ant}/10A* have the same sequences as their counterparts in Chapter 4, except there is a -8T→A bp change that decreases the rate of promoter clearance, enabling a better analysis of the effect of Arc on that rate. This graph demonstrates that 300 nM Arc-SL35 added after the formation of RNA polymerase open complexes can accelerate clearance from the P_{ant}/both* promoter as it can from the C promoter (described in Chapter 3) that contains only the left half site. However, Arc-SL35 cannot accelerate clearance from the P_{ant}/10A* promoter, even when Arc-SL35 is added before open complex formation and/or at a concentration of 1 μM (not shown).

Relative Transcription
after 1 minute with NTPS

400
300
200
100
0

P_ant/both*

P_ant/10A*

☒ No Arc
■ + Arc-SL35

165

$P_{ant} \longrightarrow$

-35   -10

GTTAGTGTATTGACAT|GATAGA|AGCAC|TCTACT|ATATTCTCAATAGGTCCA
CTTAAGCATAACTGTA|CTATCT|TCGTG|AGATGA|TATAAGAGTTATCCAGGT
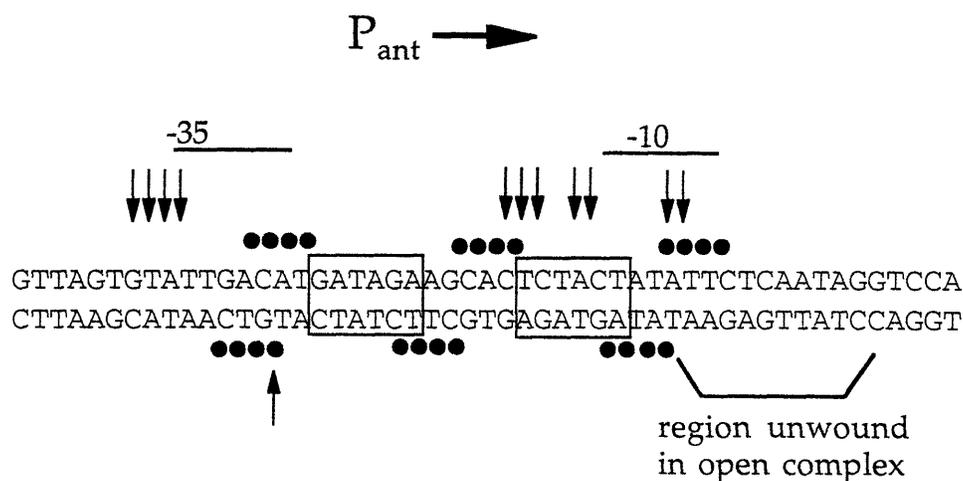
region unwound
in open complex
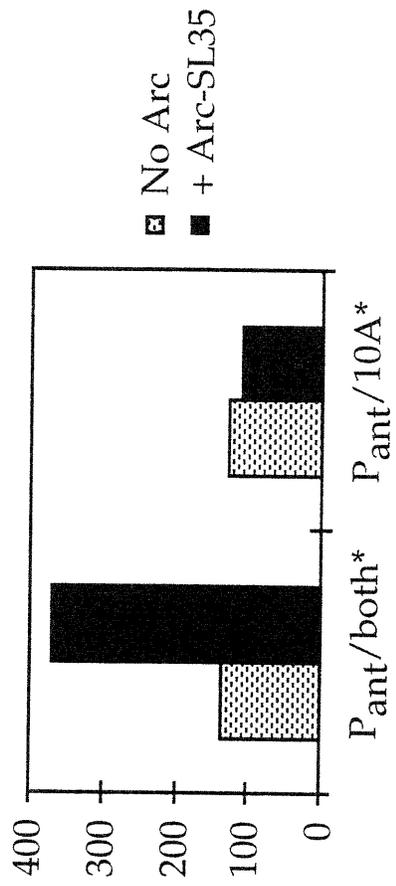
Figure 8. Depiction of overlap of Arc and RNAP polymerase contacts at the P_{ant} promoter. The -35 and -10 regions are marked. The arrows denote the positions of P_{ant} analogous to the positions of phosphate contacts made by RNA polymerase in the open complex at the lacUV5 and the T7 A3 promoters (Siebenlist *et al.*, 1980). The circles denote phosphate contacts made by Arc in the cocrystal structure (Raumann *et al.*, 1994). The six bp region where the specific base contacts are made by Arc are boxed. The -10 proximal site extensively overlaps the RNA polymerase binding site whereas the -35 site does not. Surprisingly, and despite the overlap of contacts, the results presented in Chapters 2 and 3 and the preliminary results shown in Figure 7 suggest that the -35 proximal site plays a greater role in Arc regulation of P_{ant} than the -10 proximal site.

Figure 9. P_{ant}/both* clearance assays with Arc mutants (see Chapter 2 for a full description of this assay). Arc-SL35 and Arc are used as controls that accelerate clearance two to three-fold when added to preformed open complexes at a concentration of 300 nM. Each of the other mutants has a change to alanine of one of the surface residues of helix A, a portion of the protein that could potentially be exposed to RNA polymerase in the open complex. Only Arc-RA23 does not have a strong increased rate of acceleration, and the effect of Arc-RA23 does not increase if the protein concentration is increased to 1 μM (not shown).

Clearance from P_{ant}/both* Promoter

Relative Transcription after 1 min with NTPs

200 — 150 — 100 — 50 — 0

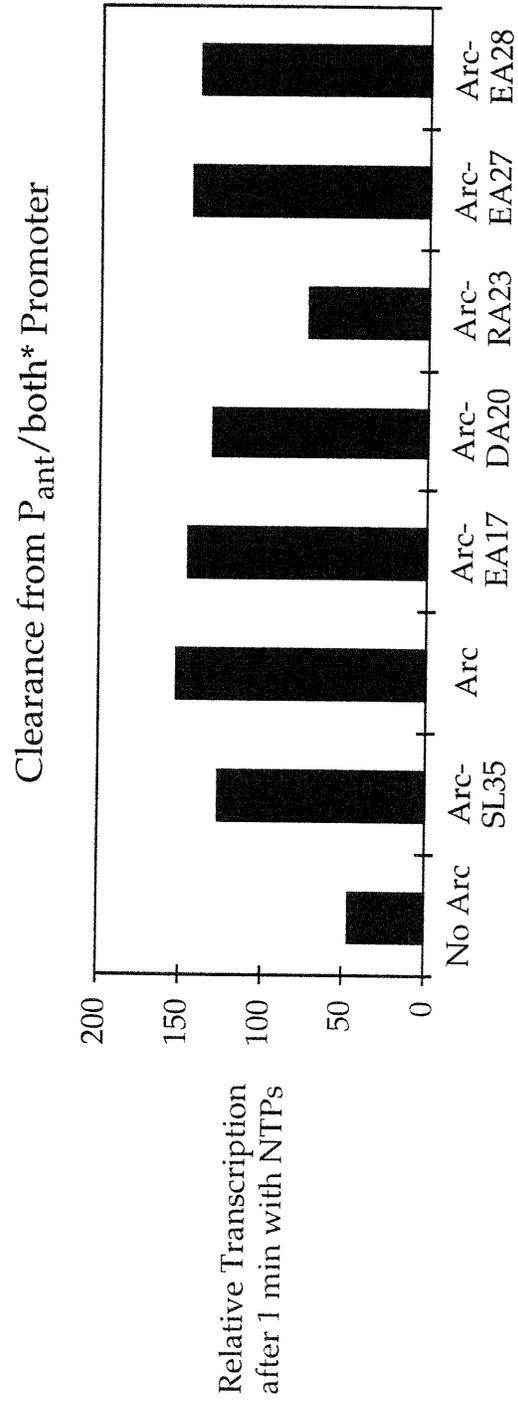No Arc — Arc-SL35 — Arc — Arc-EA17 — Arc-DA20 — Arc-RA23 — Arc-EA27 — Arc-EA28

Figure 10. The *immunity I* operon of bacteriophage P22. The top diagram illustrates the overall structure of the operon and the positions of the two promoters, P_{ant} and P_{mnt}, and the O_{arc} and O_{mnt} operators. Below is a close-up of the *arc* operator region showing the position of the arc operator with respect to the two promoters. The two *arc* binding sites are boxed. Note the different positions of the *arc* binding sites with respect to the -35 and -10 elements of each promoter.