A Structure/Function Analysis of the Interaction of the *E.coli* NusA protein with RNA polymerase, the phage λ N protein, and *nut* site RNA

by

Thien-Fah Mah

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Graduate Department of Molecular and Medical Genetics
University of Toronto

A Structure/Function Analysis of the Interaction of the *E.coli* NusA protein with RNA polymerase, the phage λ N protein, and *nut* site RNA

by Thien-Fah Mah

Department of Molecular and Medical Genetics in the University of Toronto

Submitted for a Doctor of Philosophy degree, 1999

Abstract

NusA is an *E.coli* protein that controls transcription elongation, termination and antitermination. In this thesis, I show that the functions of NusA in transcription are facilitated by interactions with RNA polymerase, RNA and the λ N protein. Use of a series of deletion constructs of NusA allowed me to identify specific regions of NusA involved in specific interactions and in various aspects of NusA function.

Genetic evidence suggested that NusA may interact with the *boxA* portion of the N-utilization site (*nut* site = *boxA*, *interbox*, and *boxB*). By constructing multiple nucleotide substitutions in the *nut* site, I showed that the identities of certain nucleotides at the 3' end of *boxA* and in the *interbox* were important for NusA to associate with an N-*nut* site complex. NusA association with RNA in the presence of N is presumably facilitated by its S1 and KH homology regions, two types of RNA-binding domains in

NusA. Elimination or mutation of the S1 homology region prevented the association of NusA with an N-*nut* site complex.

Using affinity chromatography experiments, I found that RNA polymerase bound equally well to an amino-terminal RNA polymerase-binding region in amino acids 1-137 and a carboxy-terminal RNA polymerase-binding region in amino acids 232-495 of NusA. By contrast, the α subunit of RNA polymerase only bound to the carboxy-terminal RNA polymerase-binding region of NusA. N protein also bound to a carboxy-terminal region of NusA, and both N and α allowed NusA to associate with RNA in a gel mobility shift assay. When the carboxy-terminal region of NusA was deleted in NusA (1-348), the loss of N-binding and α-binding ability did not abolish NusA function in termination and antitermination assays. This minimal functional NusA protein retained the KH and S1 homology regions and the amino-terminal RNA polymerase-binding region. Unlike full length NusA (1-495), NusA (1-416) could bind RNA on its own. These observations suggest that the carboxy-terminal region of NusA inhibits RNA binding and that this inhibition can be relieved by interaction with the λ N protein or the α subunit of RNA polymerase.

Table of Contents

Acknowledgements

I have learned an immeasurable amount about science from my supervisor, Jack Greenblatt. I started in the lab with very little experience. Through his signature style of supervision, I have grown into a scientist who feels comfortable defending her results and arguing points with peers.

I would like to thank the many people who have been in the Greenblatt lab who have filled the years with laughter, support, advice and good memories.

As fellow collaborators on the prokaryotic project, Joyce Li and Jeremy Mogridge provided valuable reagents and invaluable advice.

My committee members, Brenda Andrews, Rick Collins and Barbara Funnell, have been absolutely wonderful. I could not have asked for more caring and committed individuals to help guide me through graduate school.

This degree has taken a long time to complete. It would have been easy to get lost along the way, but I have been fortunate to have met many amazing people. This has included people both in the Department, as well as ones who are out working in the wild world, who kept me grounded with their friendship.

My family has been there all the way. Thanks for believing in me.

Last, but not least, I would like to thank Raja Bhattacharyya for unwavering love and support.

List of Figures

Chapter 1

Chapter 2

Chapter 3

List of Tables

Chapter 4

Introduction

In this thesis, I describe experiments performed to elucidate the role of the *Escherichia coli* NusA protein in the control of transcription during elongation and termination as well as during transcriptional antitermination mediated by the bacteriophage λ N protein. Therefore, I begin with a detailed description of transcription by *E.coli* RNA polymerase.

Transcription in *E.coli*

Transcription in *Escherichia coli* is carried out by a single multi-subunit RNA polymerase. This process can be separated into three distinct stages: initiation, elongation and termination. Initiation is characterized by RNA polymerase binding to promoter DNA, conversion of an unstably bound RNA polymerase to a stably bound one, and the formation of the first few phosphodiester bonds of the nascent transcript. Elongation is the stage where nucleotides are added to the growing chain of nascent RNA at a rate dictated by the template DNA sequence and the factors associated with RNA polymerase during this stage. Termination occurs when a signal in the template DNA is relayed to RNA polymerase and the entire complex of RNA polymerase, template DNA, nascent RNA, and any associated factors, dissociates. Antitermination prevents the recognition of termination signals by RNA polymerase, resulting in read-through of terminators. Below, I discuss additional control points that exist at all three stages.

Initiation

The enzyme responsible for stable transcript elongation is the core RNA polymerase, a heteromeric protein which consists of one copy each of β (151 kD) and β' (155 kD) subunits and two copies of the α (36 kD) subunit. Core RNA polymerase can associate with template DNA non-specifically. In contrast, specific initiation of transcription is carried out by the holoenzyme form of RNA polymerase (Burgess *et al.*, 1969). This form consists of the core enzyme associated with one of several σ subunits, which can specifically recognize promoter DNA sequences.

The first step of initiation involves RNA polymerase holoenzyme binding to promoter DNA. The optimal promoter recognized by the major σ factor, known as σ^{70}, has two elements: the -10 and -35 elements, and these are separated by a spacer of about 17 base pairs (Hawley and McClure, 1983; Harley and Reynolds, 1987; Lisser and Margalit, 1993). The -10 element is located 10 base pairs upstream of the transcriptional start site. The consensus sequence for this element is TATAAT. Similarly, the -35 element is generally found 35 base pairs upstream of the transcriptional start site and has the consensus sequence TTGACA. These two elements make up the core promoter and act as binding sites for the σ^{70} subunit of the holoenzyme (Losick and Pero, 1981). Early experiments showed that σ^{70} cross-links to the -3 position of the lacUV5 promoter (Simpson, 1979), and mutations in *rpoD*, the gene encoding σ^{70}, alter the sequence-specificity of the holoenzyme (Gardella *et al.*, 1989; Siegele *et al.*, 1989). The interaction of the *rpoD* mutants with altered promoter sequences suggested a relationship between two regions of σ^{70} that are conserved among σ factors and either the -10 or -35 consensus sequences. Although intact σ^{70} subunit does not bind DNA (Burgess *et al.*, 1969; Zillig *et al.*, 1970), an amino-terminally truncated form of σ^{70} can bind DNA on its own (i.e. in the absence of the RNA polymerase core enzyme), suggesting that the amino-

terminal region of σ^{70} acts as an autoinhibitory domain to prevent DNA binding (Dombroski *et al.*, 1992; Dombroski *et al.*, 1993). Thus, interactions within the RNA polymerase holoenzyme prevent autoinhibition of DNA binding by the amino-terminal region of σ^{70} and, in this way, specific promoter recognition by the holoenzyme becomes possible.

A third promoter element upstream of the -35 element was recently discovered in the promoter for an *rrn* gene. This sequence, the UP element, is found in the promoters of a subset of genes. It is located roughly 55 base pairs upstream of the transcriptional start site in these promoters (Rao *et al.*, 1994). Although there is no strict consensus sequence for the UP element, it is characterized by AT-richness. UP elements can affect the strength of a promoter by as much as thirty fold. It is postulated that the increase in promoter strength is provided by a direct interaction between the UP element and the α subunit of RNA polymerase (Ross *et al.*, 1993; Murakami *et al.*, 1996).

The initial RNA polymerase-promoter DNA complex is unstable and is called the closed complex. This closed complex has at least two conformational states. The initial closed complex has a footprint on the DNA from -55 to -5 and can only be studied at low temperature (Hofer *et al.*, 1985; Cowing *et al.*, 1989; Schickor *et al.*, 1990; Mecsas *et al.*, 1991). As the temperature is increased, this initial complex becomes an isomerized closed complex that has a footprint from -55 to +20, thus encompassing the start site of transcription at position +1 (Cowing *et al.*, 1989; Schickor *et al.*, 1990; Mecsas *et al.*, 1991). Further conformational changes within this closed complex lead to isomerization into a more stable open complex in which part of the region of DNA to which the polymerase is bound is melted, thereby exposing the nucleotides to be used as the template for the transcription reaction. The entire process of initiation can be summarized in an equation that reflects the rates at which these complexes form:

$$R+P \underset{K_B}{\leftrightarrow} RP_c \underset{k_f}{\leftrightarrow} RP_o$$

where R=RNA polymerase, P=the promoter, RP_c=the closed complex, RP_o=the open complex, K_B=the equilibrium constant that describes the binding of RNA polymerase to the promoter DNA and k_f=the rate constant for the isomerization between the closed and open complexes.

Once the open complex has formed, a process called abortive initiation can occur. This process results in the cyclical creation of short transcripts from 2 to 9 nucleotides in length without the loss of the specific RNA polymerase-promoter contacts that characterize the open complex (Johnston and McClure, 1976; Carpousis and Gralla, 1980). Initiation is complete when RNA polymerase transcribes beyond the first 8 to 10 nucleotides of initial sequence and escapes from the promoter.

Early studies in regulation of transcription focused on control at the level of initiation. Activators and repressors use different mechanisms to affect steps outlined in the equation above. For instance, the λ repressor protein, cI, activates the λ promoter P_{RM} by increasing the rate constant of the isomerization step. The cAMP-binding protein, CRP, activates the *E.coli lac* promoter by increasing the binding constant for RNA polymerase at this promoter (Malan *et al.*, 1984; Li *et al.*, 1994). Interestingly, another λ protein, cII, activates the λ promoter P_{RE} by increasing both the K_B and the k_f (Shih and Gussin, 1984). Repressors can work by simply binding to the promoter and sterically blocking access of RNA polymerase. In addition, repressors can affect the kinetics of initiation. For instance, cI inhibits the formation of an open complex at the λP_R promoter, while the lac repressor modifies the open complex such that it cannot escape the promoter (Hawley *et al.*, 1985; Lee and Goldfarb, 1991).

These effects on initiation can result from a direct interaction of the activator or repressor with one of the subunits of RNA polymerase. The target of cI at λP_{RM} is σ^{70} (Li *et al.*, 1994), while the carboxy-terminal domain (CTD) of α is the target of regulatory proteins such as the CAP protein and the FIS protein (Li *et al.*, 1994; Bokal *et al.*, 1997). FIS is a

DNA-binding protein that most notably increases the amount of transcription at the already strong *rrnB* P1 promoter (Ross *et al.*, 1990). This promoter is regulated by an UP element-α-CTD interaction and FIS may facilitate α-CTD binding to the UP element through interaction with α-CTD (Bokal *et al.*, 1997).

Another level of control of initiation of transcription is represented by σ factor-binding. There are two families of *E.coli* σ factors: the σ^{54} family is required for nitrogen-regulated gene expresssion and the σ^{70} family of primary σ factors is required for vegetative growth and response to environmental conditions and physiological states (Gross *et al.*, 1992). Association of these factors with the core enzyme directs the holoenzyme to conserved sequences in promoter DNA specifically recognized by individual σ factors. These σ factors, in turn, are regulated by the action of anti-σ factors that bind and sequester specific σ factors until they are needed. Anti-σ factors can also inhibit early steps in elongation. FlgM is a flagellar anti-σ factor from *Salmonella typhimurium*. In addition to binding to its cognate σ factor, σ^{28}, it can bind a σ^{28}-holoenzyme molecule and dissociate it (Chadsley *et al.*, 1998). Another strategy used by anti-σ factors is to prevent association of the σ subunit with DNA once it is part of the holoenzyme. The bacteriophage T_4 protein, AsiA, binds the σ^{70}-holoenzyme once is has already formed but occludes the –35 recognition site (Severinova *et al.*, 1998).

Elongation

The transition from a promoter-bound initiation complex to an elongation complex capable of rapid RNA chain elongation is characterized by the loss of the σ subunit, loss of

promoter-specific contacts, a change in the size of the RNA polymerase footprint and an increase in the stability of the complex to a variety of conditions, including high salt concentrations (Hansen and McClure, 1980; Carpousis and Gralla, 1985; Straney and Crothers, 1985; Levin *et al.*, 1987; Arndt and Chamberlin, 1990). This increase in stability has facilitated the purification of the elongation complex, resulting in the elucidation of some of its structural features.

As described below, three models have been proposed to explain how RNA polymerase transcribes DNA: the monotonic, inchworm and backsliding models. In the monotonic model the elongation complex was proposed to be relatively inflexible and to move downstream on the template in single steps. Therefore, addition of a nucleotide to the growing RNA chain results in the forward movement of the complex by one nucleotide, the maintenance of a 17 nucleotide bubble of unpaired DNA by the actions of a downstream "unwindase" and an upstream "rewindase", and the maintenance of the 8-12 nucleotide DNA-RNA hybrid (Kumar and Krakow, 1975; Gamper and Hearst, 1982; Hanna and Meares, 1983; Kainz and Roberts, 1992). This model suggested that the characteristics of the complex, such as footprint size and stability to salt and dissociation, would remain constant. However, further studies of many complexes isolated from different positions on a given template seemed to contradict the monotonic model, since some complexes had apparently different sized DNAse I footprints, different mobilities in a polyacrylamide gel, different stabilities to dissociation, and different potentials to resume transcription after being halted (Krummel and Chamberlin, 1992b; Krummel and Chamberlin, 1992a). These results suggested a more dynamic complex, one that is sensitive to template position. Thus, based on these and other observations, Chamberlin proposed the inchworm model, which

postulated that elongating RNA polymerase is flexible and could compress and expand, allowing movement of its upstream end independent of its downstream end (Chamberlin, 1992).

The discontinuity of elongation complex movement associated with inchworming was postulated as a result of nascent RNA chain extension that seemed to remain continuous even at template positions where the front end of RNA polymerase no longer moved forward. The addition of nucleotides to the nascent RNA chain seemingly caused the active center of RNA polymerase to move forward on the DNA template, without the corresponding forward movement of the front end of the elongation complex, creating internal strain that was seemingly relieved by a forward "leap" of the complex at certain positions on the template (Nudler *et al.*, 1994). Central to this model was the idea that the template conveyed information to create this inchworming effect. Comprehesive studies at every base position on a specific template showed that RNA polymerase advanced in a monontonic mode at most DNA sequences, but that, at certain DNA sequences, most notably at pause and termination sites, RNA polymerase switched to an inchworming mode (Nudler *et al.*, 1994; Nudler *et al.*, 1995; Wang *et al.*, 1995).

The backsliding model reinterpreted the above data to explain the movement of RNA polymerase along the template DNA in a different way. In this model, RNA polymerase is envisioned to oscillate between inactive, arrested complexes that have slid backwards on the DNA template and active, elongation-competent complexes (Komissarova and Kashlev, 1997b; Komissarova and Kashlev, 1997a). Weakening of the DNA-RNA hybrid increased backsliding, while strengthening the hybrid decreased the incidence of backsliding, suggesting that the signal to enter the oscillation cycle is based on the relative strength of the

DNA-RNA hybrid (Nudler *et al.*, 1997). Two lines of evidence support this model. First, the DNase I footprint of a halted elongation complex moved upstream, with the front and back ends of the complex backing up in concert with each other; second, the catalytic center of RNA polymerase also moved backwards, so that the 3'-terminus of the RNA was extruded from the active site (Komissarova and Kashlev, 1997a; Komissarova and Kashlev, 1997b; Nudler *et al.*, 1997).

The apparent contraction and relaxation of RNA polymerase postulated by the inchworming model was most likely due to misinterpreting the footprinting of only the front end of the complex at different points in its oscillation cycle. As well, it was originally assumed that the active site of RNA polymerase would remain attached to the 3' end of the nascent RNA, leading to the false assumption that the position of the catalytic centre was defined by the position of the 3' end of the RNA. However, this was formally disproved using an RNA polymerase in which the active site Mg^{2+} was replaced by Fe^{2+}, causing localized cleavage of nucleic acids (Zaychikov *et al.*, 1996). In the arrested complexes, Fe^{2+}-mediated cleavage of the transcript upstream of its 3' end strongly suggested that the active site had indeed moved upstream and that the cleaved RNA represented extruded RNA (Nudler *et al.*, 1997). The backsliding model is currently the accepted model of elongation and provides the framework for models of termination and antitermination (see below).

The elongation complex is a stable structure at most positions along the DNA template. The nature of the molecular interactions that provide stability and processivity for the elongation complex are of interest because loss of these interactions has implications for the regulation of transcription. The monotonic model of elongation suggested that the 8-12 base pair DNA-RNA hybrid confers most of the stability necessary for the synthesis of long

transcripts, whereas DNA-protein and RNA-protein interactions contribute only a small measure to the stability (Yager and von Hippel, 1987; Yager and von Hippel, 1991). Disruption of the hybrid would lead to destabilization of the complex and events such as arrest, pausing and termination. However, subsequent studies suggested that the DNA-RNA hybrid may not be the only stabilizing factor for the complex. For instance, one study reported that RNA polymerase can bind the 3' end of RNA in the absence of a DNA template providing evidence for an RNA-binding site within RNA polymerase that contributes enough stability to allow for functions that normally occur in an elongation complex (Altmann *et al.*, 1994). The RNA in the complex can be cleaved and extended, suggesting that the complex is at least somewhat stable. To explain these observations, the inchworm model had postulated that separate RNA- and DNA-binding sites hold the complex together to eliminate the need for a long hybrid (Chamberlin, 1992).

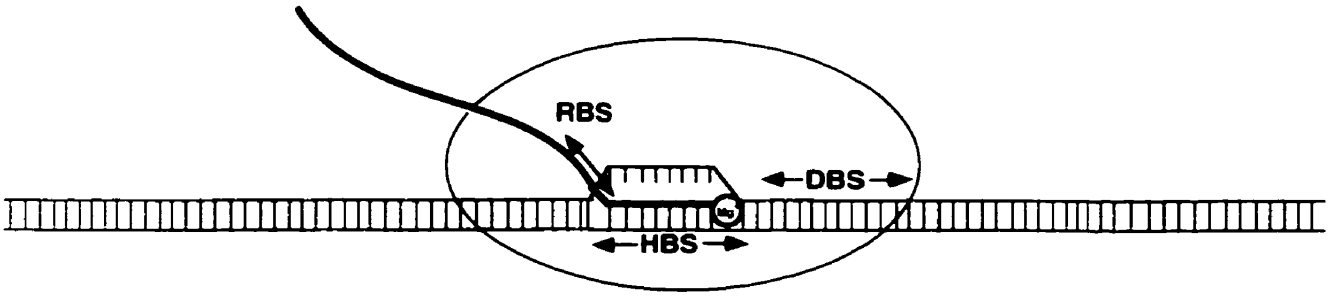


Figure 1. The Structural Features of the Elongation Complex.
A schematic of the features of the elongation complex important for stability of the complex. DBS is the DNA-binding site, RBS is the RNA-binding site, HBS is the hybrid-binding site and Mg^{+} represents the active site.

The present view of the structural aspects of the elongation complex is shown in Fig. 1. There are three main features in RNA polymerase that contribute to stability of the complex: the hybrid-binding site, the double-stranded DNA-binding site and the single-stranded RNA-binding site (reviewed in Nudler, 1999). A DNA-RNA hybrid 8-9 nucleotides in length has been confirmed by cross-linking RNA to DNA and by protection from cleavage by different RNases (Nudler *et al.*, 1997; Komissarova and Kashlev, 1998). The hybrid-binding site in RNA polymerase maintains the RNA and DNA in the elongation complex under low salt conditions (Nudler *et al.*, 1996). The double-stranded DNA-binding site interacts with 10 base pairs of duplex DNA just downstream of the active site of RNA polymerase and is responsible for the stability of the complex in high concentrations of salt. Finally, the single-stranded RNA-binding site adds more stability against high salt concentrations. The observation that RNA that is upstream of the DNA-RNA hybrid is important for stability of the elongation complex provides the framework for one of the current models of transcription termination (see below).

The interactions that hold the elongation complex together presumably serve to maintain the processivity of transcription. Once it has initiated transcription, RNA polymerase may maintain all these contacts until a terminator is reached. However, there are sites along the template that can cause RNA polymerase to pause on the DNA for varying amounts of time or to arrest transcription completely. Arrest sites are defined as sites where RNA polymerase stops transcription irreversibly *in vitro*. The elongation complex remains intact at such sites but is unable to recommence elongation, even when supplied with all the required nucleotides, because the transcript is out of register with the active site (Arndt and

Chamberlin, 1990; Komissarova and Kashlev, 1997b). This situation has not been observed *in vivo*, leading to the hypothesis that additional factors may reactivate arrested complexes *in vivo*. The cleavage factors GreA and GreB are candidates to fulfill this function.

GreA was identified in a screen for suppressors of a temperature-sensitive mutation in the β subunit of RNA polymerase (Sparkowski and Das, 1990) and was subsequently shown to induce cleavage of the 3' end of the nascent RNA in arrested elongation complexes (Borukhov *et al.*, 1992). This finding was rapidly followed by the discovery of a second cleavage factor, GreB, and the characterization of the properties of both proteins (Borukhov *et al.*, 1993). Although RNA polymerase itself possesses intrinsic cleavage activity (Orlova *et al.*, 1995), GreA induces removal by RNA polymerase of 2 to 3 nucleotides from the 3' end of the nascent transcript. GreB induces cleavage, also by RNA polymerase, of up to 9 nucleotides from the 3' end (Borukhov *et al.*, 1993). According to the backsliding model of elongation, an arrested complex is one in which the 3' end of the nascent RNA is out of register with the active site (Komissarova and Kashlev, 1997b), and the Gre factors probably induce cleavage of the extruded RNA by the active site (Nudler *et al.*, 1997).

RNA polymerase pauses for various lengths of time at pause sites, but then can resume transcriptional elongation after escape from the pause. It has been suggested that pausing is a way to maintain the tight coupling between transcription and translation that occurs in bacteria. One clear example in which pausing has this function is attenuation, where regulation of a single terminator occurs in response to a metabolic signal (Landick and Turnbough, 1992; Landick *et al.*, 1996). Attenuation often regulates the transcription of amino acid biosynthetic operons and, in these cases, there is a rho-independent terminator located between the promoter and the first gene of the operon. A pause site is located

upstream of the terminator and, since translation initiates upstream of the pause, tight coupling of transcription and translation is ensured. Once transcription resumes, alternate RNA secondary structures have the potential to form, one leading to termination and the other to terminator read-through, and this decision is based on the availability of the relevant amino acid. If the concentration of the amino acid is low, the ribosome stalls within a critical leader peptide, allowing the antiterminator hairpin to form. If the concentration of the amino acid is high, translation continues through to the stop codon and formation of the terminator hairpin is mostly favoured.

At one kind of typical pause site, a hairpin structure forms in the nascent RNA whose 3' end is separated from the 3' end of the RNA in the paused elongation complex by 10 or 11 nucleotides. Whether RNA polymerase will recognize a pause site and how long it will remain at the site is affected by the nascent RNA hairpin, the RNA located between the hairpin and the 3' end of the transcript, the identity of the incoming nucleotide and the DNA downstream of the active site (Chan and Landick, 1993; Levin and Chamberlin, 1987; Chan and Landick, 1994).

As RNA polymerase approaches a pause site, it undergoes an alteration in conformation and begins to backslide (Wang *et al.*, 1995). As transcription elongation slows due to the oscillation cycle, the pause hairpin has time to form. Furthermore, the pause hairpin is thought to block further backsliding by allowing the formation of a stable, paused complex. Recent observations suggest that there is a direct interaction between the RNA hairpin and RNA polymerase and that this interaction stabilizes the pause conformation (Chan *et al.*, 1997; Wang *et al.*, 1997; Artsimovitch and Landick, 1998).

In addition to signals derived from the DNA and RNA, the length of time that RNA polymerase remains at a pause site can be influenced by external factors. NusA associates with RNA polymerase after escape from the promoter and loss of σ^{70} (Greenblatt and Li, 1981a). As part of the elongation complex, NusA decreases the elongation rate of RNA polymerase. This phenomenon occurs in two ways: NusA increases the length of time that RNA polymerase remains at certain pause sites (Kassavetis and Chamberlin, 1981; Kingston and Chamberlin, 1981; Farnham *et al.*, 1982; Lau *et al.*, 1983; Landick and Yanofsky, 1984; Landick and Yanofsky, 1987; Levin and Chamberlin, 1987; Faus *et al.*, 1988) and it increases the substrate dissociation constant (Schmidt and Chamberlin, 1984). It is not clear how NusA influences the response of RNA polymerase to pause signals, but it is postulated that NusA interacts with the pause hairpin to stabilize it (Landick and Yanofsky, 1987; Faus *et al.*, 1988). This model is based on the observation that the RNase T_1 digestion pattern of an isolated pause hairpin is altered by the presence of the *E.coli* NusA protein (Landick and Yanofsky, 1987). These effects on the RNA polymerase elongation rate can influence gene regulation and may be necessary to couple transcription and translation (Ruteshouser and Richardson, 1989; Zheng and Friedman, 1994).

NusG is another elongation factor that interacts directly with RNA polymerase (Li *et al.*, 1992). However, its effect on transcription elongation is opposite to that of NusA. NusG increases the elongation rate of RNA polymerase *in vivo* and *in vitro* (Burova *et al.*, 1995).

Termination

Termination occurs when a signal in the DNA template causes RNA polymerase to stop the addition of nucleoside triphosphates and to dissociate from the DNA. In *E.coli*,

there are two types of termination signals, ones that do not require accessory factors in order to function and others that require the termination factor rho. The former type of terminator is called an intrinsic, simple, or rho-independent terminator, whereas the latter is called a rho-dependent terminator.

Rho-independent termination

Intrinsic terminators have some features in common with pause sites. They typically contain a GC-rich stem loop structure located 7 to 8 nucleotides upstream of the termination site. However, termination sites also contain at least 3 uridine residues downstream of the terminator hairpin and this feature is not present in pause sites.

An early model for termination at intrinsic termination sites postulated that the formation of the terminator hairpin destabilized the DNA-RNA hybrid and that this destabilization, as well as destabilizing effect of the rU-dA base pairs in the DNA-RNA hybrid, caused dissociation of the nascent transcript. Thus, in the absence of the uridine stretch, the pause hairpin would be sufficient to cause a break in elongation at a site located further downstream, but would not be sufficient to cause dissociation of the complex.

The initial view was that elongation and termination were competitive processes, and that at every position on the template there was a thermodynamic barrier to dissociation of the DNA-RNA hybrid (von Hippel and Yager, 1991; Yager and von Hippel, 1991). The energy required to maintain elongation was derived from the sum of the energy lost in the formation of the transcription bubble and the energy gained in the formation of the DNA-

RNA hybrid. This energy barrier was proposed to be higher at termination sites than at elongation sites, favouring dissociation of the complex only at a terminator.

However, recent experiments have suggested an alternative model. Instead of only destabilizing the DNA-RNA hybrid, it has been proposed that the terminator hairpin also disrupts the protein-RNA interaction within the RNA-binding site (Yarnell and Roberts, 1999). Thus, the hairpin would effectively extract the RNA from the transcription complex. Protein-RNA cross-linking experiments are consistent with this observation: formation of a terminator hairpin alters the pattern of cross-linking between RNA polymerase and the RNA in the RNA-binding site (Gusarov and Nudler, 1999). Hairpin formation would also lead to the unwinding of the DNA-RNA hybrid and transcription bubble collapse. The second termination determinant, the poly-uridine stretch, stalls RNA polymerase to allow time for the formation of the hairpin. This pause may occur when the weak rU-dA base pairs in the DNA-RNA hybrid allow RNA polymerase to slide backwards and disengage its catalytic centre from the 3' end of the RNA. If the hairpin is deleted, no termination is observed but the pause still occurs (Nudler *et al.*, 1995; Yarnell and Roberts, 1999); likewise, if the rate of elongation is very low, or the complex is nucleotide-starved and static, termination occurs in the absence of the uridine-rich region (Yarnell and Roberts, 1999). Thus, the new model postulates that the elongation complex pauses because of the poly-uridine tract in the nascent RNA. The RNA hairpin then forms and the complex dissociates because of the instability created by the loss of RNA-protein contacts in the RNA-binding site and the unwinding of the DNA-RNA hybrid (Gusarov and Nudler, 1999; Yarnell and Roberts, 1999).

Although no additional factors are required for termination at intrinsic terminators, NusA can increase the efficiency of termination at many of these sites (Greenblatt *et al.*,

1981; Farnham *et al.*, 1982; Grayhack *et al.*, 1985; Schmidt and Chamberlin, 1987). The mechanism by which NusA increases the termination efficiency is not well understood. In this thesis, I show that the ability of NusA to interact with RNA polymerase and with RNA is likely to be important for this process.

Rho-dependent termination

Rho was identified as a factor that causes transcription termination and the release of the RNA in a transcription reaction *in vitro* (Roberts, 1969). Rho is a hexameric protein that binds RNA (Finger and Richardson, 1982). It possesses DNA-RNA helicase activity and RNA-dependent ATPase activity (Lowery-Goldhammer and Richardson, 1974; Brennan *et al.*, 1987). These properties suggested that rho binds the nascent RNA upstream of the elongating RNA polymerase, translocates along the RNA toward the RNA polymerase powered by its ATPase activity, and disrupts the elongation complex by using its helicase activity (Platt and Richardson, 1992).

Sequences that are important for the generation of rho terminated transcripts have been identified by testing for the effects on rho-dependent termination of either deleting portions of the upstream DNA or using DNA oligonucleotides that hybridize with various portions of the upstream RNA. In this way, two *rut* (*r*ho-*ut*ilization) sites were identified upstream of the rho-dependent λ terminator tR1 (Chen *et al.*, 1986; Chen and Richardson, 1987). These sites, called *rutA* and *rutB*, are characterized by a relatively high cytosine content and very little secondary structure (Chen *et al.*, 1986; Chen and Richardson, 1987). In fact, examination of all defined rho-dependent termination sites, as well as of cryptic rho-

dependent termination sites within the *his* operon, led to the conclusion that rho-dependent terminators consist of cytosine-rich and guanosine-poor regions with little secondary structure upstream of defined rho-dependent 3' endpoints (Morgan *et al.*, 1985; Alifano *et al.*, 1991). In contrast, no defined sequences have been identified for the transcript 3' endpoints (Richardson and Richardson, 1996), which are often sites where RNA polymerase pauses (Lau *et al.*, 1982; Lau *et al.*, 1983). This finding suggests that slowing RNA polymerase allows rho to reach the elongation complex downstream of the site where rho initially interacts with the nascent transcript. This view is supported by the observation that there is an inverse relationship between the elongation rate of RNA polymerase and the efficiency of rho-dependent termination (Jin *et al.*, 1992). However, there is no direct relationship between the strength of a pause site and the ability of rho to cause termination, suggesting that the conformation of the paused complex is the target of rho (Richardson and Richardson, 1996). This view is supported by the finding that RNA polymerase probably slides backwards at pause and termination sites (Nudler *et al.*, 1995; Wang *et al.*, 1995).

The cryptic rho-dependent terminators in the *his* operon were identified through mutations that caused polarity and subsequent rho-dependent termination of transcription (Alifano *et al.*, 1991). Polarity is observed when premature termination of translation of an early gene in an operon results in the reduction of transcription of distal genes in the operon. The observation that *rho* mutants suppress polarity in the *gal* operon (Das *et al.*, 1976) suggested that rho binds to untranslated RNA (Adhya *et al.*, 1974).

Depletion of NusG, originally identified as a bacterial protein important for λ N-mediated antitermination (Li *et al.*, 1992)(see below), reduced rho-dependent termination at most terminators that were tested (Sullivan and Gottesman, 1992), suggesting that NusG is

important for rho function *in vivo*. Subsequent *in vitro* experiments confirmed that, at some terminators, under conditions where rho does not function well on its own, NusG enhances the termination efficiency (Burns and Richardson, 1995) by binding weakly to rho and RNA polymerase (Mason and Greenblatt, 1991; Li *et al.*, 1992; Li *et al.*, 1993). These interactions may aid rho factor in locating the transcription complex (Li *et al.*, 1993). Alternatively, the interaction between rho and NusG may stabilize the binding of both proteins to the elongation complex. In support of this model, NusG does not associate well with the elongation complex unless rho is present and bound to the nascent RNA, and NusG slows the off-rate of rho from stalled elongation complexes (Nehrke and Platt, 1994). Furthermore, NusG shifts the positions of the endpoints of the terminated transcripts in the 5' direction (Li *et al.*, 1993; Nehrke *et al.*, 1993). Perhaps, NusG enables rho to bind more quickly to its high affinity binding sites in the nascent RNA as they become available during transcription.

Regulation of Elongation by the Bacteriophage λ N protein

Antitermination is an important form of regulation during the elongation phase of transcription. As a result of this process, RNA polymerase transcribes through terminators on the DNA template. The best studied systems of antitermination are found in bacteriophage λ. When λ infects an *E.coli* cell, transcription of the phage genome is carried out by the host RNA polymerase. Initiation of transcription occurs at P_R and P_L, the promoters of the rightward and leftward phage early operons, respectively (see Fig. 2). The initial transcripts are short because the rho-dependent terminators, tR_1 and tL_1, are located downstream of the first genes in the two operons. However, the *N* gene transcript is present

in the initially transcribed P_L RNA and it is the presence of the N protein that is key to the transformation of RNA polymerase into a termination-resistant form. The action of N is required for the expression of all downstream genes in both operons and, in the particular case of the rightward operon, for the positive regulator of the late genes, the Q protein.

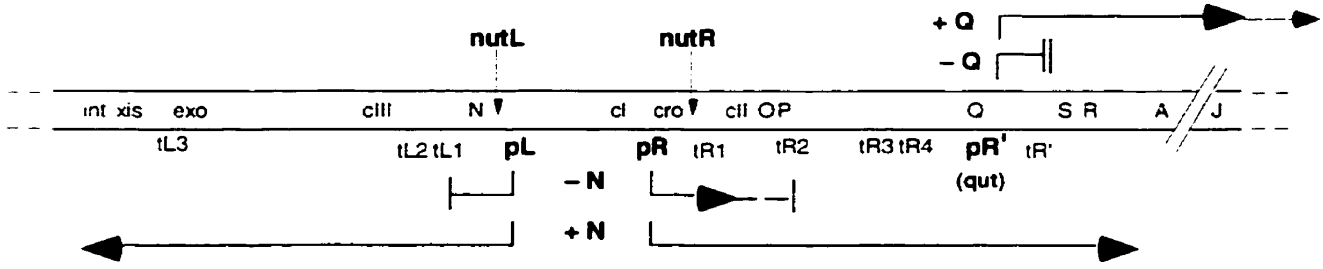


Figure 2. Regulatory elements of the λ genome

Processive N-mediated antitermination, or antitermination that persists through many terminators and for many kilobases of sequence, requires the presence of a cis-acting element, called the *nut* site (*N-ut*ilisation), that is located in both of the early operons. Processive antitermination also requires several host factors, called Nus (*N-ut*ilization substance), that were identified in screens for host mutations that prevented the growth of wild type N-dependent phage, but not the growth of N-independent mutant phage. N-mediated antitermination prevents termination by RNA polymerase at both rho-dependent and rho-independent terminators.

The *nut* site

The temperate bacteriophages λ, P22 and ϕ21 have very similar genome organizations. However, whereas the regulatory proteins of these phages have analogous functions, the sequences have diverged so that a regulatory protein from one phage often cannot function in another. Because the phage DNA sequences have limited homology, it has been possible to create hybrid phages to analyze relationships between particular regulatory proteins and their sites of action. The requirement for a cis-acting sequence for regulation by N was first postulated when it was shown that λ N could not function in other lambdoid phages (Friedman *et al.*, 1973; Friedman and Ponce-Campos, 1975). Mutations in non-coding sequences that prevented the functioning of N then led to the discovery of *nutL* and *nutR* (Rosenberg *et al.*, 1978; Salstrom and Szybalski, 1978). The *nut* sites are located between the promoters and the first terminators of both early operons. Subsequent comparison of the *nut* sequences of the related coliphages λ, ϕ21 and P22, as well as mutational studies, led to the identification of two separate elements of the *nut* site, the 12 nucleotide *boxA* element and the 15 nucleotide *boxB* element which can form a hairpin structure in the nascent RNA (Franklin, 1985a; Olson *et al.*, 1982)(Fig. 3). The sequence between *boxA* and *boxB* is called the *interbox*. Since the cloning of *nutR* downstream of a heterologous promoter in the absence of P_R did not alter the ability of the *nut* site to support antitermination by N, it was concluded that the *nut* site itself is sufficient for antitermination (de Crombrugghe *et al.*, 1979).

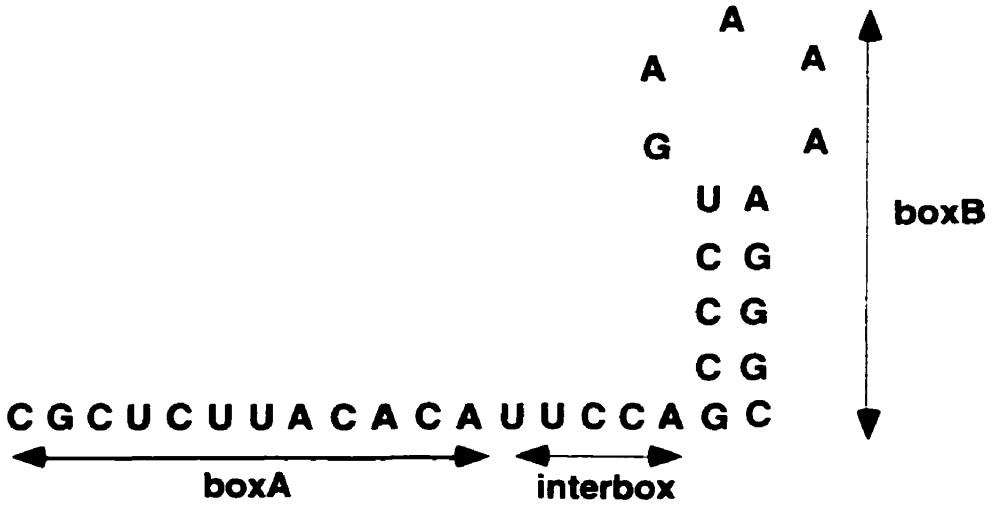


Figure 3. The λ *nutR* site

Various experiments have defined *nut* sequences required for antitermination. For example, particular mutations in any nucleotide in the loop of the *boxB* hairpin completely abolished antitermination (Doelling and Franklin, 1989). Deletion of eight nucleotides of *boxA* sequence left antitermination unaffected in one study, suggesting to the authors that *boxA* is not required for antitermination (Zuber *et al.*, 1987). However mutations in both *boxA* and *interbox* nucleotides affect the functioning of the *nut* site *in vitro* (Chapter 2). Moreover, the loss of *boxA* rendered antitermination thermosensitive *in vivo* in another study (Peltz *et al.*, 1985). Thus, *boxA* is important for antitermination while *boxB* is essential.

Several lines of evidence have suggested that the functional form of the *nut* site is RNA. First, a frameshift mutation that allowed translation through the *nut* site prevented antitermination (Olson *et al.*, 1982). Furthermore, overexpression of *boxA* RNA decreased antitermination *in vivo* (Friedman *et al.*, 1990), and high concentrations of RNase released N

from N-modified transcription complexes *in vitro* (Horwitz *et al.*, 1987). More direct evidence has been provided by gel retardation experiments in which N and the Nus factors bound to and retarded the electrophoretic mobility of the *nut* site RNA (Chattopadhyay *et al.*, 1995a; Mogridge *et al.*, 1995; Tan and Frankel, 1995; Mogridge *et al.*, 1998b) and by the observation that N and the Nus proteins can protect the *nut* site from degradation by RNases *in vitro* (Nodwell and Greenblatt, 1991).

boxB is a 15 nucleotide hairpin with a 5 nucleotide loop and a 5 base pair double-stranded stem (Chattopadhyay *et al.*, 1995a; Su *et al.*, 1997a; Su *et al.*, 1997b; Legault *et al.*, 1998; Mogridge *et al.*, 1998a). *boxA* appears to be linear at its 5' end, as determined by structure probing with single-stranded RNA-specific and double-stranded RNA-specific ribonucleases. Since nucleotides at both the 3' end of *boxA* and within the *interbox* spacer region are sensitive to both single-stranded and double-stranded RNA-specific reagents, it was suggested that they may alternate between single-stranded and double-stranded conformations (Nodwell and Greenblatt, 1991).

The N proteins of phages λ and 21 cannot functionally complement each other or the N protein of phage P22 (Dambly and Couturier, 1971). The *boxA* sequence is conserved among the lambdoid phages, as well as among the seven *E.coli rrn* operons, suggesting that *boxA* contains sites of interaction with host factors (see below). In contrast, the predicted structure but not the sequence of *boxB* is conserved among the lambdoid phages suggesting that *boxB* is the recognition site for N. The *interbox* is not well conserved and its function has not been described. However, gel mobility shift experiments with mutated *nut* sites have suggested that these nucleotides are important for the association of host factors with the N-modified antitermination complex (Chapter 4; Mogridge *et al.*, 1998b). Furthermore,

deletion of some of the nucleotides in the *interbox* prevented antitermination *in vivo* (Doelling and Franklin, 1989).

The Processive Antitermination Complex

Many protein-protein and protein-RNA interactions characterize the processive antitermination complex. Most of these interactions were initially identified genetically and then verified by various biochemical assays.

N protein

N proteins from related phages are all small, basic proteins but their sequences are highly divergent (Franklin, 1985b). However, a lysine- and arginine-rich region with high interspecies similarity is essential for N to function *in vivo* (Franklin, 1993). Arginine-rich motifs are present in several RNA-binding proteins, such as the HIV Rev and Tat proteins (Tan and Frankel, 1995). The structures of Rev bound to its cognate site, RRE (Battiste *et al.*, 1996), and Tat bound to its RNA site, Tar (Puglisi *et al.*, 1995), showed that different Arg-rich motifs adopt different structures that bind differentially to RNA, suggesting that the motif allows for flexibility in binding. Direct evidence that N binds *boxB* was provided by domain swapping experiments in which the amino-terminal region of the λ N protein was replaced by the homologous region from the ϕ21 N protein (Lazinski *et al.*, 1989). This hybrid N protein was only functional when the *boxB* sequence from ϕ21 was present (Lazinski *et al.*, 1989). Subsequently, gel mobility shift experiments demonstrated a direct

and specific interaction between N and *boxB* RNA. The electrophoretic mobility of the wild type *nut* site RNA was altered upon the addition of N protein, whereas mutations in nucleotides 1, 3, or 5 of the 5 nucleotide loop of *boxB* prevented this alteration (Chattopadhyay *et al.*, 1995a; Mogridge *et al.*, 1995). The minimal peptide that could bind *boxB* was the amino-terminal 22 amino acids, confirming the earlier suggestion that this arginine-rich region interacts with the RNA (Chattopadhyay *et al.*, 1995a; Tan and Frankel, 1995).

N is relatively unstructured until it comes into contact with the RNA (Su *et al.*, 1997a; Van Gilst *et al.*, 1997; Mogridge *et al.*, 1998a). Furthermore, the structure of N (1-22) bound to *boxB* site RNA, as solved by NMR, indicated that its arginine-rich region adopts an α-helical structure in the bound state (Su *et al.*, 1997a; Legault *et al.*, 1998; Mogridge *et al.*, 1998a). The rest of the λ N protein remains unstructured when N binds *boxB* RNA (Mogridge *et al.*, 1995) presumably until N comes into contact with other binding partners such as NusA and RNA polymerase (see below). The structure of *boxB* is also altered upon N-peptide binding (Su *et al.*, 1997b; Legault *et al.*, 1998; Mogridge *et al.*, 1998a). Instead of the predicted 5-membered loop, *boxB* forms a stable GNRA tetraloop-like structure with the guanosine in the first nucleotide of the loop and the adenine in position five of the predicted loop forming a non-Watson-Crick base pair (Su *et al.*, 1997a; Su *et al.*, 1997b; Legault *et al.*, 1998). The fourth nucleotide of the loop, a guanosine in *nutL*, is looped out, is not part of the structure recognized by N, and is not necessary for the GNRA tetraloop-like structure to form (Legault *et al.*, 1998). Since mutation of this nucleotide is detrimental to antitermination by N *in vivo* (Doelling and Franklin, 1989) and prevents the association of NusA with an N-*nut* site complex *in vitro* (Mogridge *et al.*, 1995; Legault *et*

al., 1998), it is likely that it is important for interaction with NusA. Interestingly, the base stacking interactions in the A-form double helical stem of *boxB* are extended by the G-A base pair that closes out the loop and loop nucleotides 2 and 3. These interactions are additionally stabilized by the indole ring of tryptophan 18 of the protein N (Legault *et al.*, 1998).

Protein affinity chromatography with the full length N protein demonstrated a direct interaction between N and the *E.coli* elongation factor, NusA (Greenblatt and Li, 1981b). Recently, a detailed deletion analysis of the N protein mapped the NusA-interacting region of N to amino acids 34-47 (Mogridge *et al.*, 1998a). A fragment of N (N 1-47) including this region of N as well as the amino-terminal *boxB*-binding region of N, was able to promote read-through of a rho-independent terminator, but not as well as the full length N protein, suggesting that a region important for maximal activity was missing (Mogridge *et al.*, 1998a). In fact, affinity chromatography with other regions of N identified RNA polymerase-binding regions in the carboxy-terminal region of N and in the amino-terminal 47 amino acids of N (Mogridge *et al.*, 1998a). These studies confirmed earlier suggestions that N interacts directly with RNA polymerase (Ghysen and Pironio, 1972; Rees *et al.*, 1996).

Nus Factors

Host mutations that impaired the ability of wild type phage λ, but not N-independent mutant phage, to grow in *E.coli* led to the identification of Nus (*N-ut*ilization *s*ubstance) proteins. These molecules included NusA (Friedman, 1971), NusB (Keppel *et al.*, 1974; Friedman *et al.*, 1976), NusE, or ribosomal protein S10, (Friedman *et al.*, 1981), NusC, or the

β subunit of RNA polymerase (Georgopoulos, 1971b; Georgopoulos, 1971a), and NusD, or termination factor rho (Brachet *et al.*, 1970; Simon *et al.*, 1979). Confirmation of an essential role for Nus proteins in antitermination was provided by subsequent *in vivo* and *in vitro* experiments as discussed below.

The *nusA1* mutation was the first *nus* mutation to be identified. A *nusA1* strain exhibits a strong temperature-sensitive phenotype, preventing λ growth at 42°C but not at 30°C (Friedman, 1971; Friedman and Baron, 1974). Differences between the *Salmonella typhimurium* NusA ($NusA^{St}$) and the *E.coli* NusA ($NusA^{Ec}$) explain the inability of λ to grow or the λ N protein to function in *Salmonella typhimurium* (Baron *et al.*, 1970; Friedman and Baron, 1974; Schauer *et al.*, 1987). Sequence comparisons and the construction of hybrid NusA proteins then identified functional regions of NusA (Craven and Friedman, 1991; Ito *et al.*, 1991; Craven *et al.*, 1994). $NusA^{Ec}$ and $NusA^{St}$ show high amino acid sequence similarity, with small regions of heterogeneity distributed throughout the proteins (Craven *et al.*, 1994). The most striking difference between the proteins is located in the amino-terminal one-third of the proteins and is called the 4-for-9 region. This region has four amino acids in the *E.coli* version and nine amino acids in the *S.typhimurium* version of NusA (Craven *et al.*, 1994). Interestingly, $NusA^{St}$ can substitute for $NusA^{Ec}$ to support bacterial viability, but it cannot support the growth of λ at any temperature (Baron *et al.*, 1970; Friedman and Baron, 1974). Furthermore, a NusA protein, called NusA 449, that was composed entirely of $NusA^{St}$ except for the 4-for-9 region, was able to support the action of λ N (Craven *et al.*, 1994). Thus, the few differences between the two proteins may define regions important for interaction with the antitermination complex.

nusA is an essential gene in *E.coli*. However, it is possible for the bacteria to survive in the absence of a functional NusA protein, but only when there is a mutation in the rho termination factor (Zheng and Friedman, 1994). As NusA functions as a bacterial elongation factor that enhances pausing and the efficiency of termination at certain sites (Kung *et al.*, 1975; Greenblatt *et al.*, 1980), the connection between NusA and rho probably has something to do with the postulated function of NusA in maintaining the tight coupling between transcriptiona and translation (Ruteshouser and Richardson, 1989; Zheng and Friedman, 1994).

A direct interaction between NusA and RNA polymerase has been demonstrated by affinity chromatography (Greenblatt and Li, 1981a). In Chapter 2 of this thesis, I describe deletion studies of NusA and the identification of two distinct RNA polymerase binding regions, one in the amino-terminal 137 amino acids and one in the carboxy-terminal 192 amino acids. Other experiments further identified an interaction between the α subunit of RNA polymerase and the carboxy-terminal RNA polymerase-binding region (Chapter 3). This result confirmed other *in vitro* binding experiments that concluded that NusA interacts directly with α, in addition to β and β' (Liu *et al.*, 1996). It was shown that the interaction between α and NusA is important for the regulation of transcription since RNA-protein cross-linking experiments revealed that α interacts with nascent RNA and this interaction is abolished in the presence of NusA, which itself interacts with the RNA (Liu and Hanna, 1995b). The cross-linking of NusA to the nascent RNA was lost when a variant α that lacked its entire carboxy-terminal domain was used in cross-linking experiments. Functionally, this truncated α interfered with the ability of NusA to enhance termination and Q-mediated antitermination, but not N-mediated antitermination (Liu *et al.*, 1996).

As mentioned above, NusA also interacts directly with N (Greenblatt and Li, 1981b). More specifically, a carboxy-terminal region of NusA interacts with amino acids 34-47 of N (Chapter 3; Mogridge *et al.*, 1998a). These results, in addition to those pertaining to the RNA polymerase-NusA interaction (Chapter 2), form the basis of a model for NusA function that is discussed in Chapter 3 of this thesis.

Continued study of the relationships among $NusA^{St}$ and $NusA^{Ec}$ and the *boxA* sequences from λ and P22 led to genetic evidence suggesting that NusA interacts with *boxA*. The core *boxA* sequences from both phages are almost identical with one exception: *nutL* of P22 has three thymines in positions 6-8 whereas both *nutL* and *nutR* of λ have two thymines and an adenine in the same position. A mutation in λ *boxA* that altered the sequence TTA to TTT allowed $NusA^{St}$ to support the growth of λ (Friedman and Olson, 1983).

Sequence alignments identified S1 and KH homology regions as two putative RNA-binding regions in NusA (Gibson *et al.*, 1993; Bycroft *et al.*, 1997). Other proteins that have S1 and KH domains are able to associate with RNA and have both specific and non-specific RNA-binding properties. For example, the hnRNP K protein, for which the KH domain has been named, is the major oligo-C-binding protein in vertebrates (Swanson and Dreyfuss, 1988; Siomi *et al.*, 1993). Vigilin is a protein made up of 14 KH domains that specifically binds to the 3' UTR of vigillotellin mRNA (Dodson and Shapiro, 1997). While some isolated KH domains can bind some sequences with higher affinity, they also have a lower affinity for other nucleic acids, providing some evidence for their ability to associate with nascent RNA (Dejgaard and Leffers, 1996). These observations introduced the possibility that NusA might bind with high affinity to certain specific sequences, like the *nut* site, yet also interact with lower affinity with other RNA sequences.

The S1 domain is named after ribosomal protein S1, which has 6 of these RNA-binding motifs (Subramanian, 1983). S1, as part of the 30S subunit of the ribosome, binds pyrimidine-rich RNA sequences upstream of the ribosome binding site to facilitate translational initiation (Boni *et al.*, 1991). Thus, S1 probably interacts with many sequences with low affinity. However, S1 can select specific RNA ligands from a random pool of RNAs (Ringquist *et al.*, 1995). In fact, S1 binds with higher affinity to λ *boxA* RNA since mutations in *boxA* weaken the association of S1 with *boxA*-containing RNA (Mogridge and Greenblatt, 1998).

Several lines of evidence suggest that the putative RNA-binding regions in NusA are functional in RNA-binding. λ mRNA was retained on a nitrocellulose filter by NusA, although this interaction did not require the presence of *boxA* (Tsugawa *et al.*, 1985). Protein-RNA cross-linking experiments within an elongation complex demonstrated that NusA interacts with nascent RNAs greater than 10 nucleotides in length (Liu and Hanna, 1995a; Liu and Hanna, 1995b). Interestingly, if NusA was not present, the RNA cross-linked to the α-subunit of RNA polymerase, and the addition of NusA abolished this interaction. Again, *boxA* was not required for cross-linking.

Gel retardation assays using purified proteins and a *nut* site-containing RNA probe provided *in vitro* evidence for a specific NusA-*nut* site interaction. Wild type NusA supershifted an N-*nut* site complex (Mogridge *et al.*, 1995), but this was blocked by mutations in certain nucleotides in *boxA* and *boxB* (Chapter 4; Mogridge *et al.*, 1995). Furthermore, mutation of either a conserved arginine (R199), in the S1 homology region of NusA or the *nusA1* mutation located in the S1 homology region, resulted in NusA proteins that could not supershift the N-*nut* site complex (Chapter 4; Mogridge *et al.*, 1995).

Interestingly, NusA 944, a NusA protein composed entirely of $NusA^{Ec}$ except for the 4-for-9 region, does support the formation of this complex but fails to support the formation of a complex additionally containing RNA polymerase, NusB, NusG and S10 (Chapter 4). These observations suggest that NusA interacts directly with RNA, a hypothesis that is further explored in Chapter 4 of this thesis.

NusB has apparent roles in both transcription and translation. The *nusB5* mutation partially relieves polarity suggesting that it may be involved in transcription termination (Ward and Gottesman, 1981; Sharrock *et al.*, 1985a). The same mutation affects the synthesis of rRNA (Sharrock *et al.*, 1985a). NusB has also been implicated in translational elongation because an IS10 insertion in the *nusB* gene reduces the peptide chain elongation rate by 30% (Taura *et al.*, 1992). Many *nusB* mutations exhibit the traditional Nus phenotype (Keppel *et al.*, 1974; Friedman *et al.*, 1976) and other *nusB* mutations suppress the defects in antitermination by the λ N protein of the *nusA1* and *nusE71* mutations (see below; Friedman and Baron, 1974; Friedman *et al.*, 1981). *In vitro* confirmation of the direct involvement of NusB in antitermination by N was provided when extract made from *nusB5* mutant cells that was defective for antitermination was supplemented with wild type NusB to restore the ability of the extract to support N activity (Georgopoulos *et al.*, 1980; Ghosh and Das, 1984; Swindle *et al.*, 1988).

NusB interacts directly with ribosomal protein S10 (i.e. NusE) (Mason *et al.*, 1992a) and other observations had outlined the importance of *boxA* for NusB function in antitermination (Horwitz *et al.*, 1987; Court *et al.*, 1995). The discovery that NusB and S10 together bind the *boxA* RNA sequence from the *E.coli rrn* operons, but not the closely related one from λ, helped support the hypothesis that NusB and S10 bind λ *boxA* unstably and

require the presence of the other factors in the antitermination complex to maintain a stable interaction (Nodwell and Greenblatt, 1993). In fact, NusB and S10 associate stably with λ *boxA* RNA only when N, NusA, NusG and RNA polymerase are present (Mogridge *et al.*, 1998b). Inhibition of antitermination caused by the overproduction of *boxA* RNA was restored by the overproduction of NusB suggesting that NusB interacts with *boxA* RNA (Friedman *et al.*, 1990). An NMR structure of NusB (Huenges *et al.*, 1998) is consistent with the possibility that an arginine-rich region of NusB may interact with *boxA* RNA.

When ribosomal protein S10 was identified as the NusE factor (Friedman *et al.*, 1981), it was not clear if S10 was acting alone or whether the ribosome was also involved in antitermination. In fact, addition of purified S10 or the 30S ribosomal subunit complements an antitermination-defective extract made from a *nusE71* mutant strain (Das *et al.*, 1985; Horwitz *et al.*, 1987). While it is still formally possible that the ribosome has some role in the process, two observations suggest that ribosomes are not important for antitermination by N. First, *in vitro* antitermination by N can be reconstituted with S10 in the absence of 30S subunits (Li *et al.*, 1992). Second, translation is not necessary for antitermination in crude systems *in vitro* (Goda and Greenblatt, 1985) and does not overcome the antitermination defect of *nusE71* (Warren and Das, 1984). The finding that S10 also interacts with RNA polymerase further supports the idea that the stability of the processive antitermination complex depends on multiple protein-protein interactions (Mason and Greenblatt, 1991).

NusG was identified initially as a Nus factor based on its ability to allow the reconstitution of antitermination by N with purified components (Li *et al.*, 1992). Antitermination assays using an *E.coli* S100 extract as the source of host factors for the reaction were more efficient than reactions containing only purified N, RNA polymerase,

NusA, NusB and S10, suggesting that an additional factor(s) was required (Li *et al.*, 1992). Accordingly, an activity purified to homogeneity from S100 extract that allowed efficient antitermination by N was identified as NusG (Li *et al.*, 1992). Consistent with this, genetic experiments identified a mutation in the *nusG* gene that suppressed the effects of the *nusA1* and *nusE71* mutations on antitermination by N (Sullivan *et al.*, 1992). This suggested that NusG is truly an integral part of N-mediated antitermination. However, depletion of NusG from cells had no effect on the ability of N to promote antitermination, whereas termination at a rho-dependent terminator was compromised in this situation (Sullivan and Gottesman, 1992). Therefore, some other unknown factor may be able to substitute for NusG to support antitermination by N *in vivo*. NusG binds RNA polymerase (Li *et al.*, 1992) and rho factor (Li *et al.*, 1993), adding more interactions to the intricate web of protein-protein interactions within the N-modified transcription complex. The interaction of NusG with rho is important for rho-dependent termination, but may also be a key aspect of the ability of N-modified RNA polymerase to transcribe through rho-dependent terminators (see below).

Mutations that cause RNA polymerase to be resistant to the drug rifampicin map to the *rpoB* gene that encodes the β subunit of RNA polymerase. Importantly, certain rifampicin-resistant (rif^r) mutations that alter the ability of RNA polymerase to terminate at rho-dependent and rho-independent terminators also affected N-mediated antitermination (Jin *et al.*, 1988). These rif^r mutants either enhanced or suppressed the phenotype of a *nusA* mutant, suggesting that the region of the β subunit defined by the rif^r mutations is involved in antitermination and somehow affects the functioning of NusA (Jin *et al.*, 1988)

The Das and Greenblatt laboratories took advantage of the genetic data to define the factors required to form a stable antitermination complex capable of terminator read-through

in vitro. Transcription reactions using S30 and S100 extracts as a source of host factors showed that antitermination by N did not occur with extracts prepared from *nusA*, *nusB* and *nusE* mutant strains (Das and Wolska, 1984; Goda and Greenblatt, 1985; Horwitz *et al.*, 1987). By monitoring the position of the elongation complex, Das and coworkers were able to ascertain that N became a stable component of the complex only after RNA polymerase had transcribed through the *nut* site (Barik *et al.*, 1987). These results were confirmed and expanded by experiments from the Greenblatt laboratory in which elongation complexes were purified by gel filtration from reactions containing wild type or mutant Nus factors and programmed with DNA containing no *nut* site, a wild type *nut* site, or a mutant *nut* site (Horwitz *et al.*, 1987). Although the incorporation of NusA into an elongation complex was independent of the *nut* site, presumably because it interacts directly with RNA polymerase, the incorporation of N depended on the *nut* site and the incorporation of NusB required the presence of wild type NusA, S10 and *boxA* (Horwitz *et al.*, 1987). Therefore, the assembly of a complete complex is highly cooperative.

Gel mobility shift assays are another way of testing the importance of each antitermination factor for the stability of the overall complex. A complex containing RNA polymerase, N, the Nus factors and the *nut* site RNA remains intact in a non-denaturing polyacrylamide gel (Mogridge *et al.*, 1995). However, mutations in RNA polymerase, NusA, *boxA*, or *boxB* that interfere with antitermination prevent the formation of a stable, low mobility complex containing all of the factors (Mogridge *et al.*, 1995). These experiments also defined an order of assembly of the complex where N bound first to the *nut* site RNA, followed by NusA and RNA polymerase. This quaternary complex served, in turn, as a scaffold for the further assembly of the remaining Nus factors, NusB, NusG, and S10

(Mogridge *et al.*, 1995). Interestingly, the stability of this complex does not depend on a DNA template.

A model of the processive N-modified elongation complex is shown in Fig. 4. The model is based on and summarizes extensive genetic and biochemical observations regarding the protein-protein and protein-RNA interactions within the processive antitermination complex. The dissociation constants of these interactions have been measured or estimated (summarized in Greenblatt, 1992). Although most of the interactions are relatively weak, the complete complex is stable enough to remain associated with RNA polymerase for at least 7 kilobases. In the case of the rightward operon of phage λ, the complex remains intact to the

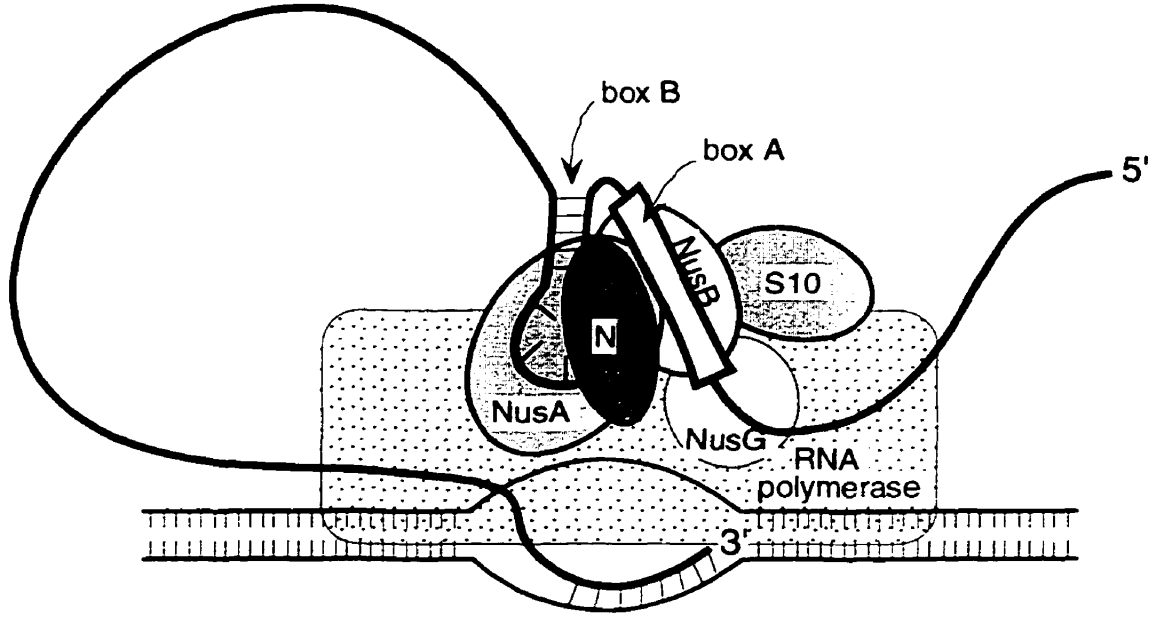


Figure 4. The Complete N-modified Antitermination Complex
The multiple protein-protein and protein-RNA interactions are decribed in the text.

end of the *Q* gene.

Mechanism of Antitermination

Despite the requirement for the Nus factors and the *nut* site for antitermination *in vivo*, N can cause RNA polymerase to ignore termination signals when present at high concentration in an *in vitro* transcription assay (Rees *et al.*, 1996). The addition of NusA and the *nut* site strengthens the ability of N to antiterminate, but only through a terminator located close to the *nut* site (Whalen *et al.*, 1988; DeVito and Das, 1994; Rees *et al.*, 1996). For processive antitermination, the presence of all of the Nus factors is required (Mason *et al.*, 1992b; DeVito and Das, 1994).

The above studies suggested that the Nus factors function primarily to provide stability to the complex, while N provides the actual antitermination activity. N appears to have at least two functions in antitermination: first, N interacts directly with RNA polymerase and probably alters its ability to recognize termination signals; second, N interacts directly with NusA and suppresses the ability of NusA to enhance termination at an intrinsic terminator (Mogridge *et al.*, 1998a). As well, since NusG interacts with rho factor, and this interaction is important for rho-dependent termination, it is likely that the rho-NusG interaction has been exploited by N to prevent termination at rho-dependent terminators. In fact, a mutant rho factor, Rho026, prevents antitermination by N at high temperatures and also causes the rho-NusG interaction to become temperature sensitive (Li *et al.*, 1993). These observations suggest that NusG may bind rho in the antitermination complex to inhibit rho's termination activity (Li *et al.*, 1993).

The exact mechanisms by which N interferes with activities of RNA polymerase and NusA in termination have not been defined. Some possible mechanisms are based on the reversal of key steps involved in termination and include preventing the formation of pause or terminator hairpins, blocking a critical site on RNA polymerase that interacts with pause and terminator hairpins, stabilizing protein-RNA interactions within RNA polymerase, or preventing backsliding by RNA polymerase.

Other Antitermination Systems

Several additional prokaryotic antitermination systems have been identified, including Q-mediated antitermination, *rrn* antitermination, Bgl antitermination and *put*-meditated antitermination. Study of these systems may help identify common features of the antitermination process. Likewise, the differences may also provide interesting clues to explain variations among the systems.

Q-mediated Antitermination

The Q protein is a product of the rightward early operon that requires the function of N for its expression. Q, in turn, is required for the expression of the late genes under the control of the late promoter, P_R'. Early investigations suggested that Q, like N, relieves polarity and functions as an antiterminator. The first *in vitro* evidence for the antitermination activity of Q was provided by transcription assays in which the only proteins present were Q, NusA and RNA polymerase. Read-through of terminators located

downstream of λ P_R' was dependent on the presence of Q (Grayhack and Roberts, 1982). Furthermore, NusA stimulated this activity, suggesting that NusA is a true component of this second antitermination system in λ (Grayhack and Roberts, 1982).

The characteristics of Q-mediated antitermination differ from those of N-mediated antitermination in many ways. First, NusA seems to be the only host factor, other than RNA polymerase, that is involved in the Q system (Grayhack and Roberts, 1982). Moreover, the recognition site for Q protein, called *qut* (*Q-ut*ilization site) is found partly in the promoter DNA, rather than entirely in the nascent RNA (Yang *et al.*, 1987; Yarnell and Roberts, 1992). Q binds the DNA in the promoter and only recognizes RNA polymerase when it has paused at a specific site downstream of the P_R' promoter that is not found downstream of most other promoters (Grayhack *et al.*, 1985; Yarnell and Roberts, 1992). Furthermore, the process by which Q interacts with RNA polymerase is unique to the system. RNA polymerase must be paused naturally at position +16/+17 downstream from P_R' in order for Q to engage the transcription complex and release it from the pause site (Yarnell and Roberts, 1992). This specific pause is dependent on sequences in the initial transcribed region of the promoter, particularly positions +2 and +6 in the non-template strand (Ring and Roberts, 1994). Interestingly, these two positions are within a region of high sequence homology to the σ^{70} consensus -10-binding site (Ring *et al.*, 1996). Although the elongation complex at +16/+17 is far enough downstream of the promoter to have released σ^{70}, σ^{70} is still present and necessary for the pause even though the σ^{70} molecule has lost contact with the -10 and -35 promoter sequences and the RNA polymerase footprint is only 30-35 base pairs long (Ring *et al.*, 1996). Thus, σ^{70} releases its contacts with the promoter only to rebind at a similar sequence just downstream of +1 to initiate the pause necessary for Q-function at +16/+17.

There are some similarities between Q-mediated and N-mediated antitermination. As

is the case for N, RNA polymerase modified by Q becomes resistant to both rho-independent and rho-dependent terminators, and this modification can persist through multiple terminators. As well, both Q and N suppress pausing by RNA polymerase (Yang and Roberts, 1989; Mason *et al.*, 1992b) and this is likely to be related to the mechanisms of antitermination in both cases.

rrn Antitermination

The seven *rrn* operons in *E.coli*, which encode 16S, 23S, and 5S rRNAs and certain tRNAs, are subject to numerous forms of regulation. Some of these regulatory mechanisms are discussed elsewhere in this thesis introduction. Here, I will focus on the antitermination system and discuss its similarities to the λ N-system.

Transcription and translation are generally coupled in prokaryotes. As ribosomes translate the nascent RNA close behind RNA polymerase, rho-dependent terminators remain masked and premature termination is avoided. However, these terminators become active if there is a disruption in the coupling and rho is able to gain access to the nascent RNA transcript. Nevertheless, termination does not occur in the untranslated *rrn* operons. The existence of an antitermination system was strongly suggested by the finding that, despite the insertion of a transposable sequence known to cause polarity in other operons, very little polarity was observed in *rrnC* (Morgan, 1980). Subsequent experiments showed that transcription initiated from an *rrn* promoter, but not the *lac* or *ara* promoters, was able to continue through a rho-dependent terminator (Holben and Morgan, 1984). This system required sequences related to the λ *nut* site, located both in the promoter/leader regions and

the spacer regions of the *rrn* operons between the 16S and 23S genes. Although a hairpin sequence superficially resembling *boxB* is present in these critical regions of the *rrn* operons, a *boxA*-like sequence is necessary and sufficient for antitermination *in vivo* and *in vitro*. However, this *boxA*-directed antitermination functions well only with rho-dependent terminators, and not with rho-indepedent terminators (Berg *et al.*, 1989; Albrechtsen *et al.*, 1990).

Because of the similarity in the sequence requirements for antitermination in the *rrn* and λ N systems, it seemed likely that some of the protein components of N- and *rrn*-mediated antitermination systems would be shared. In support of this idea, overexpression of the λ *nut* site RNA inhibits rRNA production *in vivo* (Sharrock *et al.*, 1985b). In addition, specific mutations in NusA and NusB, but not S10, had a deleterious effect on the expression of the *rrn* operons (Sharrock *et al.*, 1985a). Furthermore, NusB and S10 together bound *rrn boxA* but not lambda *boxA*, in gel mobility shift experiments (Nodwell and Greenblatt, 1993). NusA has been implicated in the system by a mutation in the *nusA* gene that affects antitermination *in vivo* (Vogel and Jensen, 1997). Reconstitution of *rrn* antitermination *in vitro* in reactions containing *E.coli* extracts showed that NusB is absolutely required for antitermination (Squires *et al.*, 1993). Although the addition of NusA, NusG and S10 improved terminator read-through in this system, it has not yet been determined whether they are absolutely required for antitermination (Squires *et al.*, 1993). Elongation complexes synthesizing rRNA *in vitro* have been shown to contain NusB and NusG (Li *et al.*, 1992).

put and nun in phage HK022

The coliphage HK022 has developed a system to prevent superinfection of a lysogen by λ. The HK022 encoded nun protein binds to the *boxB* portion of the λ *nut* site with an affinity similar to that of N protein (Chattopadhyay *et al.*, 1995b). Instead of causing antitermination, nun causes termination just downstream of the *nut* sites (Robert *et al.*, 1987; Robledo *et al.*, 1990). The nun-termination system also shares some *E.coli* factors with the N-antitermination system. Mutations in NusA, NusB and S10 that block antitermination by N also block termination by nun (Robert *et al.*, 1987; Robledo *et al.*, 1991). In fact, these mutations as well as other mutations in *boxA*, cause nun to act as an antitermination factor at the terminator, tR_1, but this alteration in function does not persist for terminators located further downstream (Robledo *et al.*, 1990; Baron and Weisberg, 1992). Thus, the nun and N systems share components, yet differ in fundamental ways.

The nun effect has been studied *in vitro*. Since one possible mechanism of antitermination is prevention of backsliding by RNA polymerase, it is intriguing to note that nun can block elongation *in vitro* by the transcription complex without causing it to dissociate (Hung and Gottesman, 1997). In this arrested complex, RNA polymerase retains catalytic activity. More specifically, the 3' OH terminal nucleotide of the nascent transcript can undergo pyrophosphorolysis and restoration by the addition of nucleoside triphosphates (Hung and Gottesman, 1997). However, only this terminal nucleotide can be modified; further translocation of the complex, either forward or backwards, is prevented by nun. HK022 also has its own antitermination system. However, this system is unique in that no phage-encoded antitermination factor is required (Oberto *et al.*, 1993). RNA polymerase

initiating from the rightward and leftward promoters becomes resistant to terminators due to the action of cis-acting RNA sites, called *put* sites (*p*olymerase *ut*ilization site), that increase the rate of RNA polymerase elongation, probably through an anti-pause mechanism (Oberto *et al.*, 1993; King *et al.*, 1996). This modification may not involve any of the Nus factors that are important for λ N-mediated antitermination and nun termination, since mutations in the Nus factors that prevent antitermination by N do not affect the ability of the *put* sites to direct terminator read-through (Oberto *et al.*, 1993). In fact, a search for host factors involved in antitermination yielded 14 *E.coli* mutants that prevented HK022 growth, all located in one of three conserved cysteine residues thought to be a part of a zinc-binding motif in the β' subunit of RNA polymerase, encoded by *rpoC* (Clerget *et al.*, 1995). One of these mutations was further studied *in vivo* and *in vitro* to show that the mutation specifically decreased terminator read-through (Clerget *et al.*, 1995). The observation that all of the mutants isolated in the screen were able to support λ growth suggests that the mechanisms of λ and HK022 antitermination are highly divergent (Clerget *et al.*, 1995).

The key to the mechanism of *put*-mediated antitermination is clearly the interaction between the *put* site RNA and RNA polymerase, since a purified system containing the template DNA, RNA polymerase and buffer allows terminator read-through (King *et al.*, 1996). Thus, a detailed analysis of the RNA structure of the *put* site may yield important information on the mechanism of antitermination in general. *putL* and *putR* are about 70 nucleotides long and have similar sequences, but are located in different positions within their respective operons (Oberto *et al.*, 1993; King *et al.*, 1996). The predicted *put* site secondary structure was tested by exposure to single- and double-strand-specific ribonucleases and by testing the effects of mutations in predicted stem regions (King *et al.*,

1996; Banik-Maiti *et al.*, 1997). The predicted functional structure consists of two hairpin stems of varying length, separated by a single nucleotide. There are internal bulges associated with the stems that seem to be important for function, but the loops at the ends of the stems are not. Exactly how an interaction between *put* RNA and RNA polymerase prevents termination by the enzyme is not clear at this time.

Thesis Rationale

In this thesis, I have analyzed the structure/function relationship of the *E.coli* elongation factor, NusA. In chapter 2, I identify regions of NusA that interact with RNA polymerase, the λ N protein and RNA. In chapter 3, I show that the inability of full length NusA to interact with RNA in the absence of N is due to an inhibitory domain located in the carboxy-terminal region of NusA. I also show that this inhibition of RNA-binding is relieved by interaction of the carboxy-terminal inhibitory region of NusA with either N or the α subunit of RNA polymerase. In chapter 4, I provide evidence that NusA interacts directly with the *boxA* and *interbox* portions of the *nut* site. Finally, in chapter 5, I summarize my results and suggest additional experiments to extend our knowledge of NusA function.

Chapter 2

Functional Importance of Regions in *E.coli* Elongation Factor NusA that Interact with RNA polymerase, the Bacteriophage λ N protein, and RNA

A version of this chapter has been accepted for publication in Molecular Microbiology.

I performed all of the experiments presented in this chapter except those shown in Figure 2, which was done by R. Muhandiram. J. Li constructed and purified the His_6-tagged NusA proteins.

Abstract

The association of the essential *E.coli* protein NusA with RNA polymerase increases pausing and the efficiency of termination at intrinsic terminators. NusA is also part of the N protein-modified antitermination complex phage λ that functions to prevent transcriptional termination. I have investigated the structure of NusA by using various deletion fragments of NusA in a variety of *in vitro* assays. Sequence and structural alignments have suggested that NusA has both S1 and KH homology regions that are thought to bind RNA. I show here that the portion of NusA containing the S1 and KH homology regions is important for NusA to enhance both termination and antitermination. There are two RNA polymerase-binding regions in NusA, one in the amino-terminal 137 amino acids and the other in the carboxy-terminal 264 amino acids; only the amino-terminal RNA polymerase-binding region provides a functional contact that enhances termination at an intrinsic terminator or antitermination by N. The carboxy-terminal region of NusA is also required for interaction with N and is important for the formation of an N-NusA-*nut* site or N-NusA-RNA polymerase-*nut* site complex; the instability of complexes lacking this carboxy-terminal region of NusA that binds N and RNA polymerase can be compensated by the presence of the additional *E.coli* elongation factors, NusB, NusG and ribosomal protein S10.

Introduction

NusA and the additional host proteins NusB, NusG and ribosomal protein S10 are important for the N protein of bacteriophage λ to modify RNA polymerase into a termination-resistant state (Das and Wolska, 1984; Goda and Greenblatt, 1985; Horwitz *et al.*, 1987; Schauer *et al.*, 1987). This modification requires a cis-acting element, called the *nut* site, that is composed of two elements, *boxA* and *boxB* (Salstrom and Szybalski, 1978; de Crombrugghe *et al.*, 1979; Olson *et al.*, 1982). N, the host proteins, and *nut* site RNA assemble into a highly stable complex that associates with elongating RNA polymerase (Barik *et al.*, 1987; Horwitz *et al.*, 1987; Mason *et al.*, 1992b).

Genetic studies on antitermination by the λ N protein have suggested that NusA may interact with the *boxA* portion of the *nut* site (Olson *et al.*, 1982; Friedman and Olson, 1983; Olson *et al.*, 1984), whereas gel mobility shift experiments have shown that mutations in both *boxA* and *boxB* specifically affect the interaction of NusA with an N-*nut* site complex (Mogridge *et al.*, 1995; Chapter 4). Protein-RNA cross-linking data have shown that NusA also interacts directly with the nascent single-stranded RNA in a transcription complex (Liu and Hanna, 1995b; Liu and Hanna, 1995a). In this context, however, *boxA* is not required for the NusA-RNA interaction.

Consistent with the evidence that NusA may interact with RNA, sequence and structural alignments have indicated that NusA has S1 and KH homology regions, both of which are thought to interact with RNA (Gibson *et al.*, 1993; Bycroft *et al.*, 1997). Relatively little is known about the S1 and KH homology regions in NusA, but some mutations in the S1 homology region of NusA prevent λ growth without affecting the viability of *E.coli* (Y. Zhou and D. Friedman, unpublished data; Friedman, 1971). These mutations often affect the ability of NusA to associate with an N-*nut* site RNA complex in

gel mobility shift experiments, but do not directly affect the interaction between NusA and N (Chapter 4; T. Mah, Y. Zhou, N. Yu, J. Mogridge, E. Olsen, J. Greenblatt and D. Friedman, unpublished data; Mogridge *et al.*, 1995).

Other studies have focused on the interaction of NusA with RNA polymerase. The core RNA polymerase enzyme consists of essential α, β and β' subunits and one non-essential subunit, ω (Burgess *et al.*, 1969; Gentry *et al.*, 1991). Genetic experiments have suggested there may be interactions of NusA with β, β' and α (Jin *et al.*, 1988; Ito *et al.*, 1991; Ito and Nakamura, 1993; Ito and Nakamura, 1996; Schauer *et al.*, 1996). More recently, it was shown that NusA can bind directly to the carboxy-terminal domain (CTD) of α, and possibly to β and β', but not to the amino-terminal domain (NTD) of α (Liu *et al.*, 1996). It has also been suggested that a direct interaction between the α-CTD and NusA is important for NusA's effects on pausing and termination, but not N-mediated antitermination (Liu *et al.*, 1996).

I have taken a systematic approach in an attempt to assign functions to the various regions of NusA. By making deletions at putative domain boundaries and assaying for effects on function, I was able to identify two RNA polymerase-binding regions, one in the amino-terminal region and another in the carboxy-terminal region of NusA. I also found that the binding of NusA to N requires amino acids in the carboxy-terminal region of NusA. Interestingly, loss of the carboxy-terminal RNA polymerase-binding region does not have an effect on termination or antitermination by N. Similarly, loss of the N-binding region does not have an effect on antitermination. However, loss of the S1 and KH homology regions or the amino-terminal RNA polymerase-binding region abolishes NusA's ability to influence either of these processes. Surprisingly, the loss of the amino-terminal RNA polymerase-binding region of NusA does not impair NusA's ability to assemble into a stable complex containing N, NusB, NusG, S10, RNA polymerase and the *nut* site RNA, indicating that

interaction of this region of NusA with RNA polymerase may have a direct role in both termination and antitermination.

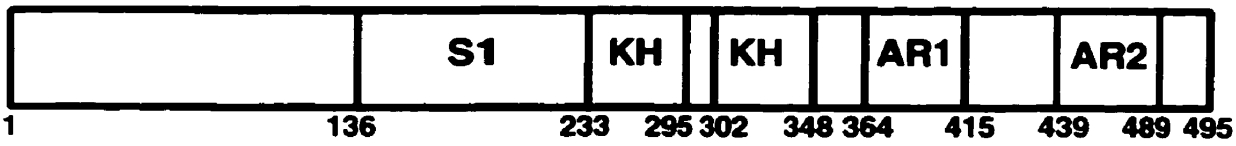
Results

Production and characterization of NusA fragments

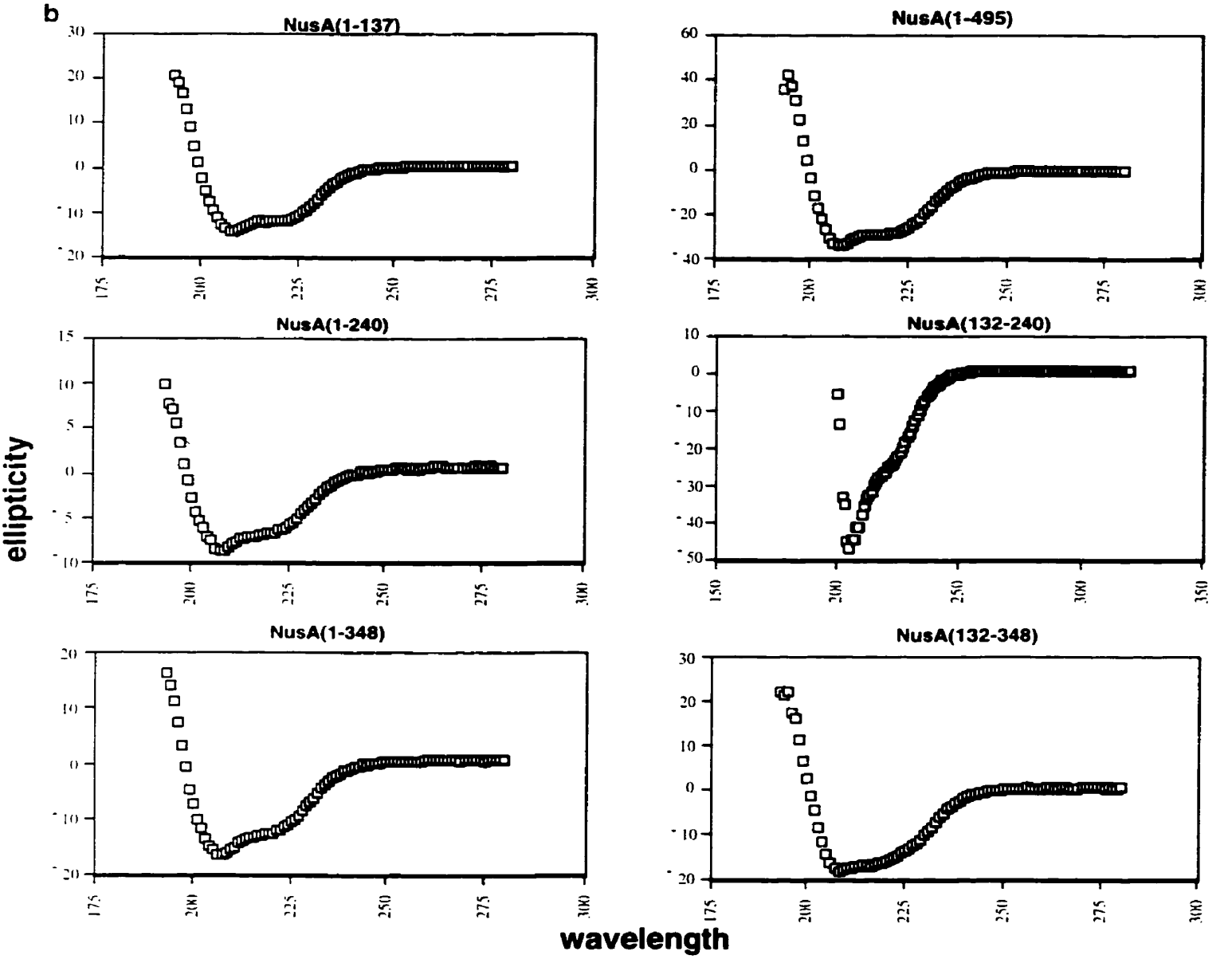
NusA binds to the core component of RNA polymerase in order to modulate transcription (Greenblatt and Li, 1981a). A schematic illustrating some of the putative domains of NusA is shown in Fig. 1a. These regions were assigned on the basis of comparisons with other proteins containing similar domains or on the basis of sequence alignments with NusA proteins from other organisms. Because of this, structural analysis must still be done to confirm the exact domain boundaries of NusA. The S1 and KH homology regions are putative RNA-binding domains, and the regions designated AR1 and AR2 are 50 amino acid repeat sequences which are 39% identical and 54% similar and contain a high percentage of acidic residues (Gibson *et al.*, 1993; Craven *et al.*, 1994; Bycroft *et al.*, 1997). In order to define structure-function relationships within NusA, J. Li used these putative domain boundaries to make various polyhistidine-tagged deletion constructs. The resulting fragments of NusA were then overproduced in *E.coli* and purified to near-homogeneity by chromatography on a nickel chelate resin (see Experimental Procedures). All the NusA fragments were highly soluble in native conditions.

I analyzed all of the purified NusA fragments by circular dichroism (CD) spectroscopy and some of the results are shown in Fig. 1b. With the exception of NusA (132-240), representing the isolated S1 homology region, there were obvious minima at 222nm and 208nm, indicating the presence of α-helices. In Fig. 1c are listed the % α-helix content of each NusA fragment and, for comparison, the % α-helix predicted by the PHDsec secondary structure prediction program (Rost and Sander, 1993; Rost and Sander, 1994).

a



b



c

Construct	% Helix	Predicted % Helix
1-137	39	46
1-240	15	31
1-348	33	28
1-416	23	33
1-495	33	38
132-240	5	12
132-348	28	15
132-416	34	26
132-495	39	34
232-348	26	17
232-416	24	34
232-495	49	42

Figure 1: Circular dichroism analysis of NusA deletion constructs. (a) Predicted domain architecture of NusA. (b) CD analysis of NusA deletion constructs. Wavelength scans were performed on selected NusA deletion constructs from 300 nm to 200 nm. (c) Comparison of predicted % helix with actual % helix. Predicted secondary structure was generated by the PHDsec program. The values in the table for % helix were calculated using the following equation: ([θ222] + 2340)/30,3000 (Chen *et al.*, 1972).

On the basis of these results, I think it is unlikely that the isolated S1 homology region in NusA (132-240) is properly folded. As well, because the 15% α-helical content of NusA (1-240) was much lower than the 31% predicted by the PHDsec program and much lower than the 30% that would be predicted by comparison with the actual α-helical contents of NusA (1-348), NusA (1-137) and NusA (232-348), it is probable that the S1 homology region in NusA (1-240) is also not properly folded. This conclusion was confirmed by NMR analysis, as shown in Fig. 2a. For this experiment, NusA (1-240) was labeled with ^{15}N and a two-dimensional ^{1}H-^{15}N quantum correlation (HSQC) spectrum was generated (Kay *et al.*, 1992; Zhang *et al.*, 1994). The lack of dispersion of the resonances in the spectrum strongly suggested that NusA (1-240) is improperly folded and highly aggregated. The likely lack of proper folding was taken into account when interpreting the negative results described below with NusA (1-240) and NusA (132-240).

In contrast, CD analysis showed that the amino-terminal fragment, NusA (1-137), has high α-helical content, similar to what is predicted. As well, this fragment was able to bind RNA polymerase (see below). In CD experiments, NusA (1-137) showed cooperative and reversible melting with a T_M of approximately 50°C (data not shown), providing strong evidence for an independently folded domain. This conclusion was confirmed by analyzing the HSQC spectrum of ^{15}N-labeled NusA (1-137), as shown in Fig.2b. In this case, the generally good dispersion of the resonances was indicative of a highly ordered, folded domain.

By comparing the 33% α-helical content of the active fragment NusA (1-348) (see below) and the 39% helical content of NusA (1-137), it can be predicted that NusA (132-348) would be about 29% α-helix, which is similar to the actual value of 28% (Fig. 1c). This suggested that the S1 and KH homology regions in NusA (132-348) should be folded for the most part. Again, this conclusion was verified by NMR (Fig. 2c). In this case, the good dispersion of the resonances in the HSQC spectrum suggested that NusA (132-348) also has

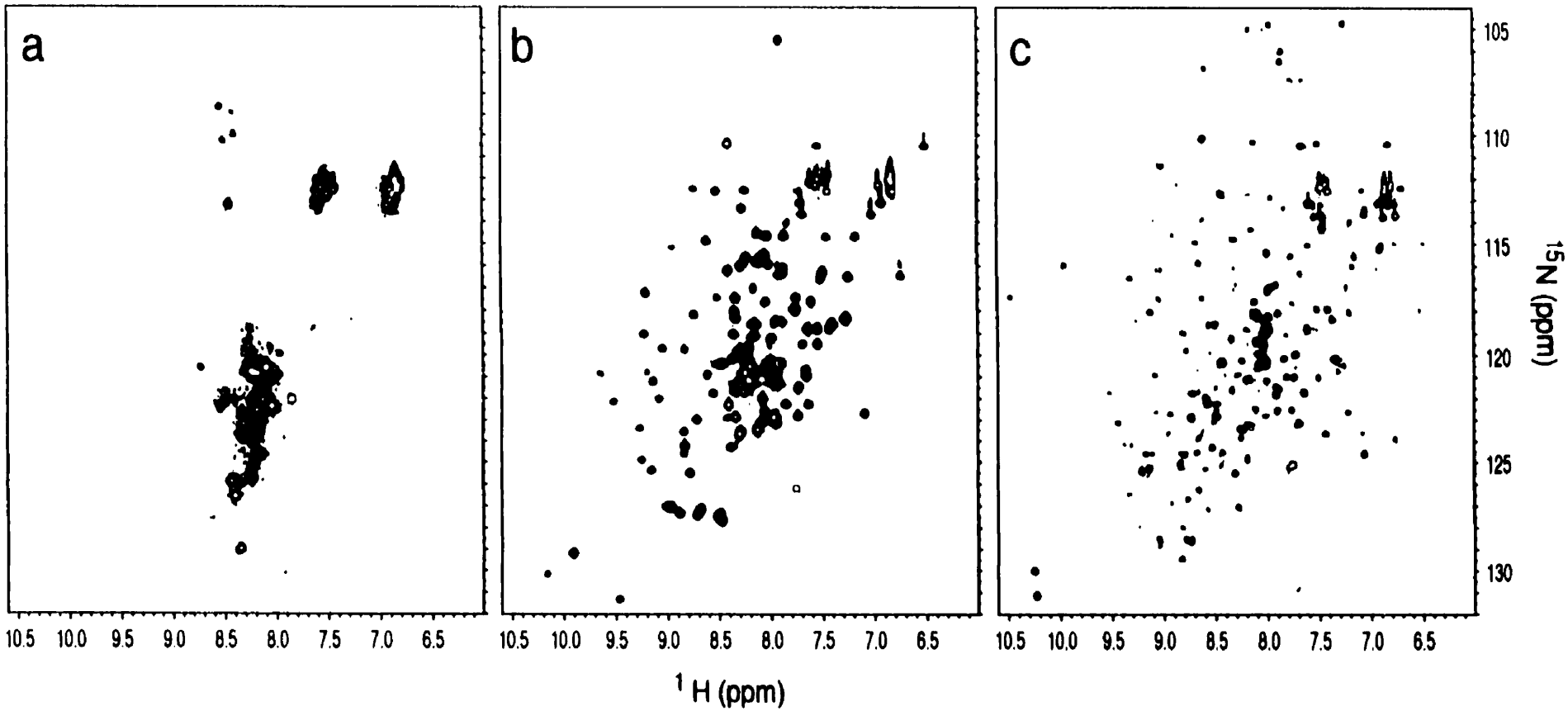


Figure 2: NMR analysis of fragments of NusA. (^{1}H, ^{15}N) HSQC spectra are shown for (a) NusA (1-240), (b) NusA (1-137), and (c) NusA (132-348). The sample concentrations were approximately 0.6 mM, 1 mM, and 0.1 mM, respectively.

a highly ordered stucture, but the variation in signal intensity was suggestive of conformational heterogeneity. The likely presence in this NusA fragment of three domains, one S1 homology region and two KH homology regions (see Fig. 1a), could generate such conformational heterogeneity if one for more of the domains has freedom of motion with respect to the others. The longer active fragments, NusA (1-416) and NusA (1-495), and the multi-domain inactive fragments, NusA (132-416) and NusA (132-495), which can interact with both N and with N-*nut* site complexes (see below), are also likely to be folded. I have made the assumption that these fragments of NusA, as well as NusA (1-137) and NusA (132-348), are properly folded when interpreting the results of some of the experiments that are described below.

Two RNA polymerase-binding regions in NusA

To determine which part of NusA is required for the NusA-RNA polymerase interaction, carboxy-terminal deletion mutants of NusA with His_6 tags at the amino-terminus were used as column ligands in affinity chromatography experiments (Fig. 3a). *E.coli* extract containing additional RNA polymerase core component (Fig. 3a, lane 2) was passed over the columns. The core polymerase subunits, β, β' and α, bound specifically to all of the deletion constructs that were tested (Fig. 3a, lanes 4-8), over and above the background binding to the Ni-agarose column matrix (Fig. 3a, lane 3). None of the NusA fragments had acquired an ability to bind proteins non-specifically, because few, if any, of the other proteins in the *E.coli* extract bound exclusively to any of the NusA columns (Fig. 3a, compare lane 3 with lanes 4-8). The smallest NusA fragment that bound RNA polymerase, NusA (1-137) (Fig. 3a, lane 4), corresponds to a moderately α-helical amino-terminal domain of NusA that folds independently (see Figs. 1b and 2b) and had not yet been assigned a function.

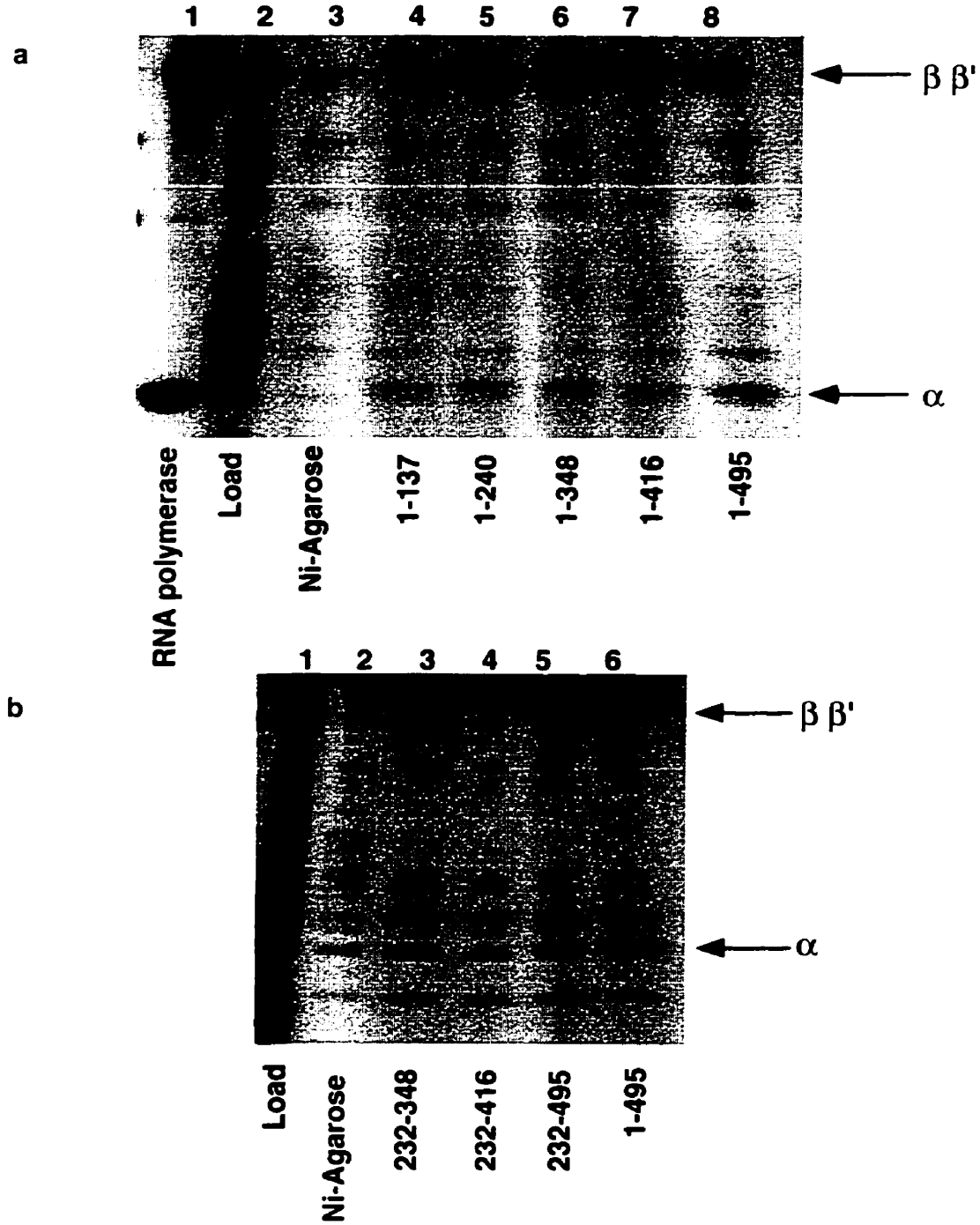


Figure 3: Two RNA polymerase binding-regions in NusA. (a) NusA (1-137) binds RNA polymerase. *E.coli* extract containing additional RNA polymerase core enzyme (lane 2; RNA polymerase is shown in lane 1) was passed over columns containing Ni-agarose (lane 3), 1.3 mg/mL His6-tagged NusA (lane 8) or various His6-tagged NusA deletion mutants (lanes 4-7). The concentration of each NusA deletion mutant on the column was adjusted so that each had the same molar concentration as the full length NusA. Bound proteins were eluted with buffer containing 1 M NaCl, subjected to SDS-PAGE and stained with silver. (b) NusA (232–495) binds RNA polymerase. *E.coli* extract containing additional RNA polymerase core enzyme (lane 1) was passed over columns containing Ni-agarose (lane 2), 1.3 mg/mL His6-tagged NusA (lane 6) or equivalent molar concentrations of various His6-tagged NusA deletion mutants (lanes 3-5). Bound proteins were eluted with buffer containing 1 M NaCl, subjected to SDS-PAGE and stained with silver.

Since direct binding experiments (Liu *et al.*, 1996) and NusA-RNA polymerase crosslinking experiments (J. Li and J. Greenblatt, unpublished data) have shown that NusA may be able to bind to more than one subunit of RNA polymerase, I reasoned that there could be another RNA polymerase-binding site on NusA. Therefore, His_6-tagged constructs that lacked either the first 136 amino acids of NusA (data not shown) or its first 232 amino acids (Fig. 3b) were coupled to Ni-agarose and used as affinity chromatography ligands. In this case, the β, β' and α subunits of RNA polymerase were specifically retained on a column containing full length NusA (Fig. 3b, lane 6), as well as on a column containing NusA (232-495) (Fig. 3b, lane 5), but deletion of the carboxy-terminal 79 amino acids in NusA (232-416) resulted in the loss of RNA polymerase binding (Fig. 3b, lane 4). Similarly, a NusA protein beginning at amino acid 132 would not bind RNA polymerase when the carboxy-terminal 79 amino acids of NusA were deleted (data not shown). Thus, there are two RNA polymerase-binding regions in NusA, one in the amino-terminal 1-137 amino acids and another in the carboxy-terminal 264-amino acids. These conclusions were confirmed by two additional experiments showing that both highly purified RNA polymerase core enzyme and the endogenous RNA polymerase in a crude extract bind specifically to columns containing immobilized NusA (1-137), NusA (232-495), and NusA (1-495) (data not shown). Therefore, NusA (1-137) and NusA (232-495) can bind directly and independently to RNA polymerase.

Only the amino-terminal RNA polymerase-binding region of NusA is necessary for enhancement of termination at an intrinsic terminator

NusA increases the termination efficiency of RNA polymerase at many intrinsic terminators (Greenblatt *et al.*, 1981; Ward and Gottesman, 1981; Farnham *et al.*, 1982; Schmidt and Chamberlin, 1987). It was therefore of interest to determine the impact of the loss of one or other of the RNA polymerase-binding regions on NusA's function in

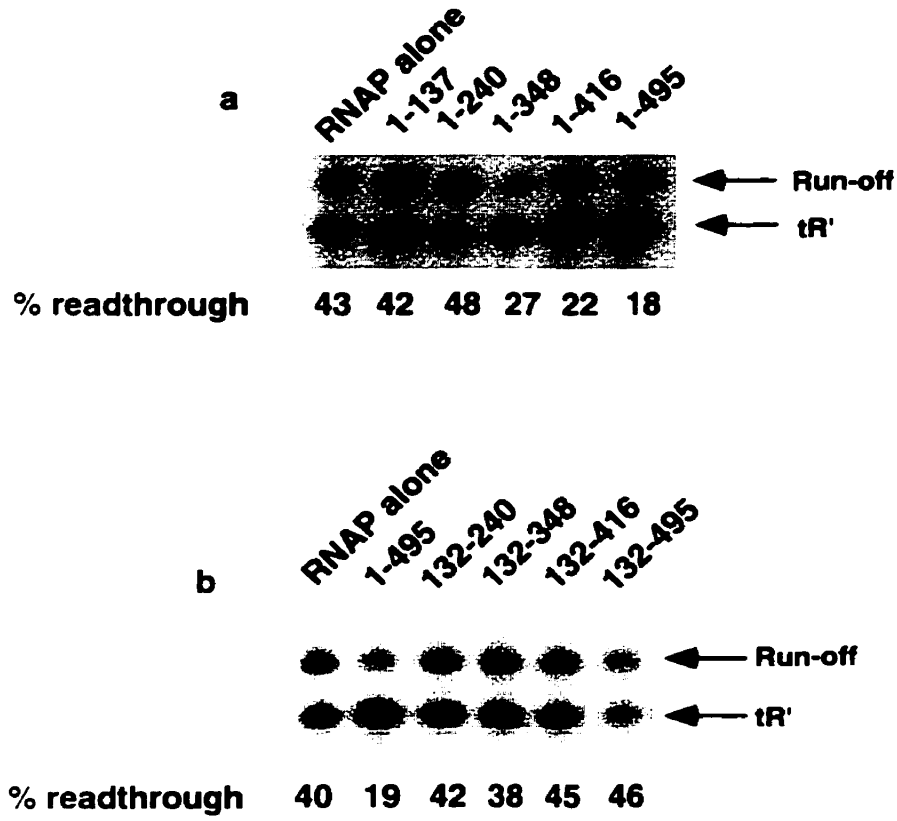


Figure 4: Regions of NusA important for enhancement of termination *in vitro*. (a) The carboxy-terminal RNA polymerase-binding region of NusA is not necessary for termination enhancement. (b) The amino-terminal RNA polymerase binding-region of NusA is necessary for termination enhancement. Transcription reactions containing 25 nM RNA polymerase and 50 nM NusA, or various NusA deletion mutants (as indicated), were electrophoresed on 6M urea 4% polyacrylamide gels, dried and exposed to film. Positions of the terminated and run-off transcripts are indicated.

termination. To this end, *in vitro* transcription was performed on a template containing a promoter, a wild type *nut* site, and the simple λ terminator, tR' (Fig. 4a). When no NusA was added, RNA polymerase alone gave 43% readthrough of the tR' terminator. NusA (1-495) substantially decreased the amount of readthrough to 18%. In contrast to this effect, the addition of NusA (1-137) or NusA (1-240) had no effect on the efficiency of termination. Surprisingly, even though NusA (1-348) and NusA (1-416) had lost the ability to bind RNA polymerase via their carboxy-terminal RNA polymerase-binding regions, both proteins enhanced termination almost as well as the full length NusA protein. Thus, the presence of the carboxy-terminal RNA polymerase-binding region of NusA is not required for enhancement of termination. The ability of NusA (1-137) to bind RNA polymerase and the evidence from CD spectroscopy that this fragment of NusA is folded (see above), combined with the inability of NusA (1-137) and NusA (1-240) to enhance termination, indicated that the region containing the S1 and KH homology regions (amino acids 132-348) is required for enhancement of termination. Since the S1 homology region in NusA (1-240) is unlikely to be folded (see above), it is still unclear which of the S1 and KH homology regions are individually required for NusA to enhance termination.

The fact that NusA constructs containing only the amino-terminal RNA polymerase-binding region and the S1 and KH homology regions were able to enhance termination suggested that the amino-terminal RNA polymerase-binding region might be very important for this process. To test this idea, *in vitro* transcription was carried out using four NusA constructs of various lengths that lacked this region (Fig. 4b), and none was able to enhance termination. Since CD and NMR experiments on NusA (132-348), which cannot enhance termination, had indicated that this portion of NusA is folded on its own, these results showed that the amino-terminal RNA polymerase-binding region is necessary for NusA to enhance termination and cannot be compensated by the carboxy-terminal RNA polymerase-binding region that is present in NusA (132-495).

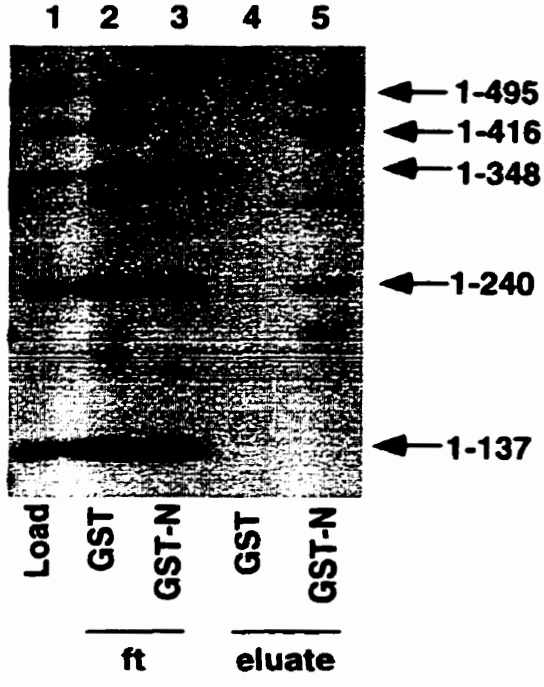


Figure 5: N binds a carboxy-terminal region of NusA. A mixture of His6-tagged NusA and His6-tagged NusA deletion mutants (as indicated. lane 1) was passed over affinity columns containing 2 mg/mL GST (lanes 2 and 4) or 0.5 mg/mL GST-N (lanes 3 and 5). The flow through (lanes 2 and 3) and the 1 M NaCl eluate (lanes 4 and 5) fractions were subjected to SDS-PAGE and stained with silver.

A carboxy-terminal region of NusA is necessary for the binding of N

NusA also binds the bacteriophage λ N protein (Greenblatt and Li, 1981b) and these proteins can together associate with RNA polymerase to prevent termination of transcription when a *nut* site is present (Whalen *et al.*, 1988; DeVito and Das, 1994; Mogridge *et al.*, 1995; Rees *et al.*, 1996). In order to determine which part of NusA is required for a direct interaction with N, GST-N was used as an affinity ligand in binding experiments (Fig. 5). A mixture of His_6-tagged full length NusA (1-495) and four carboxy-terminally deleted proteins, NusA (1-137), NusA (1-240), NusA (1-348) and NusA (1-416), as shown in lane 1, was passed over a GST control column and a GST-N column with sufficient capacity to bind all the NusA fragments. All of the proteins flowed through the GST column (Fig. 5, lane 2) and were absent from the salt eluate of this column (Fig. 5, lane 4). In contrast, the flow-through from the GST-N column contained only the three smallest proteins, NusA (1-137), NusA (1-240) and NusA (1-348) (Fig. 5, lane 3), whereas the salt eluate from this column contained significant amounts of only full length NusA and the deletion construct that had the smallest carboxy-terminal truncation, NusA (1-416) (Fig. 5, lane 5). Since NusA (1-348) is active in termination and antitermination assays (see above and below), it must be properly folded. Therefore, these results indicated that the N-binding ability of NusA requires a carboxy-terminal portion of NusA, but the extreme carboxy-terminal 80 amino acids are not required for this interaction.

The amino-terminal RNA polymerase-binding region of NusA is essential for enhancing antitermination by N

NusA enhances an intrinsic ability of N to antiterminate transcription (Whalen *et al.*, 1988; DeVito and Das, 1994; Rees *et al.*, 1996). Therefore, *in vitro* transcription was performed in order to assess the effects of progressive carboxy-terminal deletions in NusA on

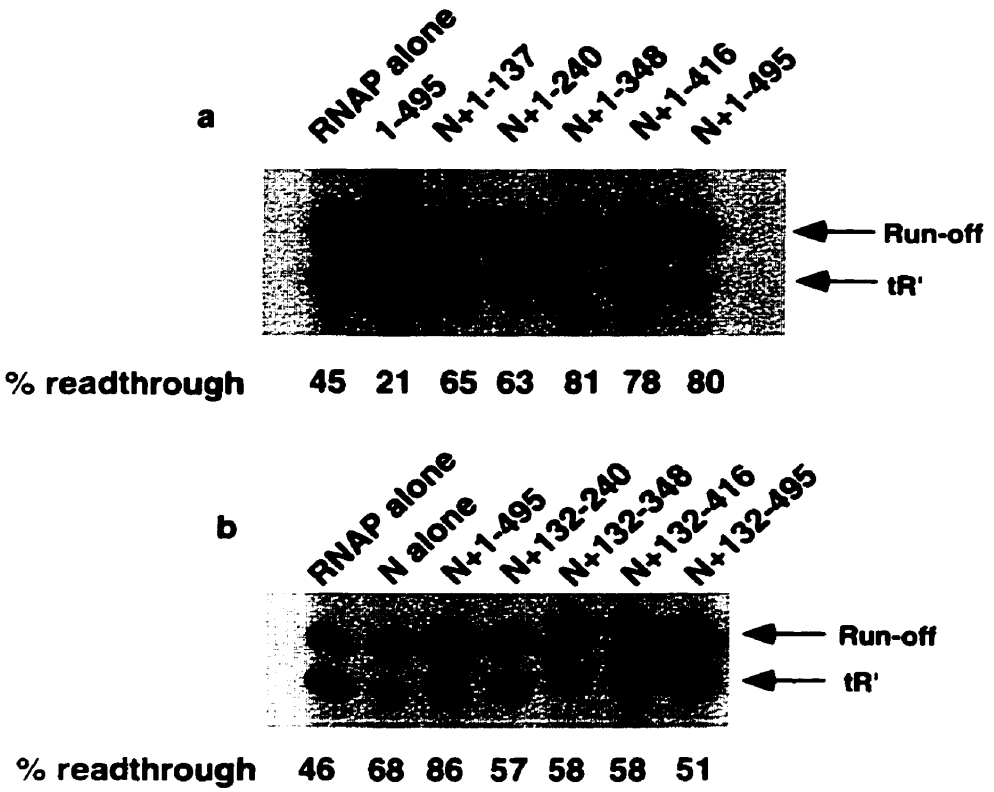


Figure 6: Regions of NusA necessary for enhancement of antitermination by N. (a) The carboxy-terminal N-binding region of NusA is not necessary for enhancement for antitermination by N. (b) The amino-terminal RNA polymerase binding-region of NusA is essential for enhancement of antitermination by N. Transcription reactions containing 25 nM RNA polymerase, 100 nM N and 50 nM NusA or various NusA deletion mutants (as indicated) were electrophoresed on 6 M urea 4% polyacrylamide gels, dried and exposed to film. Positions of the terminated and run-off transcripts are indicated.

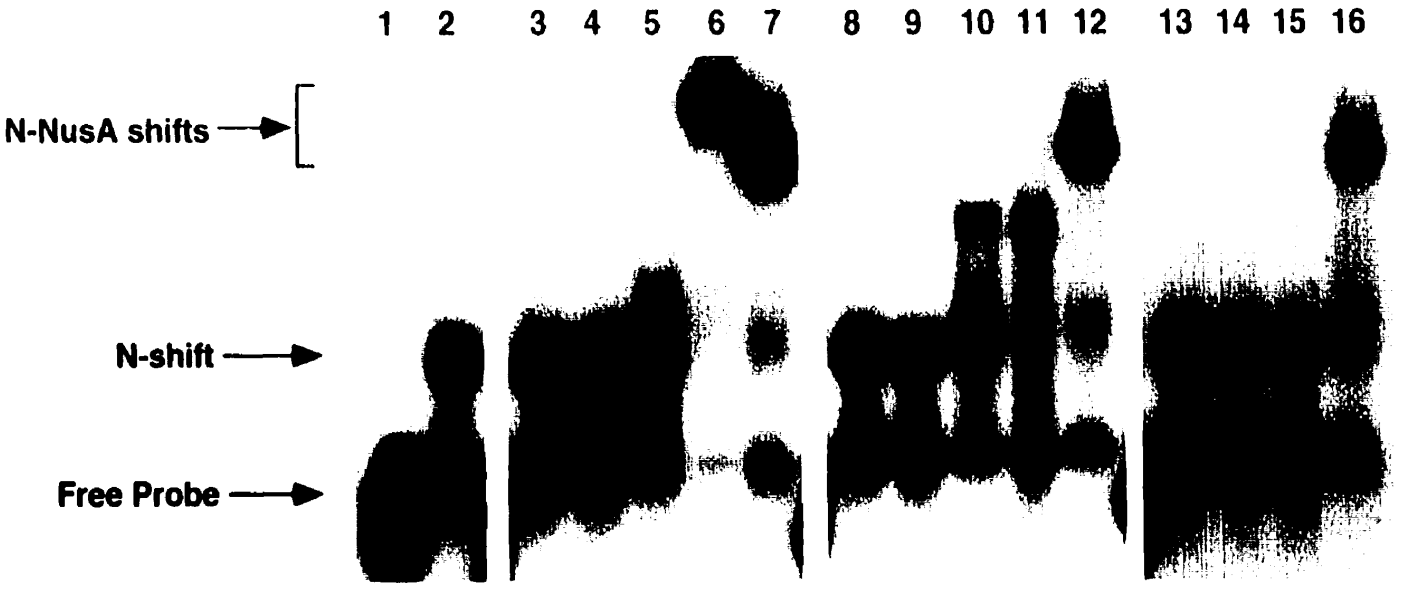
the ability of NusA to influence antitermination (Fig. 6). Addition of N protein alone to a reaction increased the amount of readthrough from 46% to 68% (Fig. 6b). When NusA was added to a reaction containing N and RNA polymerase, the readthrough was further increased to 86%. Addition of NusA (1-137) or NusA (1-240) to the reaction had no effect on the ability of N to antiterminate (Fig. 6a). Surprisingly, however, NusA (1-348), which had lost its ability to bind directly to N, enhanced the effect of N on readthough of the tR' terminator as well as the two N-binding constructs, NusA (1-416) and NusA (1-495) (Fig. 6a). Combined with CD and NMR evidence that NusA (1-137) is folded and other evidence that it binds RNA polymerase, the inability of NusA (1-240) and NusA (1-137) to enhance antitermination by N indicated that the region containing the KH and S1 homology regions (amino acids 132-348) must be critical for this process.

Experiments were also carried out to assess the effect of deleting the amino-terminal RNA polymerase-binding region of NusA on antitermination by N (Fig. 6b). Loss of the amino-terminal RNA polymerase-binding region of NusA in NusA (132-348), NusA (132-416) and NusA (132-495) greatly affected the ability of NusA to enhance antitermination by N. The levels of readthrough observed with these amino-terminally truncated NusA constructs were not increased beyond the level that N could promote on its own. In fact, there was actually less readthrough, suggesting that these mutant NusA proteins may be interfering with antitermination by N via an unknown mechanism. Thus, even the amino-terminally deleted N-binding constructs, NusA (132-416), which can interact with an N-*nut* site complex (see below), and NusA (132-495), which can even assemble into complexes containing RNA polymerase (see below), were unable to influence antitermination, suggesting that the amino-terminal RNA polymerase-binding region of NusA is critical for this process.

Effects of NusA deletions on the formation of N-NusA-*nut* site complexes

Gel mobility shift experiments can be used to assess the interaction of N, the *E.coli* Nus factors, and RNA polymerase with ^{32}P-labeled *nut* site-containing RNA (Mogridge *et al.*, 1995). In order to assess the importance of the various NusA domains on the binding of NusA to N-*nut* site complexes, gel shift experiments were done with N and various NusA deletion constructs (Fig. 7). N protein alone is sufficient to bind and retard the mobility of RNA containing a wild type *nut* site (Fig. 7, lanes 1 and 2), whereas full length NusA cannot shift the RNA on its own (Mogridge *et al.*, 1995). When full length NusA was added to the reaction containing N, the RNA was supershifted from the N-*nut* site complex (Fig. 7, lanes 7, 12, and 16). NusA (1-416), which can also bind N directly (see Fig. 5), was also capable of supershifting the N-RNA complex (Fig 7, lane 6). In this case, the supershifted complex had lower mobility, perhaps because deleting the last 80 amino acids of NusA alters its conformation or causes it to dimerize. As expected, the deleted, non-functional NusA proteins that were not able to bind N, namely NusA (1-137) and NusA (1-240), were not able to supershift the N-RNA complex (Fig. 7, lanes 3 and 4), and even NusA (1-348), which supports antitermination but cannot bind N, interacted more weakly with the complex (Fig. 7, lane 5). None of the NusA fragments except NusA (1-416) can directly bind the RNA in the absence of N (Chapter 3). Thus, only NusA constructs that can bind N can interact with an N-*nut* site complex strongly enough to create a supershift in a non-denaturing gel. Because the carboxy-terminal N-binding region of NusA is important for the supershift, I could not determine whether the KH and S1 homology regions of NusA are also necessary for complex formation.

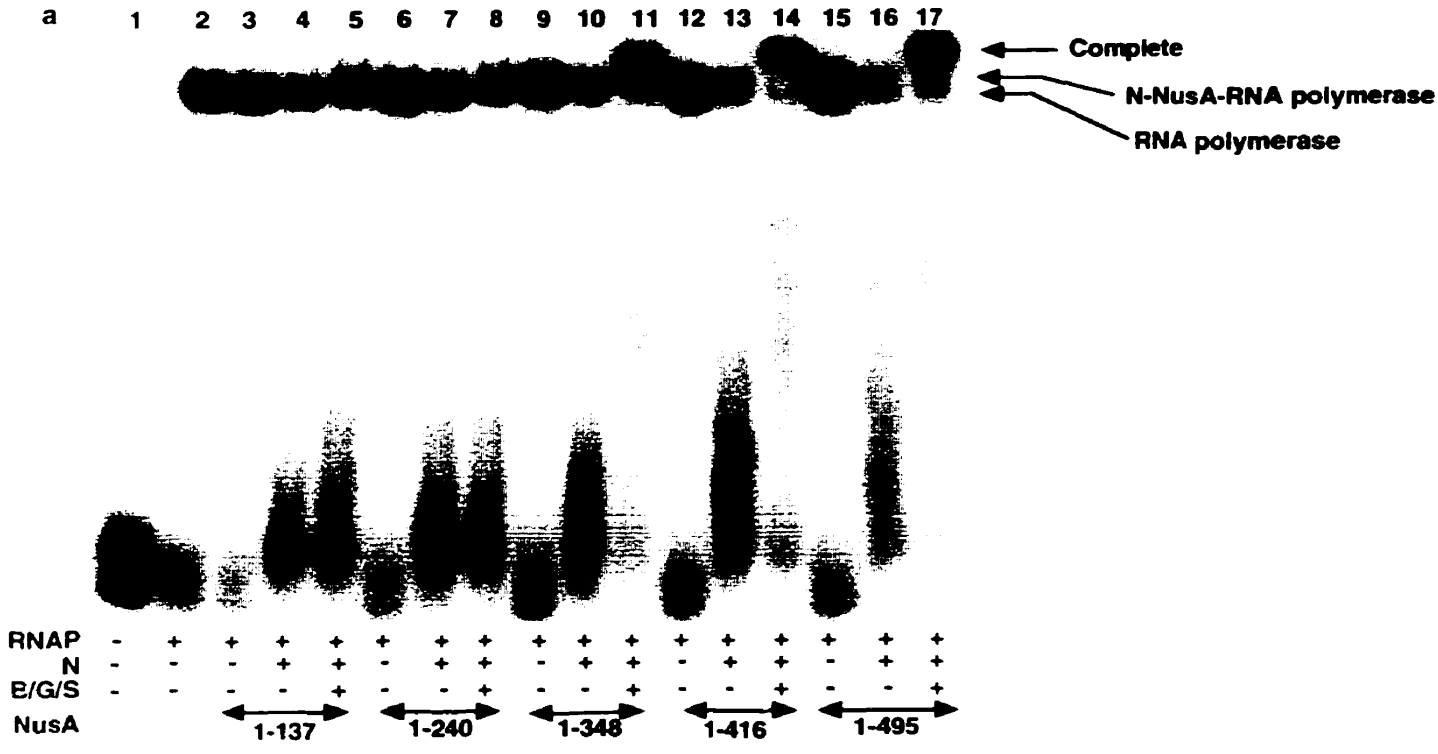
To assess the effect of progressive amino-terminal deletions on the ability of NusA to supershift an N-*nut* site complex, gel mobility shift experiments were done with NusA constructs truncated either at amino acid 132 or at amino acid 232, and thus missing either the amino-terminal RNA polymerase-binding region (amino acids 1-137) or this region plus

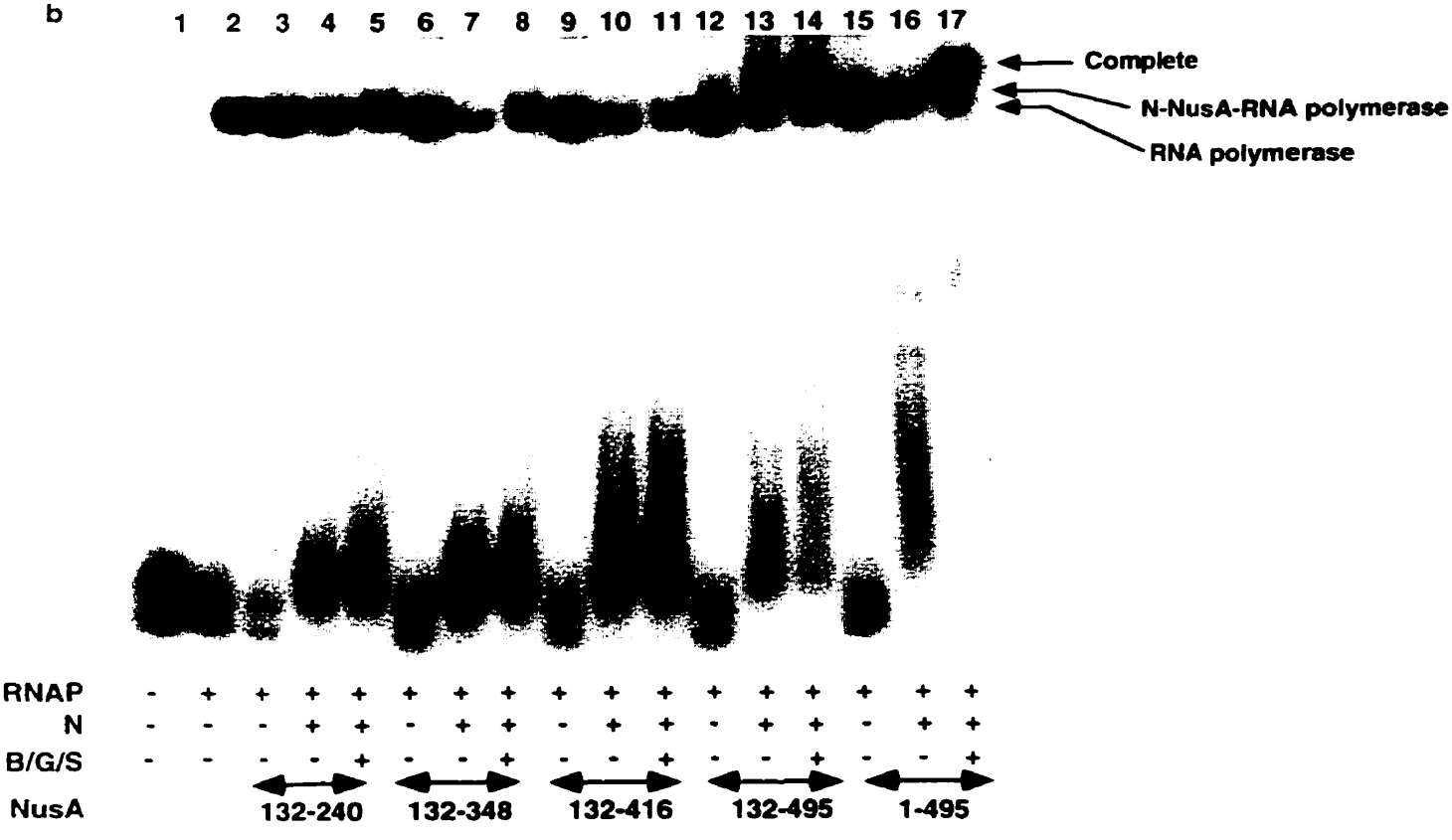


	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
N	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
NusA	-	-	-	-	-	-	+	-	-	-	-	+	-	-	-	+
NusA (1-137)	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-
NusA (1-240)	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-
NusA (1-348)	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-
NusA (1-416)	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-
NusA (132-240)	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-
NusA (132-348)	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-
NusA (132-416)	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-
NusA (132-495)	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-
NusA (232-348)	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-
NusA (232-416)	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-
NusA (232-495)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-

Figure 7 (preceding page): The effects of NusA deletions on the formation of N-NusA-*nut* site complexes. Reactions containing ^{32}P-labelled *nut* site RNA, 1.5 μM N, and 1.0 μM NusA or various NusA deletion mutants (as indicated) were electrophoresed on 7.5% non-denaturing gels, dried and exposed to film. The results from three separate gels were combined, but control lanes 1 and 2, originally present in all the gels, are shown only once.

Figure 8 (following page): The effects of NusA deletions on the formation of complete complexes containing N, the Nus factors and RNA polymerase. (a) The portion of NusA containing the KH and S1 homology regions is important for the formation of the complete complex. (b) Either RNA polymerase binding-region is sufficient to support complete complex formation. Reactions containing ^{32}P-labelled *nut* site RNA and various combinations of 500 nM N, 50 nM NusB, 50 nM NusG, 50 nM S10, 25 nM RNA polymerase and 100 nM NusA or NusA deletion mutants (as indicated) were electrophoresed on 5% non-denaturing gels, dried and exposed to film.


a
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Complete
N-NusA-RNA polymerase
RNA polymerase
RNAP
N
B/G/S
NusA
1-137
1-240
1-348
1-416
1-495


b
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Complete
N-NusA-RNA polymerase
RNA polymerase
RNAP
N
B/G/S
NusA
132-240
132-348
132-416
132-495
1-495

the S1 homology region (amino acids 136-233). NusA (132-416) and NusA (132-495) were able to support the formation of an N-NusA-*nut* site complex (Fig. 7, lanes 10 and 11). These proteins retained the S1 and KH homology regions and the N-binding region of NusA. Therefore, the amino-terminal RNA polymerase-binding region is not necessary for binding to an N-*nut* site complex even though it is necessary for NusA to enhance antitermination. As before, further carboxy-terminal deletions to amino acid 240 or 348, which eliminated the N-binding site of NusA, reduced or eliminated the ability of NusA to join the N-*nut* site complex (Fig. 7, lanes 8 and 9). None of the constructs missing the S1 homology region, as well as the amino-terminal RNA polymerase-binding region, namely NusA (232-348), NusA (232-416) and NusA (232-495), could interact with the N-*nut* site complex (Fig. 7, lanes 13-15). Since CD experiments indicated that the KH homology regions in these fragments were likely to be folded, and since constructs containing the S1 homology region and extending to amino acid 416 or 495 were able to support complex formation (Fig. 7, lanes 10 and 11), this result suggested that the S1 homology region is required for interaction with the N-*nut* site complex. This conclusion would be consistent with the earlier observation that the *nusA1* mutation in the S1 homology region prevents NusA from binding to the N-*nut* complex (Mogridge *et al.*, 1995). Thus, at least the S1 homology region of NusA and an N-binding region between amino acids 348 and 416 are required in order for NusA to associate strongly with an N-*nut* site complex.

Effects of NusA deletions on the formation of complete complexes containing N, the Nus factors, and RNA polymerase

A NusA mutant truncated at amino acid 343, which lacks the N-binding site of NusA, is partially functional for antitermination *in vivo*, and somewhat temperature-sensitive for *E.coli* growth (Tsugawa *et al.*, 1988), suggesting that most of the wild type function is

retained in this protein. To investigate whether the presence of the remaining Nus factors, NusB, NusG, and ribosomal protein S10, as well as RNA polymerase, could somehow stabilize the association of such a truncated NusA molecule with the N-*nut* site complex, I performed gel mobility shift experiments in the presence of all of the Nus factors, RNA polymerase and RNA containing a wild type *nut* site (Fig. 8a; Mogridge *et al.*, 1995). Wild type NusA supports the formation of a complete, lower mobility complex containing N, all of the Nus factors, and RNA polymerase (Fig. 8a, compare lane 17 with lanes 15 and 16; Mogridge *et al.*, 1995). The low mobility complex was still formed with NusA (1-416), which lacks the carboxy-terminal RNA polymerase-binding region but still retains the carboxy-terminal N-binding region and the amino-terminal RNA polymerase-binding region (Fig. 8a, lane 14). Interestingly, although NusA (1-348) seemed unable to support formation of the unstable N-NusA-RNA polymerase-*nut* site complex (Fig. 8a, lane 10) that was observed with NusA (1-416) (Fig. 8a, lane 13) or full length NusA (Fig. 8a, lane 16), it was able to support complete complex formation (Fig. 8a, lane 11), although not as efficiently as NusA (1-495) or NusA (1-416) (Fig. 8a, compare lane 11 with lanes 14 and 17). This result correlated well with the *in vivo* data (Tsugawa *et al.*, 1988) and suggests that the additional stability provided by NusB, NusG and S10 can partially compensate for instability caused by the loss of the N-binding and RNA polymerase-binding regions in the carboxy-terminal region of NusA. The shorter NusA deletion constructs, NusA (1-137) and NusA (1-240), were not able to support complete complex formation (Fig. 8a, lanes 5 and 8), indicating that the portion of NusA containing the KH and S1 homology regions is important for complete complex formation.

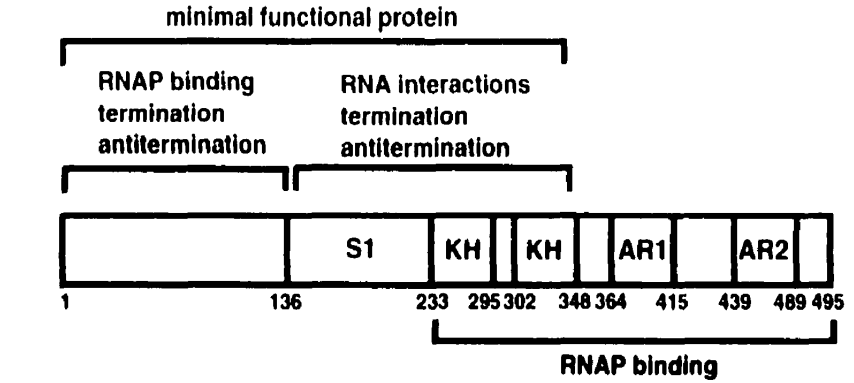
To test directly whether the amino-terminal RNA polymerase-binding region is important for forming a low mobility complex containing RNA polymerase, I used NusA constructs that lack this region in gel mobility shift experiments with N, RNA polymerase and the rest of the Nus factors (Fig. 8b). The constructs that lacked the carboxy-terminal RNA polymerase-binding region, as well as the amino-terminal RNA polymerase-binding

region, NusA (132-240), NusA (132-348) and NusA (132-416), were unable to support the formation of the complete complex (Fig. 8b, lanes 5, 8 and 11). Interestingly, the presence of the carboxy-terminal RNA polymerase-binding region in NusA (132-495) compensated for the loss of the amino-terminal RNA polymerase-binding region such that a low mobility complex was formed (Fig. 8b, compare lanes 14 and 17) whose stability was increased by NusB, NusG and S10 (Fig. 8b, compare lanes 13 and 14). Therefore, the presence of either RNA polymerase-binding region is sufficient to support formation of a low mobility stable complex even though only the amino-terminal RNA polymerase-binding region is needed for NusA to enhance antitermination (Fig. 6).

Discussion

I used deletion constructs of NusA in various assays to identify functional regions of the protein (see Fig. 9). Two RNA polymerase-binding regions were identified by affinity chromatography: one was localized to amino acids 1-137 of NusA and appears to be a folded domain; the other in amino acids 232-495 was eliminated by deleting amino acids 417-495. Although both of these regions bound about equally well to the RNA polymerase core enzyme, *in vitro* transcription assays revealed that the amino-terminal RNA polymerase-binding domain of NusA is very important for NusA's effects on termination at an intrinsic terminator and on antitermination by N, while the carboxy-terminal RNA polymerase-binding region is dispensable for both activities.

In addition to the previously observed interaction between N and full length NusA (Greenblatt and Li, 1981b), N protein also bound NusA (1-416). However, further truncation of NusA, to amino acid 348, resulted in the loss of N-binding. There are two acidic repeats in the carboxy-terminal region of NusA (Craven *et al.*, 1994). One of these repeats is lost when NusA is truncated to amino acid 416, while the second is lost when NusA is further truncated



	Binding to RNAP	Term.	Binding to N	N-NusA-*nut* complex	N-*nut*-NusA-RNAP complex	N-*nut*-NusA,B,E,G complex	Antiterm.
HIS 1-137	+	-	-	-	-	-	-
HIS 1-240	+	-	-	-	-	-	-
HIS 1-348	+	+	-	-	-	+	+
HIS 1-416	+	+	+	+	+	+	+
HIS 1-495	+	+	+	+	+	+	+
HIS 132-240	-	-	nd	-	-	-	-
HIS 132-348	-	-	nd	-	-	-	-
HIS 132-416	-	-	nd	+/-	-	-	-
HIS 132-495	+	-	nd	+	+	+	-
HIS 232-348	-	nd	nd	-	nd	nd	nd
HIS 232-416	-	nd	nd	-	nd	nd	nd
HIS 232-495	+	nd	nd	-	nd	nd	nd

Figure 9: Summary of the Results of the Deletion Analysis of NusA.
See text for details.

to amino acid 348. It may be that the basic N protein interacts with both of these regions equally well and that loss of one of the two acidic repeats has only a minor effect on N-binding.

Results from the *in vitro* antitermination assay, where loss of N-binding ability did not affect NusA function, calls into question the importance of the N-NusA interaction. However, full length NusA cannot interact with RNA on its own, and I speculate that the binding of N to the carboxy-terminal region of NusA facilitates the binding of NusA to RNA in the context of the antitermination complex. Indeed, interaction of NusA with RNA appears to be essential for antitermination because the *nusA1* point mutation in the S1 homology region of NusA (NusA R183A; Friedman, 1971) or deletion of the S1 and KH homology regions in NusA (1-137) each can impair NusA's ability to support antitermination by N. Thus, I think that N may transform wild type NusA from a termination factor into an antitermination factor by promoting a NusA-*nut* site RNA interaction that would lead to antitermination rather than an alternative NusA-RNA interaction that would lead to termination or pausing.

Formation of complexes in gel mobility shift experiments gives an indication of the stability of a complex. By using the NusA deletion constructs with this assay, it was possible to determine what regions of NusA are important for interaction with the other components of the system. When gel shifts were performed with N, NusA and the *nut* site, loss of the N-binding region or the S1 homology region prevented the formation of the N-NusA-*nut* site complex, but neither RNA polymerase-binding region was essential for the formation of this complex. When the reactions contained RNA polymerase, as well as N, NusA, and *nut* site RNA, the N-binding region and one or the other of the RNA polymerase-binding regions of NusA were critical for the formation of the characteristically unstable RNA polymerase-containing complex that functions in antitermination only over a short distance (see Figs. 8 and 9; Whalen *et al.*, 1988; DeVito and Das, 1994; Mogridge *et al.*, 1995; Rees *et al.*, 1996).

The ability of NusA (132-495) to form this complex, as well as a more stable complex in the presence of NusB, NusG, and S10, even though this complex lacks the amino-terminal RNA polymerase-binding region of NusA and is non-functional in antitermination, implies that N alters the amino-terminal termination-enhancing region of NusA so that it actually interferes with termination by RNA polymerase.

NusA (1-348) was unable to support the formation of observable amounts of the N-NusA-RNA polymerase-*nut* site complex in a gel mobility shift assay. However, as more Nus factors were added to the reaction, so that there were additional protein-protein and protein-RNA interactions, NusA (1-348) was able to support the formation of a low mobility, stable complex. Thus, NusB, NusG, and S10 can restore the stability of the complex and compensate for loss of the carboxy-terminal RNA polymerase-binding region and N-binding region in NusA (1-348). Interestingly, although the carboxy-terminal RNA polymerase-binding region could not compensate for the loss of the amino-terminal RNA polymerase-binding region in transcription assays, it could do so in the gel mobility shift experiment, suggesting that, whereas either protein-protein contact between NusA and RNA polymerase is enough to stabilize the complex, only the amino-terminal interaction has very critical consequences for termination and antitermination. This is consistent with the observation that an *E.coli* strain containing NusA (1-343), which lacks the carboxy-terminal RNA polymerase-binding region, is viable, although temperature-sensitive, and supports partial antitermination by N (Tsugawa *et al.*, 1988). Whereas NusA (1-348) supports both termination and antitermination, neither NusA (1-137), which binds RNA polymerase, nor NusA (132-348), which contains the S1 and KH homology regions, can function on its own, even though CD and NMR analysis indicate that these NusA fragments have highly ordered, folded structures. I have not determined whether a combination of NusA (1-137) and NusA (132-348) is functional.

The S1 and KH homology regions are predicted RNA-binding domains (Gibson *et al.*, 1993; Bycroft *et al.*, 1997). Gel mobility shift experiments with *nut* site RNA have

shown that point mutations in the S1 homology region are detrimental to the formation of various RNA-bound complexes containing N just as they are detrimental for antitermination (Chapter 4: Mogridge *et al.*, 1995; Y. Zhou and D. Friedman, unpublished data). My observations here that NusA (232-416) and NusA (232-495), two constructs which lack the S1 homology region and are likely to be folded, do not bind N-*nut* site complexes, also suggest that the S1 homology region is important for complex formation. Although NusA (1-240), which lacks the KH homology regions, did not support termination, antitermination, or the formation of RNA-bound complexes, CD and NMR experiments indicated that the S1 homology region in this construct is unlikely to be folded. Therefore, it is still unclear whether the KH homology regions of NusA participate in termination, antitermination, or RNA-binding.

J. Mogridge and I have shown that there are nucleotides in both *boxA* and *boxB* that are important for NusA association with the N-*nut* site complex (Chapter 4: Mogridge *et al.*, 1995). It has also been shown previously that ribosomal protein S1 specifically binds *boxA* RNA (Mogridge and Greenblatt, 1998), suggesting that the S1 homology region of NusA may interact with *boxA*. If the S1 homology region interacts with one portion of the *nut* site, it is possible that at least one of the KH homology regions interacts with the other. As there is evidence that individual KH domains have different sequence specificities (Dejgaard and Leffers, 1996), it is possible that one of the KH homology regions in NusA is required for *nut* site binding and the other is required to bind to nascent mRNA. An alternative possibility is provided by the observation that the hnRNP K protein has three KH domains and mutation or deletion of any one of them severely affects the ability of the protein to interact with C-rich sequences (Siomi *et al.*, 1994). This suggests that both KH homology regions in NusA may be required for binding one type of RNA sequence.

The amino-terminal RNA polymerase-binding domain and the S1 and KH homology regions of NusA are the minimal regions required for wild-type function in our *in vitro* assays for termination and antitermination. Further work needs to be done to elucidate the

RNA sequence preferences of the S1 and KH RNA-binding regions in NusA and to identify the precise functions of the critical interaction between the amino-terminal domain of NusA and RNA polymerase.

Experimental procedures

Plasmids, enzymes, and strains

RNA polymerase, N, NusA, NusB, and NusG were purified as previously described (Burgess and Jendrisak, 1975; Greenblatt *et al.*, 1980; Greenblatt *et al.*, 1981; Swindle *et al.*, 1988; Li *et al.*, 1992). Purified S10 was generously provided by Dr. Volker Nowotny. All plasmids were prepared in the bacterial strain DH5α (Life Technologies, Inc.) and purified on Qiagen-tip 500 columns (Qiagen). The oligonucleotides used for cloning were purchased from ACGT Corp. (Toronto). RNAguard was bought from Pharmacia Biotech. Restriction enzymes and DNA ligase were purchased from New England Biolabs. T7 RNA polymerase was obtained from Life Technologies, Inc.

Construction of His_6-tagged NusA Proteins

PCR primers were designed to amplify *NusA* or fragments of *NusA* from the plasmid pJL4. Forward and reverse primers contained *Nco*I and *Bam*HI restriction sites, respectively, for subsequent cloning into the vector, pET-11d (Novagen). The forward primers also contained the sequence for six histidines.

Purification of His_6-tagged NusA Proteins

The *E.coli* strain BL21 ($pREP_4$) containing the His_6-tagged NusA plasmids was grown in 1l of LB medium to an A_{600} of 0.5 and induced for 3 hours with 2 mM isopropyl-1-thio-β-D-galactopyranoside. Cells were harvested by centrifugation for 20 minutes at 4500 rpm in a Sorvall H6000A swinging bucket rotor, resuspended in 10 mL binding buffer (5

mM imidazole, 0.5 M NaCl, 20 mM Tris [pH 7.9], 10% glycerol, 5 mM β-mercaptoethanol), sonicated, and centrifuged for 30 minutes at 15 000 rpm in a Sorvall SS34 rotor. The extracts were mixed with 0.25 mL His Bind beads (Qiagen) and incubated at 4°C for 2 hours. The beads were added to 1ml Bio-Spin columns (BioRad) and washed sequentially with 2 ml binding buffer, 4 ml wash buffer (45 mM imidazole, 0.5 M NaCl, 20 mM Tris, [pH 7.9], 10% glycerol, 5 mM β-mercaptoethanol) and 1 ml ACB (10 mM HEPES [pH 7.0], 10% glycerol, 0.1 mM EDTA, 5 mM β-mercaptoethanol containing 0.1 M NaCl). Columns were eluted with 1 ml ACB containing 0.1 M NaCl, 0.4 M imidazole [pH 7.9], and 1 mM DTT.

Affinity Chromatography

For experiments with GST-N, extract containing GST or GST-N (Mogridge *et al.*, 1998a) was incubated with Glutathione Sepharose 4B beads (Pharmacia Biotech) and rotated at room temperature for 15 min. The beads were washed sequentially with 1M NaCl Buffer A (20 mM Tris-HCl [pH 7.8], 0.2 mM EDTA, 1 mM DTT, 1 mM PMSF) and 0.1 M NaCl Buffer A. Beads (20 µL) were added to 200 µl pipette tips that contained 10 µl of 212-300 micron glass beads (Sigma). The column bed was washed with 10 column volumes of 1 M NaCl ACB (10 mM HEPES [pH 7.0], 10% glycerol, 0.1 mM EDTA, 1 mM DTT) and then washed with 10 column volumes of 100 mM NaCl ACB. Columns were loaded with 100 µL of a mixture of His_6-tagged NusA and His_6-tagged NusA deletion mutants (3.0 µg each) in 100 mM NaCl ACB buffer supplemented with 0.2 mg/ml insulin. The columns were washed with 10 column volumes of 100 mM NaCl ACB and eluted with 1 M NaCl ACB. One quarter of the eluate was loaded in the gel.

For experiments with RNA polymerase and the His_6-tagged NusA and the His_6-tagged NusA deletion mutants, purified His_6-tagged protein was incubated with Ni-agarose (Qiagen) for 30 min. 20 µl of beads were added to 200 µl pipette tips and treated as described above except

that these columns were loaded with 100 µL of 1 mg/mL *E.coli* extract containing 6 µg additional RNA polymerase core enzyme.

In vitro Transcription

Transcription reactions were performed as previously described (Whalen and Das, 1990) using the template pJD12 (generous gift of A. Das). Concentrations of proteins are indicated in the figure legends. Levels of terminated and run-off transcripts were quantitated using a phosphoimager.

Gel Mobility Shift Experiments

Gel mobility shift assays were performed as previously described (Mogridge *et al.*, 1995). Concentrations of proteins are indicated in the figure legends.

CD and NMR Spectroscopy

Circular dichroism was performed in an Aviv 62A DS circular dichroism spectrometer. All measurements were collected at 25°C in 10 mM Tris-HCl, 0.1 mM EDTA, 250 mM NaCl and 0.1 mM DTT in a 0.1 cm cuvette. Spectra were collected at a scanning speed of 1 nm/sec from 300 nm to 200 nm. ^{15}N-labeled fragments of NusA were prepared and HSQC NMR spectra were generated as described previously for the λ N protein (Mogridge *et al.*, 1998).

Chapter 3

The α Subunit of *E.coli* RNA Polymerase Activates RNA-Binding by NusA

Abstract

NusA modulates pausing, termination, and antitermination by associating with the *E.coli* RNA polymerase core enzyme. Although NusA cross-links to nascent RNA within a transcription complex, it does not associate directly with RNA on its own. I have shown here that both the carboxy-terminal domain of the α subunit of RNA polymerase and the bacteriophage λ *N* gene antiterminator protein bind to carboxy-terminal regions of NusA. Partial deletion of these same regions allowed NusA to bind RNA on its own. Interaction with α (or the λ N protein) resulted in RNA-binding that involves the S1 homology region of NusA. The interaction of α with NusA may enable NusA to bind the nascent transcript and thus stimulate pausing and termination. The λ N protein may reverse the effects of NusA on pausing and termination by inducing NusA to interact with N-utilization (*nut*) site RNA rather than RNA near the 3' end of the nascent transcript.

Introduction

The NusA protein of *E.coli* binds to core RNA polymerase shortly after the initiation of transcription and stimulates pausing and termination at certain sites (reviewed in Richardson and Greenblatt, 1996). The mechanism by which NusA influences pausing during transcription is not yet clear but RNase protection experiments suggest that NusA may bind and stabilize the stem-loop RNA structures often associated with pause sites (Landick and Yanofsky, 1987).

NusA binds directly to the α subunit of RNA polymerase (Liu *et al.*, 1996). It can also be cross-linked to the large β and β' subunits of RNA polymerase (J. Li and J. Greenblatt, unpublished data) and may be capable of binding directly to these subunits as well (Liu *et al.*, 1996). The α subunit of RNA polymerase has two domains: the amino-terminal domain (NTD) is required for dimerization and for interaction with the β and β' subunits of RNA polymerase, whereas the carboxy-terminal domain (CTD) is a contact surface for DNA-binding activator proteins (reviewed in Ebright and Busby, 1995). The CTD of α also binds the UP element, a DNA element that enhances initiation of transcription at certain promoters (Ross *et al.*, 1993; Blatter *et al.*, 1994). Additionally, it has been suggested that a direct interaction between the α-CTD and NusA is important for NusA's ability to control pausing and termination (Liu *et al.*, 1996).

NusA influences not only pausing and termination by RNA polymerase but also transcriptional antitermination by the bacteriophage λ N protein (Friedman, 1971). Antitermination by N requires a cis-acting RNA element, called the *nut* site, which consists of two functional components, *boxA* and *boxB* (Salstrom and Szybalski, 1978; de Crombrugghe *et al.*, 1979; Olson *et al.*, 1982; Das and Wolska, 1984; Horwitz *et al.*, 1987). NusA, as well as NusB, NusE (ribosomal protein S10), NusG, RNA polymerase and the *nut*

site RNA, take part in multiple interactions within the N-modified transcription complex (reviewed in Friedman, 1988; Greenblatt *et al.*, 1993). The resulting highly stable ribonucleoprotein complex is capable of suppressing transcription termination over long distances and through multiple terminators (Mason *et al.*, 1992; Mogridge *et al.*, 1995). Within this complex, NusA interacts with the N protein (Greenblatt and Li, 1981b), and both the amino- and carboxy-terminal regions of NusA interact with RNA polymerase (Chapter 2).

There is evidence that NusA may interact directly with nucleotides in both the *boxA* and *boxB* components of the *nut* site (Olson *et al.*, 1982; Friedman and Olson, 1983; Olson *et al.*, 1984; Mogridge *et al.*, 1995). The effects of mutations in the S1 homology region of NusA between amino acids 136 and 240 suggest that this region is important for antitermination. The *nusA1* (L183R) and *nusA R199A* mutations both cause temperature-sensitive λ growth because of an inability of N to function at high temperature (Y. Zhou and D. Friedman, unpublished data; Friedman, 1971; Friedman and Baron, 1974). Unlike wild type NusA, both mutant proteins are unable to supershift an N-*nut* site RNA complex in a gel mobility shift experiment, even though they bind N directly with wild-type affinity (Chapter 4; T. Mah and J. Greenblatt, unpublished results; Mogridge *et al.*, 1995). This suggests that both mutations cause a defect in the interaction of the S1 homology region of NusA with *nut* site RNA. Other experiments have shown that both an amino-terminal RNA polymerase-binding region in amino acids 1-137 and a portion of NusA that contains the S1 and KH homology regions are essential for NusA to enhance both termination at an intrinsic terminator and antitermination by N (Chapter 2).

Despite this abundant evidence suggesting that NusA can interact with RNA, it does not associate directly with RNA in a gel retardation assay (Mogridge *et al.*, 1995). Rather, its ability to interact with phage λ *nut* site RNA depends on the presence of N (Mogridge *et al.*, 1995), which interacts directly with NusA (Greenblatt and Li 1981). Since NusA does cross-link to RNA in a transcription complex in the absence of N (Liu and Hanna, 1995), it seemed

likely that an interaction of NusA with RNA polymerase might also alter the conformation of NusA so as to allow for RNA-binding. I report here that interaction of the RNA polymerase α subunit with NusA promotes the interaction of NusA with RNA. My observation that α and the λ N protein bind to similar regions near the carboxy-terminus of NusA suggests that α and N may act in similar ways to control the binding of NusA to RNA. My results suggest that an interaction of NusA with α in a transcription complex would allow NusA to bind the nascent transcript and stimulate pausing and termination by RNA polymerase. The λ N protein may reverse the effects of NusA on pausing and termination by causing NusA to interact with *nut* site RNA rather than the RNA near the 3' end of the nascent transcript.

Results

Binding of NusA to RNA in the presence of the RNA polymerase α subunit

In view of the known involvement of the α subunit of RNA polymerase in NusA function (Liu *et al.*, 1996), I tested whether α could promote RNA-binding by NusA in gel mobility shift assays containing ^{32}P-labeled RNA with a wild type λ *nut* site (Fig. 1). The addition of increasing amounts of α to a constant amount of NusA resulted in the appearance of two bands, and sometimes a weak third band, with lower mobility than the free RNA (Fig. 1a, compare lane 1 with lanes 7-10). These bands were absent in lanes containing either NusA alone (Fig. 1a, lane 2) or α alone (Fig. 1a, lanes 3-6), indicating that formation of a complex on the RNA required both proteins. In the converse experiment, the concentration of α was held constant and that of NusA was varied (Fig. 1b). In this case, NusA alone was not able to bind *nut* site-containing RNA even at very high concentrations (Fig. 1b, lanes 2-4), but complexes of lower mobility appeared and increased in intensity as the concentration

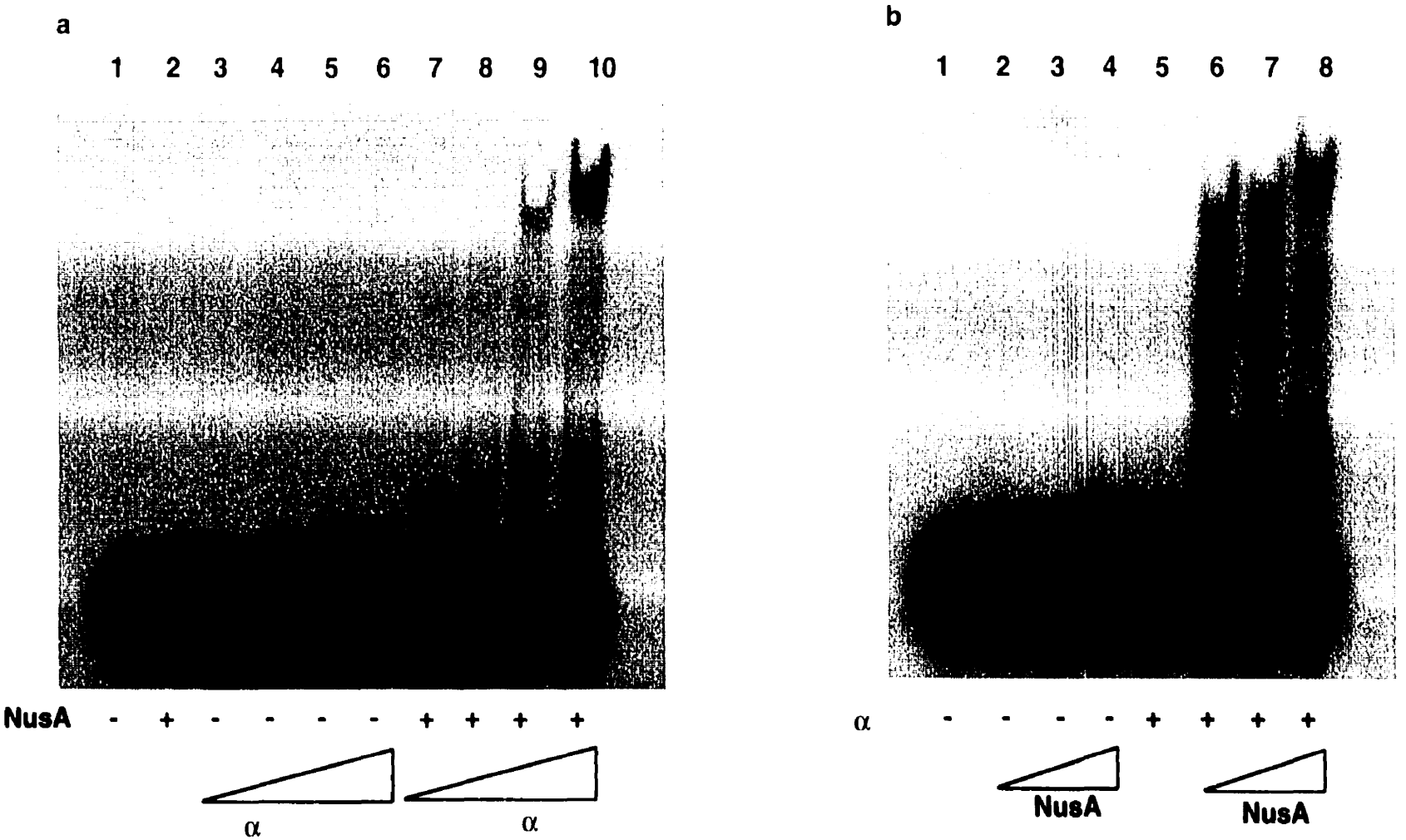


Figure 1: Binding of NusA to RNA in the presence of the RNA polymerase α subunit. (a) Addition of increasing amounts of α to a constant amount of NusA results in an increase in complex formation. Reactions containing ^{32}P-labelled *nut* site RNA and various combinations of 14 μM NusA and 1.25, 2.5, 5, or 10 μM α (as indicated) were electrophoresed on 7.5% non-denaturing gels, dried, and exposed to film. (b) Addition of increasing amounts of NusA to a constant amount of α results in increased complex formation. Reactions containing ^{32}P-labelled *nut* site RNA and various combinations of 9 μM α and 3.5, 7, or 14 μM NusA (as indicated) were electrophoresed on 7.5% non-denaturing gels, dried, and exposed to film.

of NusA was increased in the presence of α (Fig. 1b, lanes 6-8). In view of the known ability of α to dimerize primarily via its amino-terminal domain (Blatter *et al.*, 1994; Kimura *et al.*, 1994), these complexes may represent different combinations of α and NusA.

The CTD of the RNA polymerase α subunit stimulates RNA-binding by NusA

Because the CTD of the RNA polymerase α subunit is known to be important for NusA activity in pausing and termination (Liu *et al.*, 1996), I tested whether the α-CTD (amino acids 249-329) or α-NTD (amino acids 1-235) alone could promote RNA-binding by NusA (Fig. 2). I added full length α, α-CTD or α-NTD to gel mobility shift reactions with NusA and ^{32}P-labeled *nut* site-containing RNA. Neither the α subunit nor its isolated amino-terminal and carboxy-terminal domains alone could bind to the RNA (Fig. 2, lanes 3-6). By contrast, a carboxy-terminally truncated NusA derivative containing amino acids 1-416 is capable of direct binding to *nut* site-containing RNA (Fig. 2, lane 2; also see below). The addition of either intact α (Fig. 2, lanes 7 and 8) or the α-CTD (Fig. 2, lanes 9 and 10) to intact NusA (amino acids 1-495) caused shifts in the mobility of the RNA. The mobility of the complex formed with the α-CTD and NusA (Fig. 2, lanes 9 and 10) was similar to that of the more rapidly migrating complex obtained when full length α was incubated with NusA (Fig. 2, lanes 7 and 8). Since the α-CTD lacks the principle dimerization domain of α (Blatter *et al.*, 1994; Kimura *et al.*, 1994), the complex obtained with NusA and the α-CTD likely contains only one molecule each of the α-CTD and NusA. The additional lower mobility complex in the reactions containing α and NusA (Fig. 2, lanes 7 and 8) likely contains two molecules of α and one or more molecules of NusA as a consequence of dimerization of the α subunits. In contrast, no distinct shift was obtained with NusA and the

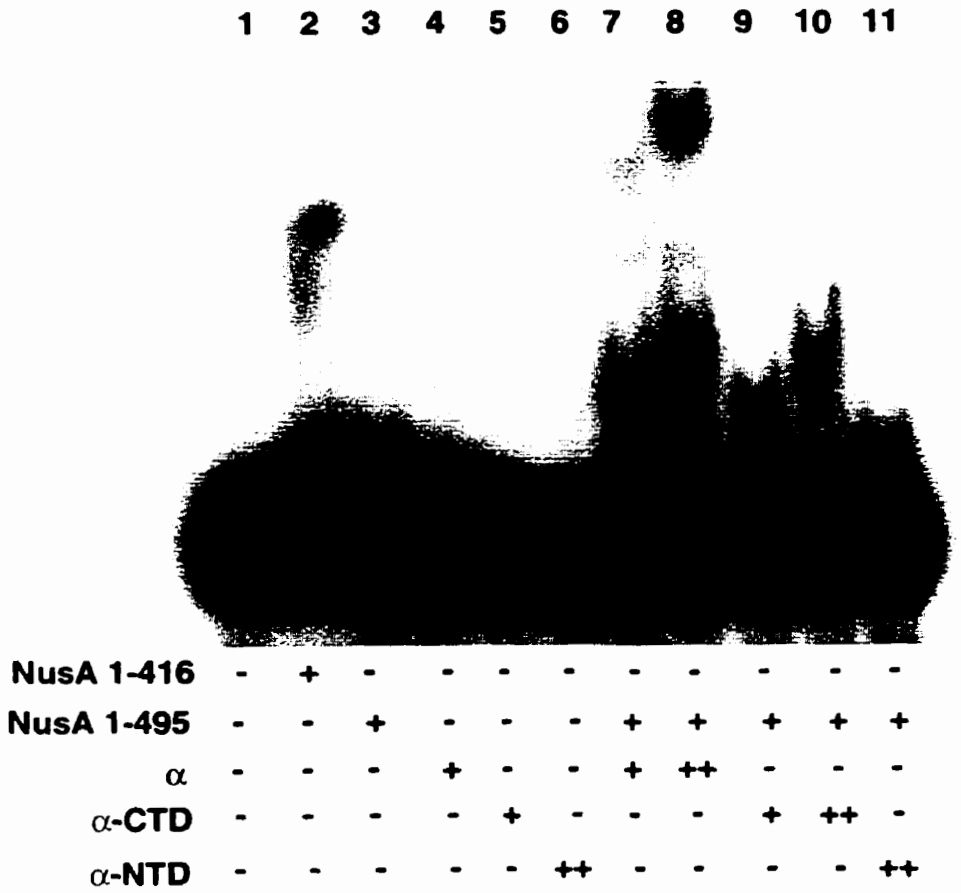


Figure 2: α-CTD stimulates RNA-binding by NusA. Reactions containing 32P-labelled *nut* site RNA and various combinations of 13 μM NusA (1-416), 4.5 or 9 μM α, 4.5 or 9 μM αCTD or 11 μM αNTD (as indicated) were electrophoresed on 7.5% non-denaturing gels, dried, and exposed to film.

α-NTD (Fig. 2, lane 11). These results indicate that the CTD of the RNA polymerase α subunit is capable of stimulating RNA-binding by NusA.

RNA-binding by NusA in the presence of α is sequence-specific and sensitive to a mutation in the S1 homology region of NusA

To evaluate potential RNA sequence or structure specificity in the RNA-binding observed with NusA and α, I compared the abilities of RNAs containing either a wild type *nut* site or a *nut* site in which the sequence of the *boxA* element had been switched from 5'–3' to 3'-5' (*boxA* reverse) (Mogridge *et al.*, 1995) to support formation of NusA-α-*nut* site complexes (Fig.3a). Whereas N only requires the *boxB* RNA element for binding, NusA is unable to supershift an N-*nut* site complex when *boxA* is reversed (Mogridge *et al.*, 1995). This suggests that there is a direct and specific interaction between *boxA* and an RNA-binding domain in NusA. As shown in Fig. 3a, the low mobility complexes formed with the wild type probe, NusA, and α were not present in reactions when the reverse *nut* site probe was used (Fig. 3a, compare lane 4 with lane 8). Thus, the NusA-binding promoted by α in these gel mobility shift experiments has structure or sequence-specificity. Moreover, the importance of the *boxA* element for RNA-binding provided additional evidence that the RNA-binding observed in experiments containing NusA and α involves NusA.

The *nusA R199A* mutation causes a defect in antitermination by N *in vivo*, as well as a defect in the ability of NusA to supershift an N-*nut* site complex, even though the NusA R199A mutant protein binds with normal affinity to N (Chapter 4; T. Mah, Y. Zhou, J. Greenblatt and D. Friedman, unpublished data). Therefore, this mutation in the S1 homology region of NusA appears to cause a defect in the binding of NusA to *nut* site RNA. The *nusA R199A* mutation also prevented NusA from binding the *nut* site RNA in the presence of α (Fig. 3b, compare lanes 6 and 7). This result indicates that the S1 homology region of NusA

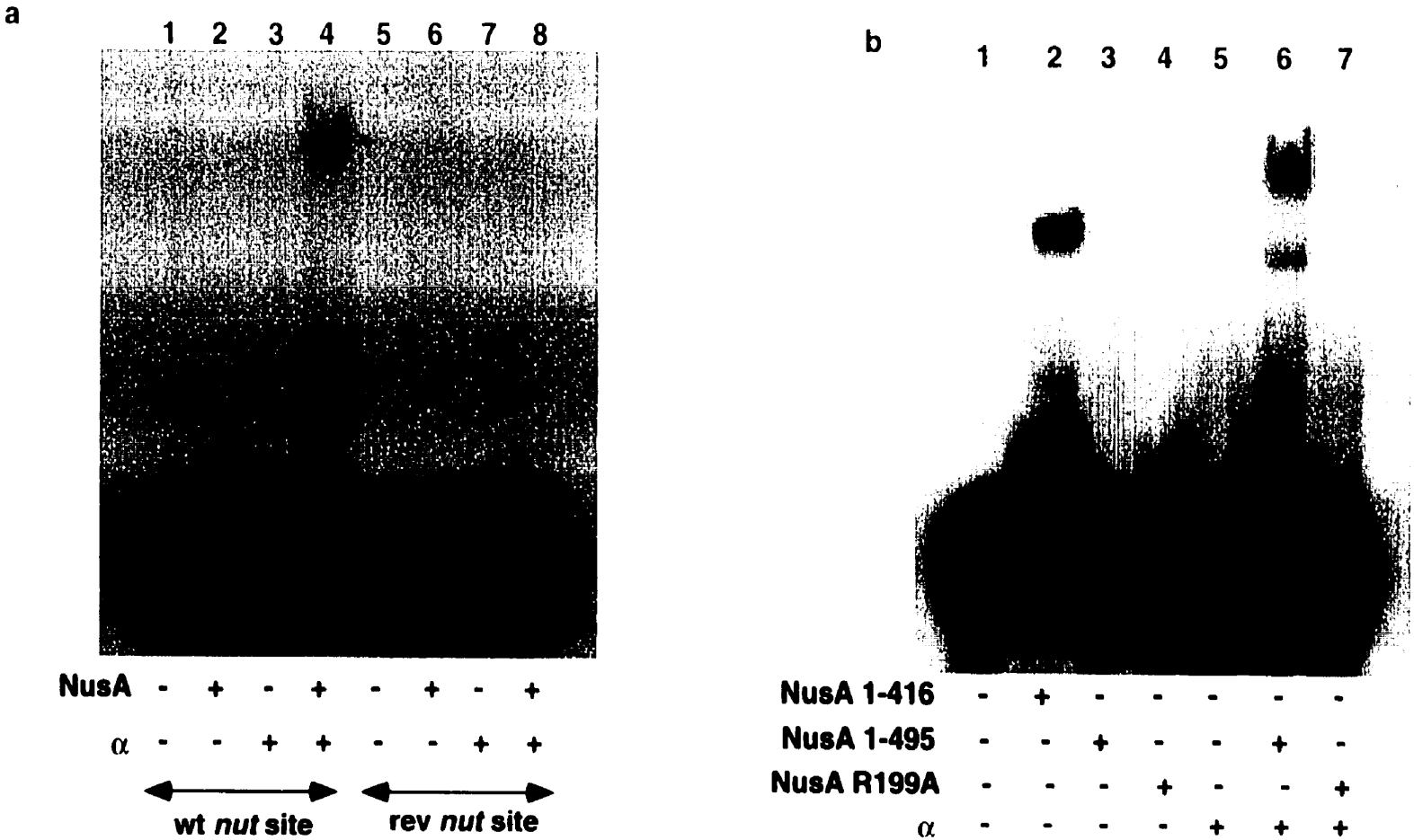


Figure 3: RNA-binding by NusA in the presence of α is sequence-specific and sensitive to a mutation in the S1 domain of NusA. (a) RNA-binding by NusA in the presence of α is prevented by a mutation in the *boxA* portion of the *nut* site. Reactions containing wild type or mutant 32P-labelled *nut* site RNA (as indicated) and various combinations of 10 μM α and 14 μM NusA (as indicated) were electrophoresed on 7.5% non-denaturing gels, dried, and exposed to film. (b) RNA-binding by NusA in the presence of α is prevented by a mutation in the S1 domain of NusA. Reactions containing 32P-labelled *nut* site RNA and various combinations of 14 μM NusA, 14 μM NusA (1-416), or 12 μM NusA R199A and 11 μM α (as indicated) were electrophoresed on 7.5% non-denaturing gels, dried, and exposed to film.

is likely to participate in *nut* site-binding stimulated by α and provides further evidence that RNA-binding observed in the presence of α and NusA reflects direct RNA-binding by NusA.

Interaction of α with the carboxy-terminal region of NusA

My observation that NusA required α to bind *nut* site RNA suggested that there may be a direct interaction between NusA and α. To test for a direct interaction between α and NusA, a mixture of full-length NusA and three carboxy-terminally deleted mutant proteins (Fig. 4a, lane 2) was passed over columns containing various concentrations of covalently bound α. Specific binding to α was observed only for the full-length NusA protein, whose binding increased in concert with the α concentration on the column (Fig. 4a, lanes 4-6). Since NusA (1-416) did not bind to α, the carboxy-terminal 79 amino acids of NusA are necessary for the binding of α to NusA. To further establish which regions of NusA are sufficient for interaction with α, purified α was tested for binding to various covalently-immobilized portions of NusA: an amino-terminal region, NusA (1-137), that I have shown binds RNA polymerase (Chapter 2); a carboxy-terminal fragment, NusA (303-495); and the full length protein (Fig. 4b). α did not bind to the amino-terminal fragment of NusA (Fig. 4b, lane 4). Appreciable binding to the carboxy-terminal region of NusA was observed (Fig. 4b, lane 5), and the binding of α was best with full length NusA (Fig. 4b, lane 6). It appears that α interacts primarily with the carboxy-terminal region of NusA, although the amino-terminal two-thirds of NusA may also contribute to the binding.

Since the α-CTD but not the α-NTD allowed RNA-binding by NusA (Fig.2), I also tested which region of α was involved in the direct binding of α to NusA (Fig. 4c). Full length α, α-NTD and α-CTD were mixed together (Fig. 4c, lane 2) and loaded onto columns containing GST (Fig. 4c, lane 3) or various concentrations of GST-NusA (303-495) (Fig. 4c,

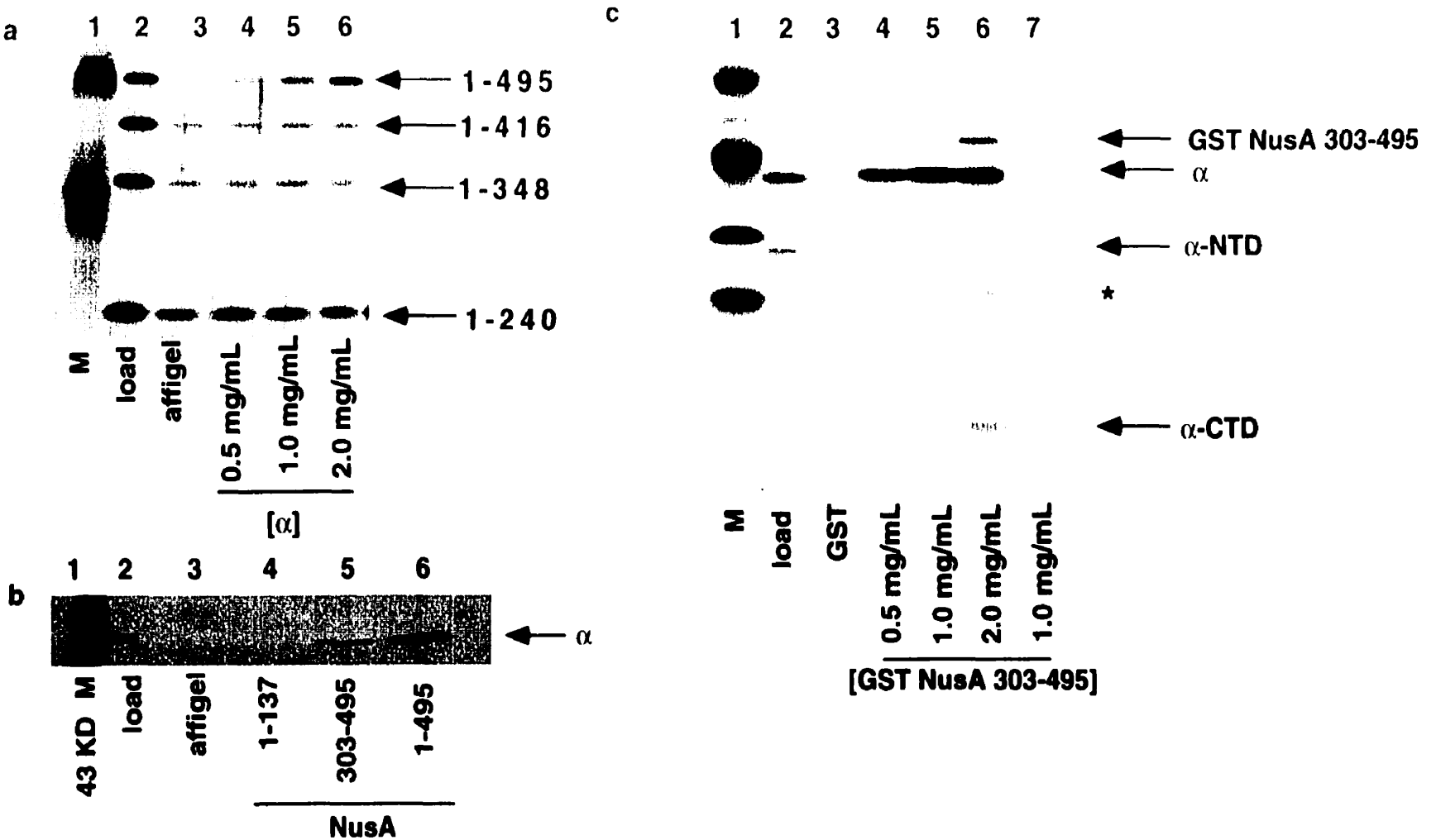


Figure 4: Interaction of α with the carboxy-terminal region of NusA. (a) Carboxy-terminal turncation of NusA prevents the α-NusA interaction. A mixture of four His_6-tagged NusA proteins (lane 2) was passed over columns containing affigel (lane 3) or increasing amounts of affigel-coupled α (lanes 4-6). (b) The carboxy-terminal region of NusA interacts directly with α. Buffer containing α and 0.2 mg/ml insulin (lane 2) was passed over columns containing affigel (lane 3) or affigel-coupled NusA (2 mg/ml) (lane 6) or affigel-coupled regions of NusA (lanes 4 and 5) (as indicated). The concentrations of amino- and carboxy-terminal regions of NusA on the columns were adjusted so that each had the same molar concentration as the full length NusA. (c) α_{CTD} interacts with the carboxy-terminal 192 amino acids of NusA. A mixture of α, α_{CTD}, and α_{NTD} (lane 2) was passed over columns containing GST (lane 3) or increasing amounts of GST- NusA (303-495) (lanes 4-6). As a control, buffer alone was passed over a GST-NusA (303-495) column (lane 7). Bound proteins were eluted with buffer containing 1M NaCl, subjected to SDS-PAGE, and stained with silver.
* indicates a degredation product of α, as identified by mass spectrometry.

lanes 4-6). The bound proteins were eluted with salt. None of the α fragments were retained on the control GST column. In contrast, as the concentration of immobilized GST-NusA (303-495) on the columns was increased, increasing amounts of full length α and α-CTD were present in the salt eluates from this matrix. The binding of α-NTD to the immobilized GST-NusA (303-495) was barely detectable (Fig. 4c, lanes 4-6). Therefore, the result of this direct protein-protein binding study was consistent with the results of the gel mobility shift experiments, which also indicated that only full-length α and the α-CTD would interact with NusA.

A carboxy-terminally truncated NusA also binds specifically to *nut* site RNA

My observations that α could provoke RNA-binding by NusA and interact with NusA (1-495), but not with NusA (1-416) suggested that the carboxy-terminal 79 amino acids of NusA might inhibit the RNA-binding activity of NusA. Unlike full-length NusA, NusA (1-416) could bind RNA containing a wild type *nut* site in a gel mobility shift experiment (Fig. 5a, lanes 2-7), but bound only weakly to RNA containing a *nut* site with a reversed *boxA* sequence (Fig. 5a, lanes 12-14). This result indicates that the binding of NusA to *nut* site RNA is indeed inhibited by the carboxy-terminal 79 amino acids of NusA and suggests that this inhibition could be relieved by an interaction of the carboxy-terminal region of NusA with the CTD of the RNA polymerase α subunit.

In order to further characterize the α-independent RNA-binding by NusA (1-416), other deletion mutants of NusA were tested for their ability to bind RNA containing a wild type *nut* site in a gel mobility shift experiment (Fig. 5b). Truncation of NusA to amino acid 348 weakened RNA-binding, even though NusA (1-348) retains the S1 and KH homology regions (Fig. 5b, lane 4). The deletion of NusA's amino-terminal RNA polymerase-binding region, as in NusA (132-416), prevented RNA binding by NusA (Fig. 5b, lane 2). Therefore,

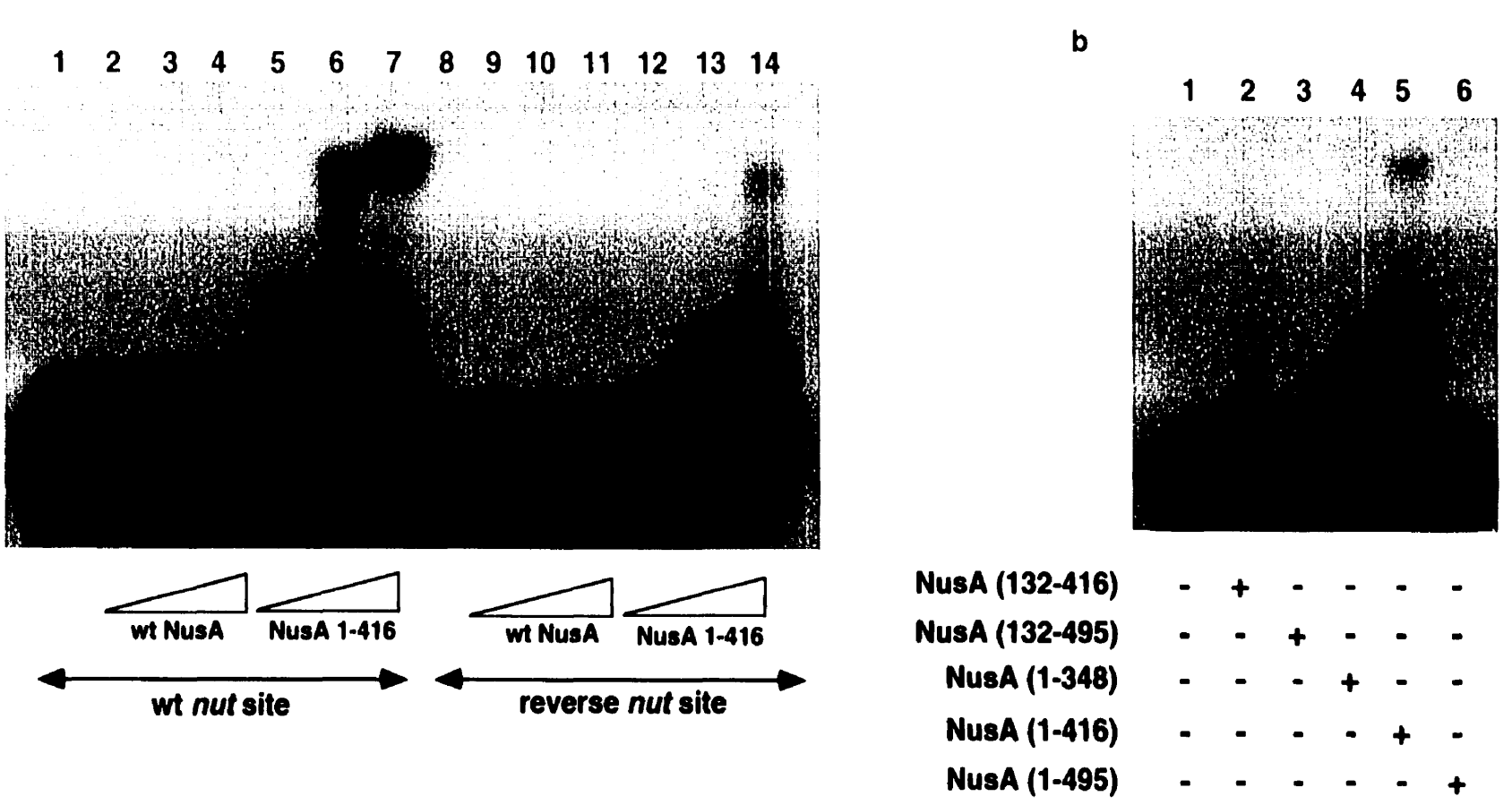


Figure 5: A carboxy-terminally truncated NusA binds specifically to *nut* site RNA
(a) RNA-binding by NusA (1-416) is prevented by a mutation in the *boxA* portion of the *nut* site. Reactions containing wild type or mutant ^{32}P-labelled *nut* site RNA (as indicated) and 3.5, 7, or 14 μM NusA or NusA (1-416) (as indicated) were electrophoresed on 7.5% non-denaturing gels, dried, and exposed to film. (b) RNA-binding by NusA (1-416) is stabilized by amino acids 1-137 and 348-415 of NusA. Reactions containing ^{32}P-labelled *nut* site RNA and 10 μM NusA (132-416), NusA (132-495), NusA (1-348), NusA (1-416) or NusA (as indicated) were electrophoresed on 7.5% non-denaturing gels, dried, and exposed to film.

the amino-terminal RNA polymerase-binding region in amino acids 1-137 of NusA, as well as amino acids 348 to 415 of NusA, appear to be required to stabilize the RNA-binding ability of NusA (1-416).

The λ N protein also binds the carboxy-terminal region of NusA

The N protein of the bacteriophage λ also binds NusA (Greenblatt and Li, 1981b) and I have shown that this binding requires a carboxy-terminal region of NusA (Chapter 2). Moreover, NusA (1-348), which lacks this region, cannot supershift an N-*nut* site complex (Chapter 2). In view of the possibility that α and N might act in similar ways to provoke RNA-binding by NusA, I tested various fragments of NusA for binding to N. Full length NusA, a carboxy-terminally truncated NusA, NusA (1-399), and a carboxy-terminal fragment of NusA, NusA (303-495), were mixed with *E.coli* extracts and passed over GST and GST-N affinity columns (Fig. 6). None of the major *E.coli* proteins present in the extracts applied to the columns bound to the GST or GST-N columns (Fig. 6, lanes 2, 3, 5, 6, 8 and 9). Both NusA and NusA (303-495) bound selectively to GST-N and therefore were present in the high salt eluates from the GST-N columns, but not in GST control column eluates (Fig. 6, lanes 2, 3, 8, and 9). NusA (1-399) was absent from the high salt eluates of both columns (Fig. 6, lanes 5 and 6). Thus, N binds directly to the carboxy-terminal region of NusA. This result suggested that N might activate the RNA-binding ability of NusA by binding to the same region of NusA that I have shown binds α.

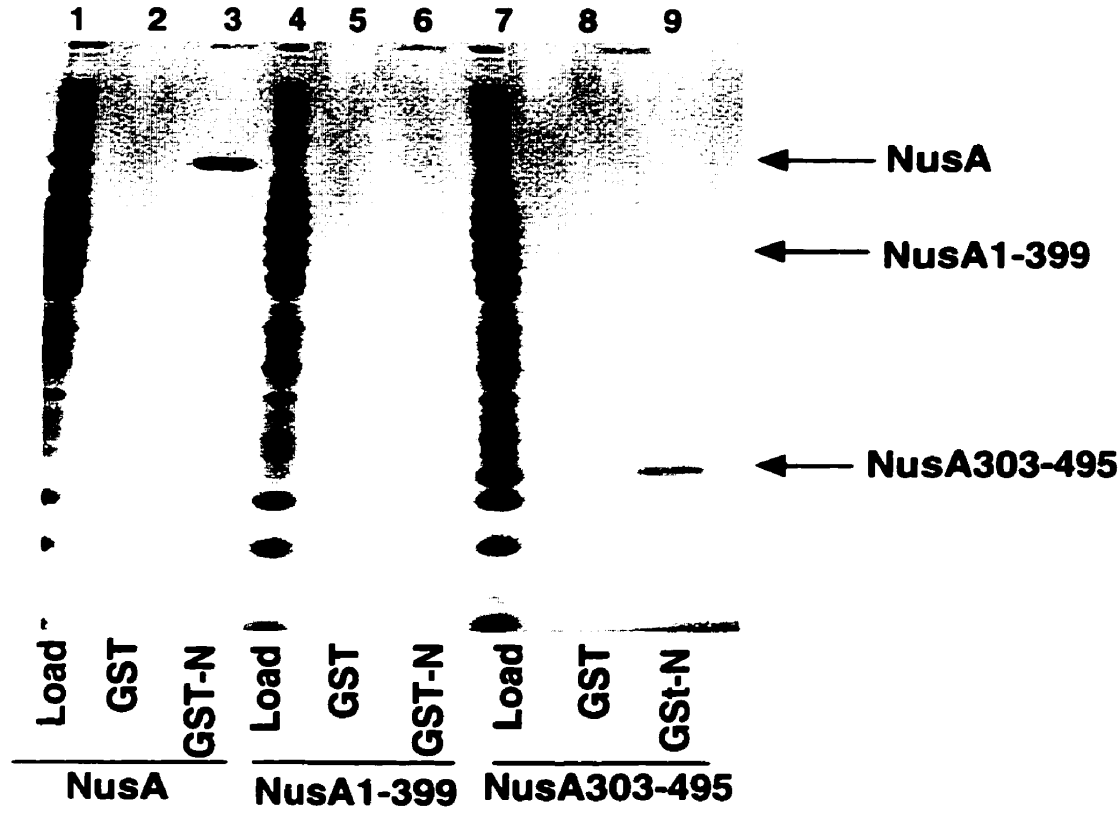


Figure 6: The λ N protein binds directly to the carboxy-terminal 192 amino acids of NusA. *E.coli* extract containing additional NusA (lane 1), NusA (1-399) (lane 4) or NusA (303-495) (lane 7) was passed over columns containing 2 mg/ml GST (lanes 2, 5, and 8) or 0.5 mg/ml GST-N (lanes 3, 6, and 9). Bound proteins were eluted with buffer containing 1M NaCl, subjected to SDS-PAGE, and stained with silver.

Discussion

An Autoinhibition Domain May Inhibit RNA-binding by NusA

Whereas full length NusA does not bind RNA, I have shown here that a carboxy-terminal deletion mutant, NusA (1-416), which retains the S1 and KH homology regions of NusA, can bind RNA in the absence of α or N. This suggests that one or more of the RNA-binding domains of NusA may be occluded by the carboxy-terminal 79 amino acids of NusA (see Fig. 7a). Since neither NusA (1-348) nor NusA (132-416) binds RNA as strongly as NusA (1-416), amino acids 1-131 and 349-416 must also alter the folding, alignment, or accessibility of the RNA-binding domains. Interaction of the carboxy-terminally truncated NusA (1-416) with RNA is sensitive to alteration of the *boxA* portion of the *nut* site (Fig. 5a). This is also the case when full length NusA binds the *nut* site RNA in the presence of N (Mogridge *et al.*, 1995) or α (Fig. 3a). RNA-binding by NusA (1-416) is also inhibited when alterations in the loop of *boxB* prevent *boxB* from forming a GNRA tetraloop-like structure (see Chapter 4; Legault *et al.*, 1998). It is unclear, however, which domains of NusA (1-416) contribute to the sequence- or structure-specific interactions of NusA with *nut* site RNA. The inability of full-length NusA to bind RNA resembles the inability of the intact initiation subunit σ^{70} of RNA polymerase to bind DNA unless it is part of RNA polymerase holoenzyme (Dombroski *et al.*, 1992; Dombroski *et al.*, 1993). Just as deletion of the carboxy-terminal region of NusA enables NusA to bind RNA in the absence of RNA polymerase, deletion of the amino-terminal 130 amino acids of σ^{70} enables σ^{70} to interact specifically and non-specifically with DNA in the absence of the other RNA polymerase subunits (Dombroski *et al.*, 1992). The autoinhibition of DNA-binding by σ^{70} may be relieved when σ^{70} interacts with the core polymerase subunits and undergoes a conformational change that uncovers or reorients its DNA-binding domains (Dombroski *et*

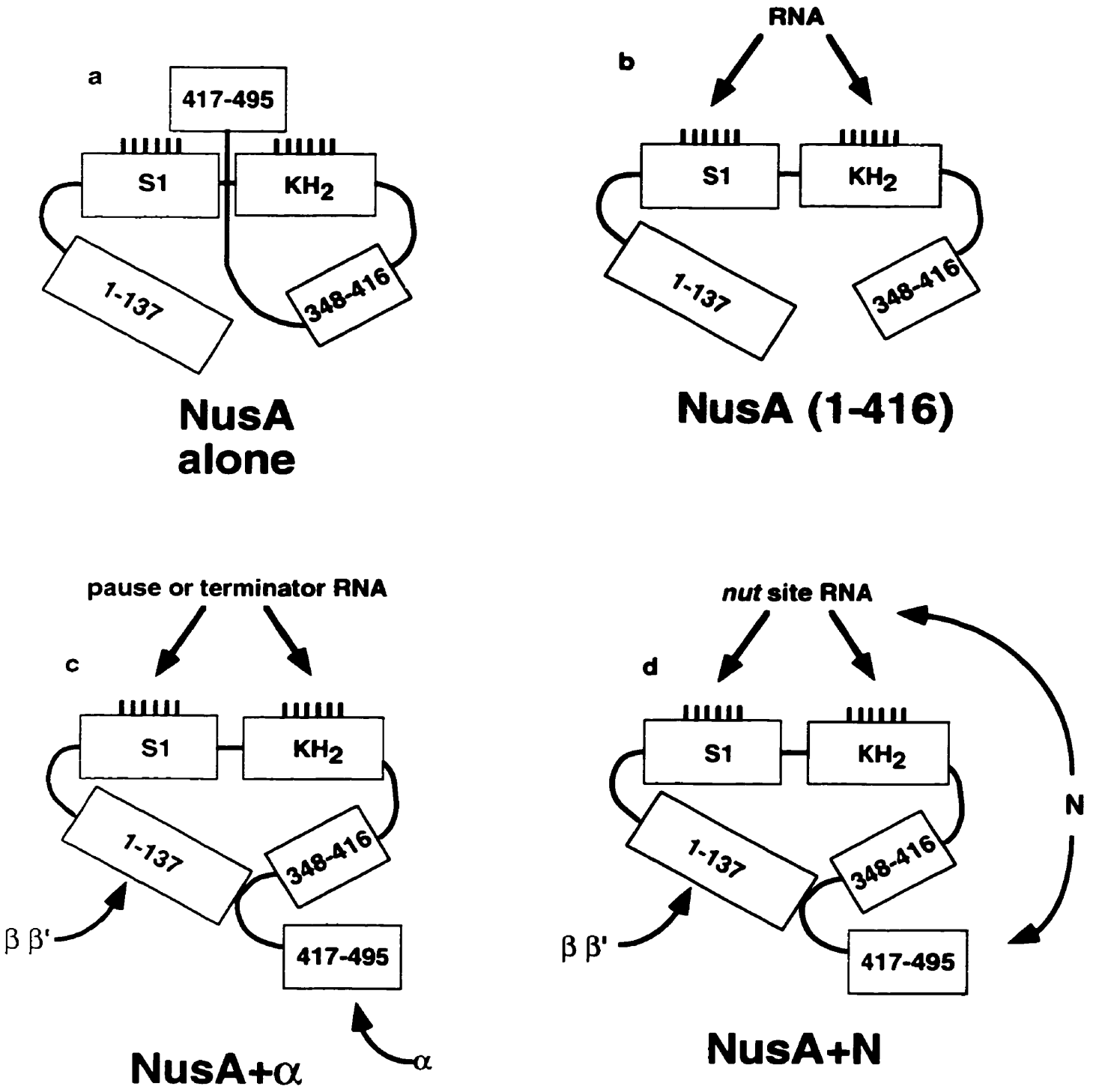


Figure 7: Model for NusA function in elongation, termination and antitermination. See text for details.

al., 1992; Dombroski *et al.*, 1993; Malhotra *et al.*, 1996). It is intriguing that two prokaryotic proteins that compete for binding to the RNA polymerase core enzyme (Greenblatt and Li, 1981a) may use similar mechanisms to control their nucleic acid binding ability.

Interaction of NusA with α : Relationship to RNA-binding and Transcription Termination

My observation that the α subunit of RNA polymerase stimulates RNA-binding by NusA is consistent with previous studies showing that NusA can be cross-linked to the nascent transcript in transcription complexes (Liu and Hanna, 1995). I propose that this effect of α on RNA-binding by NusA is the consequence of a direct interaction between the two proteins (see Fig. 7b). The binding of NusA to α and stimulation of RNA-binding by α are mediated by the carboxy-terminal domain of α (Figs. 2 and 4c). The α-CTD was implicated previously in the ability of NusA to stimulate pausing and termination by RNA polymerase (Liu *et al.*, 1996). RNA-binding either by truncated NusA or by full length NusA in the presence of α is weak. Nevertheless, the dependence of RNA-binding by truncated NusA or by full length NusA in the presence of α on the *boxA* sequence in the *nut* site is similar to the effect of *boxA* on RNA-binding by NusA in the presence of N (Mogridge *et al.*, 1995). My ability to abrogate RNA-binding by deleting the α-CTD or portions of NusA also suggests that RNA-binding genuinely involves both α and NusA.

From the data presented here, I hypothesize that full-length NusA is prevented from interacting with RNA by an autoinhibition mediated by its carboxy-terminal 79 amino acids (compare Fig. 7a with Fig. 7b). I suggest that during elongation, NusA uses its RNA polymerase-binding region in amino acids 1-137 (Chapter 2) to interact with RNA polymerase subunits β and β', and its carboxy-terminal region to interact with α (Fig. 7c). The interaction with α may then cause a conformational change in NusA such that its RNA-binding domains either fold or become exposed and competent to bind the nascent RNA.

Thus, as part of the transcription complex, NusA would be in a position to bind and stabilize pause and termination motifs in the nascent RNA, leading to enhancement of pausing and termination at certain sites. I suggest that the interaction of the α-CTD with NusA is essential for NusA to stimulate termination only if the inhibitory carboxy-terminal region of NusA is present and not if it is deleted. In other experiments I have shown that NusA (1-416) is able to stimulate termination (Chapter 2). This fragment of NusA cannot interact with α, but it lacks the inhibitory region and can bind RNA on its own.

The interaction between the amino-terminal RNA polymerase-binding region of NusA (amino acids 1-137) and RNA polymerase is clearly essential for function because the loss of the amino-terminal RNA polymerase-binding region results in the inability of NusA to participate in termination and antitermination (Chapter 2). It is likely that one or both of the two large RNA polymerase subunits contact the amino-terminal RNA polymerase-binding region of NusA. I found that α alone does not bind the amino-terminal region of NusA. Also crosslinking data (J. Li and J. Greenblatt, unpublished data) and other evidence (Liu *et al.*, 1996) suggest that β and β' interact directly with NusA. Nevertheless, a weak but important interaction between α and this region of NusA cannot be ruled out, because the interaction between α and full length NusA may be stronger than the interaction between α and the carboxy-terminal region of NusA.

Transcriptional Antitermination by N

I also showed that N interacts with this same carboxy-terminal autoinhibitory region of NusA. Therefore, I propose that N activates the RNA-binding activity of NusA in a manner similar to that of α, as modelled in Fig. 7d. In this scheme, the close proximity of NusA to both the *nut* site and N could result in exclusive binding of NusA to the *nut* site RNA. Such an interaction might then serve two purposes. First, the interaction of NusA

with *nut* site RNA would prevent pause and termination sequences in the nascent RNA from binding to NusA. This would prevent NusA from enhancing pausing and termination. Second, the NusA-*nut* site interaction, together with interactions involving other Nus factors, N, RNA polymerase, and the *nut* site RNA would increase the overall stability of the antitermination complex containing N.

Friedman and colleagues have shown that a point mutation in the α-CTD or deletion of the entire α-CTD enhances antitermination *in vivo* (Schauer *et al.*, 1996). Furthermore, even though loss of the α-CTD prevents NusA from stimulating termination *in vitro*, it has no effect on the ability of NusA to stimulate antitermination mediated by the N protein *in vitro* (Liu *et al.*, 1996). This data suggests that there may be a competition between α and N for binding to the carboxy-terminal region of NusA. Interaction of α with NusA may direct NusA to the nascent transcript near its 3' end and facilitate pausing and termination, whereas interaction of N with NusA would block the interaction between NusA and α, direct NusA to the *nut* site RNA, and prevent NusA from stimulating pausing and termination.

Materials and methods

Plasmids, strains and reagents

RNA polymerase, NusA, GST-NusA proteins, GST-N, His_6-tagged NusA proteins, α and α_{235} were purified as previously described (Chapter 2; Burgess and Jendrisak, 1975; Greenblatt and Li, 1981b; Mogridge *et al.*, 1998a; Mogridge *et al.*, 1998b). The plasmids His α wt and His α_{235} for α and α_{235} (α-NTD), respectively, were kindly provided by K. Severinov. Purified α_{CTD} was provided by G. Zhang and S. Darst. NusA R199A was provided by Ying Zhou and David Friedman. Purification for this protein will be described elsewhere (T. Mah, Y. Zhou, N. Yu, J. Mogridge, E. Olsen, J. Greenblatt and D. Friedman, unpublished data).

The oligonucleotides used for cloning were purchased from ACGT Corp. (Toronto). RNAguard was bought from Pharmacia Biotech. Restriction enzymes and DNA ligase were purchased from New England Biolabs. T7 RNA polymerase was obtained from Life Technologies.

Construction of GST-NusA Proteins

PCR primers were designed to amplify fragments of *NusA* from the plasmid J1150. Forward and reverse primers contained BamHI and EcoRI restriction sites, respectively, for subsequent cloning into the vector pGEX-2T (Pharmacia).

Purification of GST-NusA Proteins

The *E.coli* strain DH5α (Life Technologies) containing the GST-NusA fusion plasmids was grown in 1l of LB medium to an A_{600} of 0.5 and induced for 3 hours with 0.5 M isopropyl-1-thio-β-D-galactopyranoside. Cells were harvested by centrifugation, resuspended, and sonicated in 10 ml of 1 M NaCl Buffer A (20 mM Tris-HCl, pH 7.8, 0.2

mM EDTA, 1 mM dithiothreitol, 1 mM phenylmethylsulfonyl fluoride) and then centrifuged for 20 minutes at 12,000 rpm in a Sorval SS34 rotor. Glutathione-sepharose 4B beads (1ml: Pharmacia) were added to the supernatant and this slurry was rotated for 1 hour at 4°C. The beads were washed and resuspended in 500 µl thrombin cleavage buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 2.5 mM $CaCl_2$). Thrombin was added to the slurry and incubated at room temperature for 1 hour. After centrifugation at 3000 rpm in a tabletop eppendorf microfuge, the supernatant was collected and dialyzed into 0.1 M NaCl ACB (10 mM HEPES pH 7.0, 10% glycerol, 0.1 mM EDTA, 1 mM DTT).

Affinity Chromatography

Experiments with GST, GST-N, and GST-NusA (303-495) were performed as previously described (Chapter 2 of this thesis) except that the GST and GST-N columns were either loaded with *E.coli* extract containing NusA or NusA fragments cleaved from GST with thrombin and the GST and GST-NusA (303-495) columns were loaded with a mixture of α, α_{NTD}, and α_{CTD}, buffer and 0.2 mg/ml insulin.

Experiments with the α subunit of RNA polymerase were done two ways. In the first case, α was coupled to affigel 10 matrix at three different concentrations. 20 ml of beads were added to 200 ml pipette tips that contained 10 µl of 212-300 micron glass beads (Sigma). The column bed was washed with 10 column volumes of 1 M NaCl ACB (10 mM HEPES [pH 7.0], 10% glycerol, 0.1 mM EDTA, 1 mM DTT) and then washed with 10 column volumes of 100 mM NaCl ACB. Columns were loaded with a mixture of his-tagged NusA deletion proteins as well as full length his-tagged NusA, buffer and 0.2 mg/ml insulin. In the second case, his-tagged NusA (1-495), GST-NusA (1-137) and GST-NusA (303-495) were coupled to affigel 10 at equimolar concentrations. 20 ml of beads were added to 200 ml

pipette tips. The columns were treated as described above except that these columns were loaded with buffer containing α and 0.2 mg/ml insulin.

Gel Mobility Shift Experiments

Gel mobility shift assays were performed as previously described (Mogridge *et al.*, 1995) except that the proteins were added together and incubated on ice for 20 minutes. Radiolabelled probe was then added and incubation on ice continued for an additional 10 minutes. Concentrations of proteins are indicated in the figure legends.

Chapter 4

Interaction between the *E.coli* NusA elongation factor and the λ *nut* site RNA

Abstract

The *nut* site (*boxA*, *interbox* and *boxB*) is a cis-acting RNA element necessary *in vivo* for antitermination by the λ N protein. N binds directly to *boxB* whereas *boxA* is the postulated interaction site for the *E.coli* co-factors for antitermination NusA, NusB and S10. By using a series of clustered substitution mutations in *boxA* and the *interbox* in gel mobility shift assays, I have found that the *interbox* and the 3' end of *boxA* are important for the ability of NusA to associate with an N-*nut* site complex. Moreover, the sequence requirements are similar for binding of wild type NusA to an N-*nut* site complex and for the binding to *nut* site RNA of NusA (1-416), a carboxy-terminally truncated form of NusA that can bind RNA in the absence of N. As well, the effects of point mutations in *boxB* on RNA-binding by NusA (1-416) in the absence of N suggested that the formation of a GNRA tetraloop-like structure in *boxB* RNA is important for the binding of NusA to *nut* site RNA. The inability of NusA with a mutation at R199 in its S1 homology region to bind to an N-*nut* site complex indicated that the S1 homology region of NusA is important for the association of NusA with RNA. In contrast, the 944 mutation in the S1 homology region allowed NusA to bind the N-*nut* site complex, but blocked the subsequent association of the resulting complex with RNA polymerase, NusB, NusG, and S10. The ability of NusA 944 to inhibit the formation of a complete complex containing NusA R199A indicated that NusA normally binds to the N-*nut* site complex prior to the assembly of the other factors.

Introduction

Transcriptional antitermination mediated by the N protein of bacteriophage λ requires several *E.coli* host factors and a cis-acting element known as the *nut* site. The host factors are called Nus, for N-utilization substance, and include NusA, NusB, NusG and S10 (reviewed in Friedman, 1988; Greenblatt *et al.*, 1993). These factors, along with the N protein and the *nut* site RNA, form a stable complex that modifies RNA polymerase into a termination-resistant form that persists over several kilobases of λ sequence and through many terminators in a process known as "processive antitermination" (Das and Wolska, 1984; Mason *et al.*, 1992b; DeVito and Das, 1994; Mogridge *et al.*, 1995). Many protein-protein and protein-RNA interactions have been identified that are likely to stabilize this antitermination complex.

There is indirect evidence to suggest that NusA interacts with *boxA* or the *interbox* sequence that separates *boxA* from *boxB*. Genetic observations that mutations in nucleotides 8 and 9 in *boxA* suppressed the effect of the *nusA1* mutation on antitermination by N suggested that NusA may interact with *boxA* (Friedman *et al.*, 1990), whereas gel mobility shift experiments showed that inversion of *boxA* and the *interbox*, in the context of a wild type *boxB*, prevents the association of NusA with an N-*nut* site complex (Mogridge *et al.*, 1995). Furthermore, NusA, as part of the transcription complex, could be cross-linked to nascent RNA greater than 10 nucleotides in length (Liu and Hanna, 1995). Consistent with this evidence suggesting that NusA interacts with RNA, sequence comparisons have shown that NusA contains one S1 and two KH homology regions, which are putative RNA-binding domains (Chapter 3; Gibson *et al.*, 1993; Bycroft *et al.*, 1997).

There is no direct evidence to show that either the S1 or KH homology regions in NusA are involved in RNA-binding. However, mutations in the S1 homology region of NusA have been isolated and their effects on *E.coli* growth and antitermination by N have been assessed. The *nusA1* mutation in this region prevents antitermination by N at 42°C (Friedman, 1971) and prevents NusA from binding to an N-*nut* site complex (Mogridge *et al.*, 1995). *nusA R199A* has a mutation of a conserved arginine to alanine at position 199 in the S1 homology region and *nusA R199K* has a mutation at the same position, but in this case, the arginine was replaced by lysine (Y. Zhou and D. Friedman, unpublished results). Substitution of an alanine or lysine for arginine at this position has minimal effect on the viability of *E.coli* and antitermination by N. Conversely, a strain containing *nusA 944*, a hybrid between $nusA^{Ec}$ and $nusA^{Sc}$ in which 4 amino acids from *E.coli* NusA are replaced by 9 amino acids from *S.typhimurium* NusA in the only major region of heterogeneity between the two proteins, is defective in supporting antitermination by N. The 944 substitution is also located in the S1 homology region. Interestingly, *nusA 944* is dominant to *nusA R199A* in the sense that it prevents antitermination by N in the presence of *nusA R199A*, which would otherwise allow antitermination (Y. Zhou and D. Friedman, unpublished data).

Gel mobility shift experiments with a collection of point mutations in each nucleotide in *rrn boxA*, a sequence that directs antitermination in *E.coli rrn* operons and differs from λ *boxA* at positions 1, 8 and 9, mapped the binding site for the NusB/S10 heterodimer to the *boxA* sequence UGCUCUUUACACA (Nodwell and Greenblatt, 1993). A similar mutational study showed that the sequence GCUCUU near the 5' end of the *boxA* sequence CGCUCUUACACA in the λ *nutR* site is important for the binding of NusB, S10, and NusG to a complex containing N, NusA, RNA polymerase and the *nut* site RNA (Mogridge *et al.*,

1998b). Ribosomal protein S1, a protein consisting of six S1 domains, also binds specifically to *rrn boxA* RNA (Mogridge and Greenblatt, 1998). Whereas S1 can compete with NusB/S10 for binding to *boxA*, the nucleotides most important for the binding of S1 to *boxA* RNA are located mostly toward the 3' end of *boxA*, at positions 3, 8, 9, 10, and 12 (Mogridge and Greenblatt, 1998).

Given the genetic observations that had connected NusA with *boxA*, my goal was to use gel mobility shift assays to test whether there was likely to be a direct interaction between NusA and *boxA*. In previous studies, I used a series of three- and five-nucleotide substitution mutations in *boxA* and the *interbox* to identify a region of the *nut* site that was important for the association of NusA with an N-*nut* site complex. These results revealed that the 3' end of *boxA* and the *interbox* region were most important for the binding of NusA (T. Mah, Masters Thesis, 1995). In the present study, I have used these same substitution mutations to characterize the binding to *nut* site RNA of a carboxy-terminally truncated form of NusA, NusA (1-416), that can bind *nut* site RNA on its own (i.e. in the absence of N). I have also used NusA proteins with mutations at R199 in the S1 homology region of NusA to provide evidence that this predicted RNA-binding region is important for the interaction between NusA and RNA. Moreover, experiments with the 944 mutation in the S1 homology region have enabled me to conclude that NusA normally associates with the N-*nut* site complex prior to the association of this complex with RNA polymerase and the other Nus factors.

Results

Inversion of the *boxA* and *interbox* nucleotides in the *nut* site prevents the association of NusA with an N–*nut* site complex in a gel mobility shift assay (Mogridge *et al.*, 1995). In order to further investigate the relationship between this portion of the *nut* site and NusA, I made a series of three nucleotide substitution mutations along *boxA* and a five nucleotide substitution mutation that altered the entire *interbox* region. I then used these mutated *nut* site constructs as RNA probes in gel mobility shift assays. The results from these experiments were published in my Masters thesis and a summary of them is shown in Table 1. My results suggested that the nucleotides important for the association of NusA with an N-*nut* site complex are located in the region of *boxA* starting at nucleotide 7 and extending into the *interbox* region.

Full length NusA does not associate with *nut* site RNA on its own (Mogridge *et al.*, 1995). Rather, it requires interaction with N or the α subunit of RNA polymerase in order to activate its RNA-binding ability (see Chapter 3). However, a carboxy-terminally truncated form of NusA, NusA (1-416), is capable of independent association with *nut* site RNA and this interaction is severely compromised by inversion of the *boxA* and *interbox* sequences (see Chapter 3). To determine whether the nucleotides that are important for the association of NusA with an N-*nut* site complex are similar to those that are important for NusA (1-416) to bind the *nut* site on its own, I performed gel mobility shift experiments with NusA (1-416) using the three nucleotide substitution mutants of *boxA* and the five nucleotide substitution mutant of the *interbox* as RNA probes (Figs. 1 and 2). The binding of NusA (1-416) to the *nut* site RNA was much weaker than the binding of intact NusA to an N-*nut* site complex.

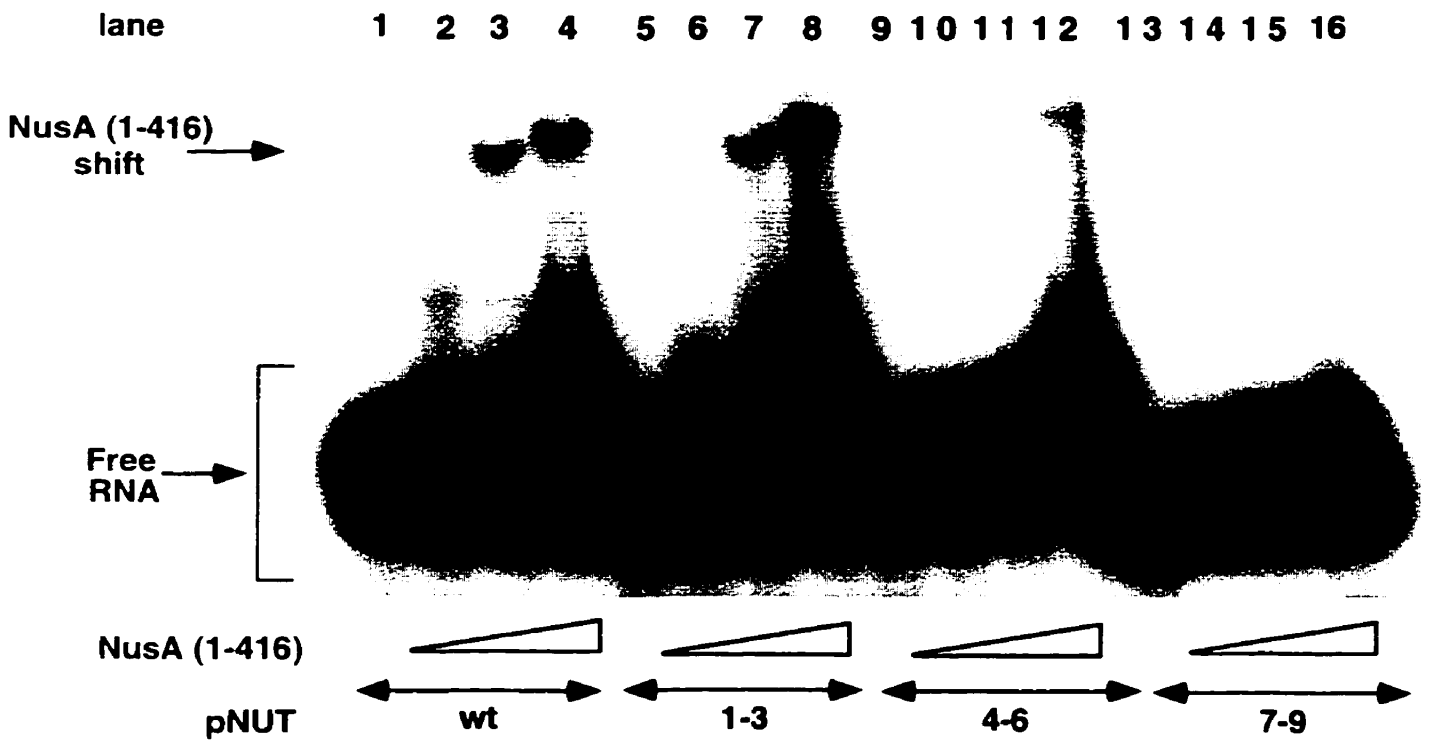


Figure 1. Substitution of nucleotides 1 to 3 in *boxA* does not affect the ability of NusA (1-416) to interact with *nut* site RNA, whereas substitution of nucleotides 4 to 6 or nucleotides 7 to 9 substantially reduces the interaction.
Reactions containing 0.8, 1.6, or 3.2 μM NusA (1-416) and ^{32}P-labeled wild type or mutant *nut* site RNAs (as indicated) were electrophoeresed on non-denaturing gels, dried, and exposed to film.
Positions of free and complexed RNAs are indicated.

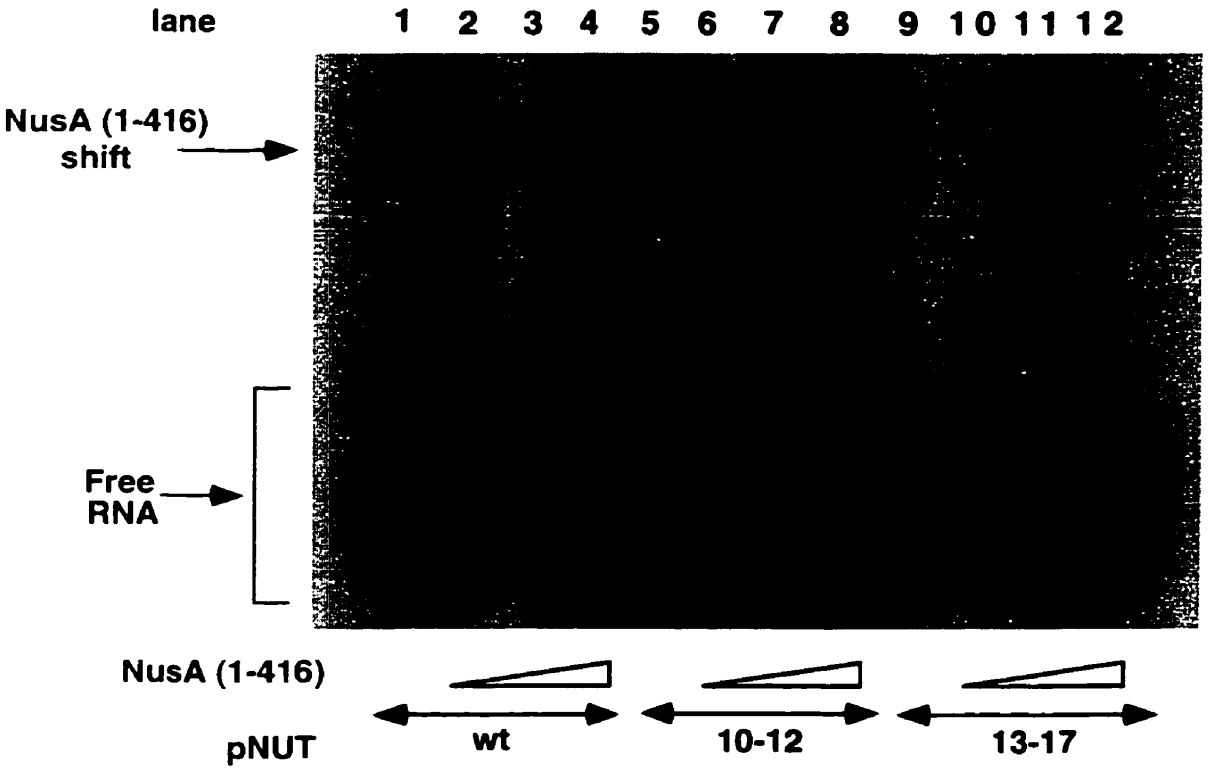


Figure 2. Substitution of nucleotides 10 to 12 of *boxA* and 13 to 17 of the *interbox* both severely affect the ability of NusA (1-416) to interact with *nut* site RNA

Reactions containing 0.8, 1.6, or 3.2 μM NusA (1-416) and ^{32}P-labeled wild type or mutant *nut* site RNAs (as indicated) were electrophoresed on non-denaturing gels, dried, and exposed to film.

Positions of free and complexed RNAs are indicated.

This result suggests that the presence of N may stabilize the interaction between NusA and RNA. Nevertheless, the nucleotides in *boxA* and the *interbox* that were important for the binding of NusA (1-416) to *nut* site RNA were similar to those that were important for the binding of NusA to an N-*nut* site complex. Inversion of *boxA* prevented association (see Chapter 3), mutation of the 5' end of *boxA* had either no effect or minimal effect, and mutations of the 3' end of *boxA* from nucleotides 7 to 12 and the *interbox* had more profound effects on the binding of NusA (1-416) to the RNA probe (Table 1). I conclude that the interaction of NusA with an N-*nut* site complex is likely to involve a direct interaction of NusA with nucleotides in the *boxA* and *interbox* regions.

Since mutations at position 2 (*boxB* A24C) and 4 (*boxB* A26C) in the loop of *boxB* prevent the association of NusA with an N-*nut* site complex (Mogridge *et al.*, 1995), I wanted to determine whether the same nucleotides were important for the binding of NusA (1-416) to the *nut* site RNA (Figs. 3 and 4; Table 1). Surprisingly, the binding of NusA (1-416) to both the A24C and A26C *boxB* mutants was only marginally reduced. Nucleotides G23 and A27 form a sheared G-A base pair that is essential for forming the GNRA tetraloop-like structure in *boxB* (Cai *et al.*, 1998; Legault *et al.*, 1998). Mutations that affected the ability of *boxB* to form the GNRA tetraloop-like structure, namely *boxB* G23A and *boxB* A27C, greatly inhibited the ability of NusA (1-416) to bind the RNA. Thus, the binding of NusA (1-416) to *nut* site RNA requires specific nucleotides at the 3' end of *boxA* and the *interbox*, and also requires the formation of a GNRA tetraloop-like structure in *boxB*.

To analyze the involvement of the S1 homology region in this interaction between NusA and *nut* site RNA, I made use of three NusA mutants, NusA R199A, NusA R199K and NusA 944, that were altered in the S1 domain. These proteins were used in gel mobility shift

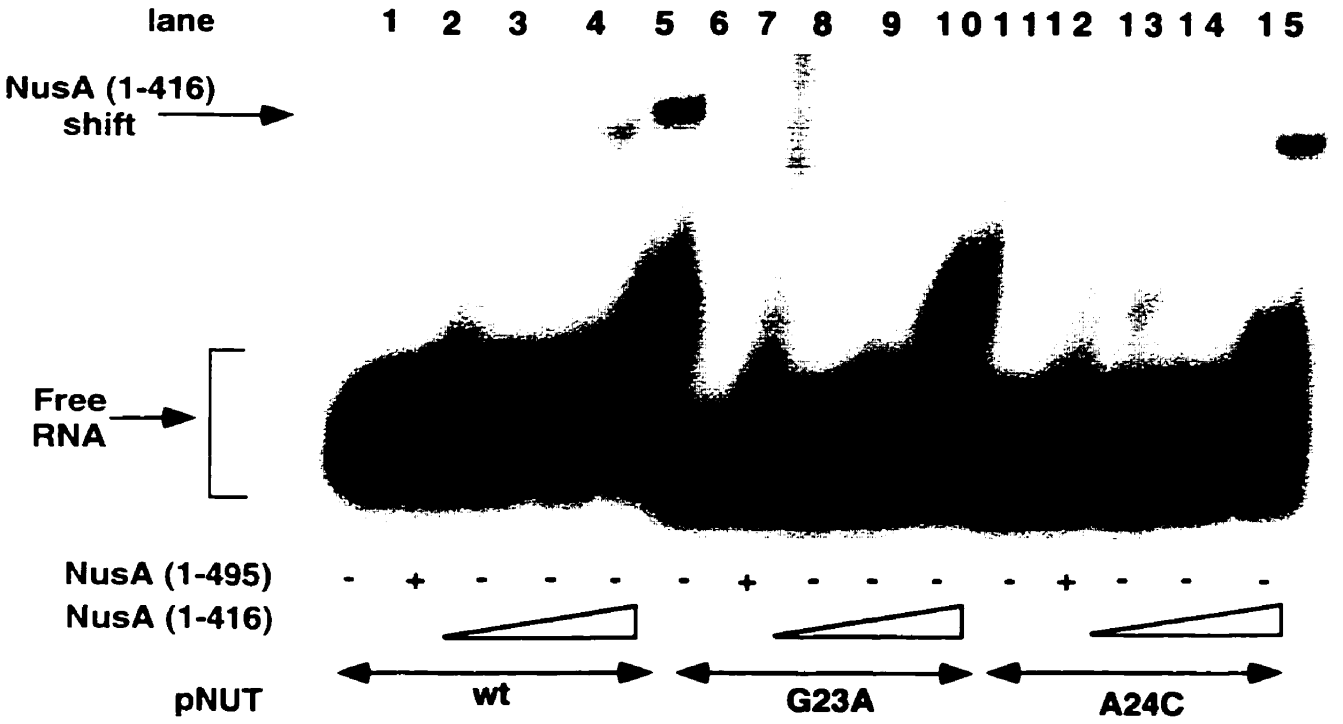


Figure 3. Substitution of nucleotide 23 in the loop of *boxB* reduces the ability of NusA (1-416) to interact with *nut* site RNA whereas substitution of nucleotide 24 has little effect.
Reactions containing 3.2 μM NusA and 0.8, 1.6, or 3.2 μM NusA (1-416) and ^{32}P-labeled wild type or mutant *nut* site RNAs (as indicated) were electrophoresed on non-denaturing gels, dried, and exposed to film.
Positions of free and complexed RNAs are indicated.

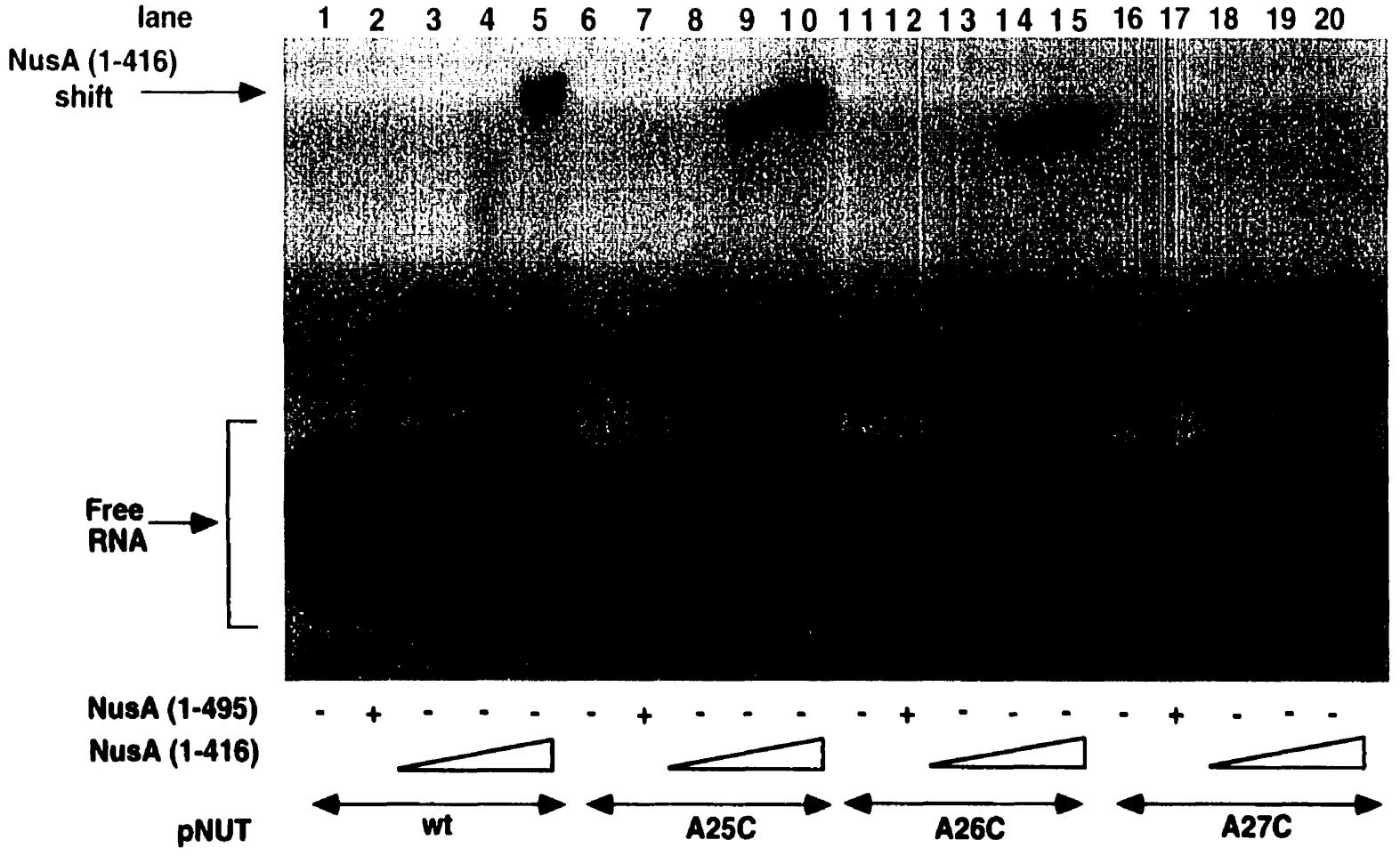


Figure 4. Substitution of nucleotide 27 in the loop of *boxB* severely reduces the ability of NusA (1-416) to interact with *nut* site RNA whereas substitution of nucleotides 25 and 26 has little effect. Reactions containing 3.2 μM NusA or 0.8, 1.6, or 3.2 μM NusA (1-416) and ^{32}P-labeled wild type or mutant *nut* site RNAs (as indicated) were electrophoresed on non-denaturing gels, dried, and exposed to film. Positions of free and complexed RNAs are indicated.

nut site variant	N-NusA-*nut* site	NusA (1-416)-*nut* site
wild type	+ [a]	+
reverse *boxA*	- [a]	-
boxA 1-3	+ [b]	+
boxA 4-6	+ [b]	+
boxA 7-9	- [b]	-
boxA 10-12	- [b]	-
boxA 13-17	- [b]	-
boxB G23C	ND	-
boxB A24C	- [a]	+
boxB A25C	ND	+
boxB A26C	- [a]	+
boxB A27C	ND	-

Table 1. Comparison of the abilities of NusA (1-416), in the absence of N, and wild type NusA, in the presence of N, bind to wild type and mutant nut sites.
A plus sign indicates that the *nut* site mutation has less than a two-fold effect on the formation of either the N-NusA-*nut* site complex or the NusA (1-416)-*nut* site complex.
A negative sign indicates that there is a three- to five-fold reduction in the formation of the complex.
ND indicates that an experiment with that particular mutant *nut* site was not done.
[a] Mogridge *et al.* 1995
[b] T. Mah, Masters Thesis, 1995

assays with N, RNA polymerase, and the rest of the Nus factors to assess the impact of these mutations on the formation of various complexes containing these factors (Figs. 5-8). The addition of N protein to wild type *nut* site RNA resulted in the formation of an N-*nut* site complex (Fig. 5a, lane 2; Fig. 5b, lane 2; T. Mah, Masters Thesis, 1995; Mogridge *et al.*, 1995), while the addition of increasing amounts of NusA to this complex further retarded the mobility of the RNA, resulting in the formation of an N-NusA-*nut* site complex (Fig. 5a, lanes 3-5; Fig. 5b, lane 3; T. Mah, Masters Thesis, 1995; Mogridge *et al.*, 1995). However, NusA R199A and NusA R199K were unable to cause this supershift of the N-*nut* site complex (Fig. 5a, lanes 6-11).

A consensus *nut* site which has alterations in nucleotides 8 and 9 that makes λ *boxA* resemble *rrn boxA* was able to suppress the effect of the *nusA1* mutation on antitermination by N (Friedman *et al.*, 1990). It was suggested that the alteration of nucleotides 8 and 9 in *boxA* would improve the interaction between NusA and *boxA* (Friedman *et al.*, 1990). However, use of the consensus *nut* site RNA as a probe with the R199 mutant NusA proteins did not improve the ability of these proteins to associate with an N-*nut* site complex (Fig. 5a, lanes 16-19). Although the NusA mutants altered at position 199 are unable to bind an N-*nut* site complex, NusA 944 was able to supershift an N-*nut* site complex (Fig. 5b, lane 4). These experiments suggest that the conserved arginine at position 199 is important for the ability of NusA to interact with an N-*nut* site complex but that not all mutations in the S1 homology region result in this type of defect.

Mutation of the arginine at position 199 of NusA did not alter *E.coli* viability nor did it affect antitermination by N. However, substitution of nine *S.typhimurium* amino acids for four *E.coli* amino acids in the 4-for-9 region resulted in a NusA 944 protein defective in

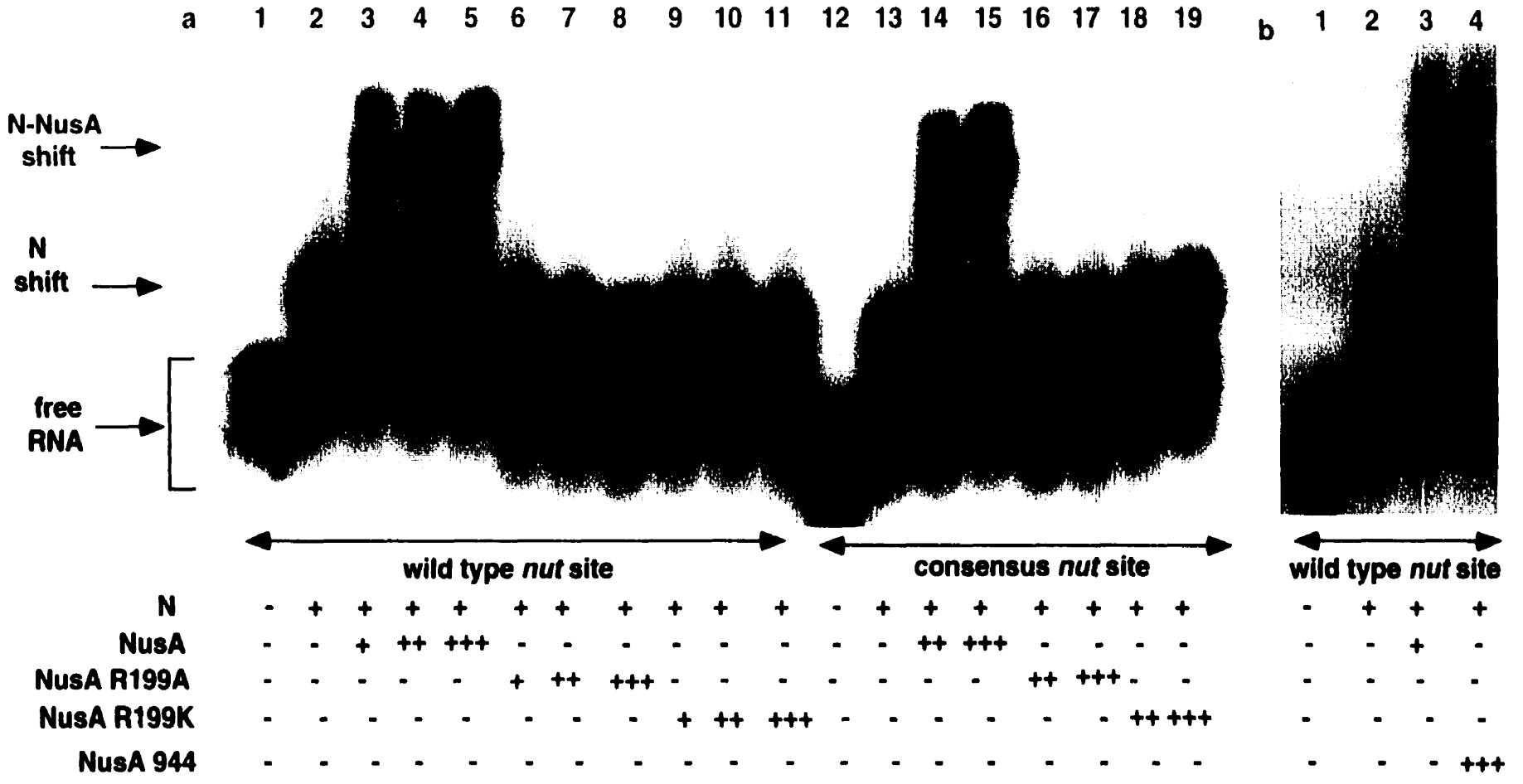
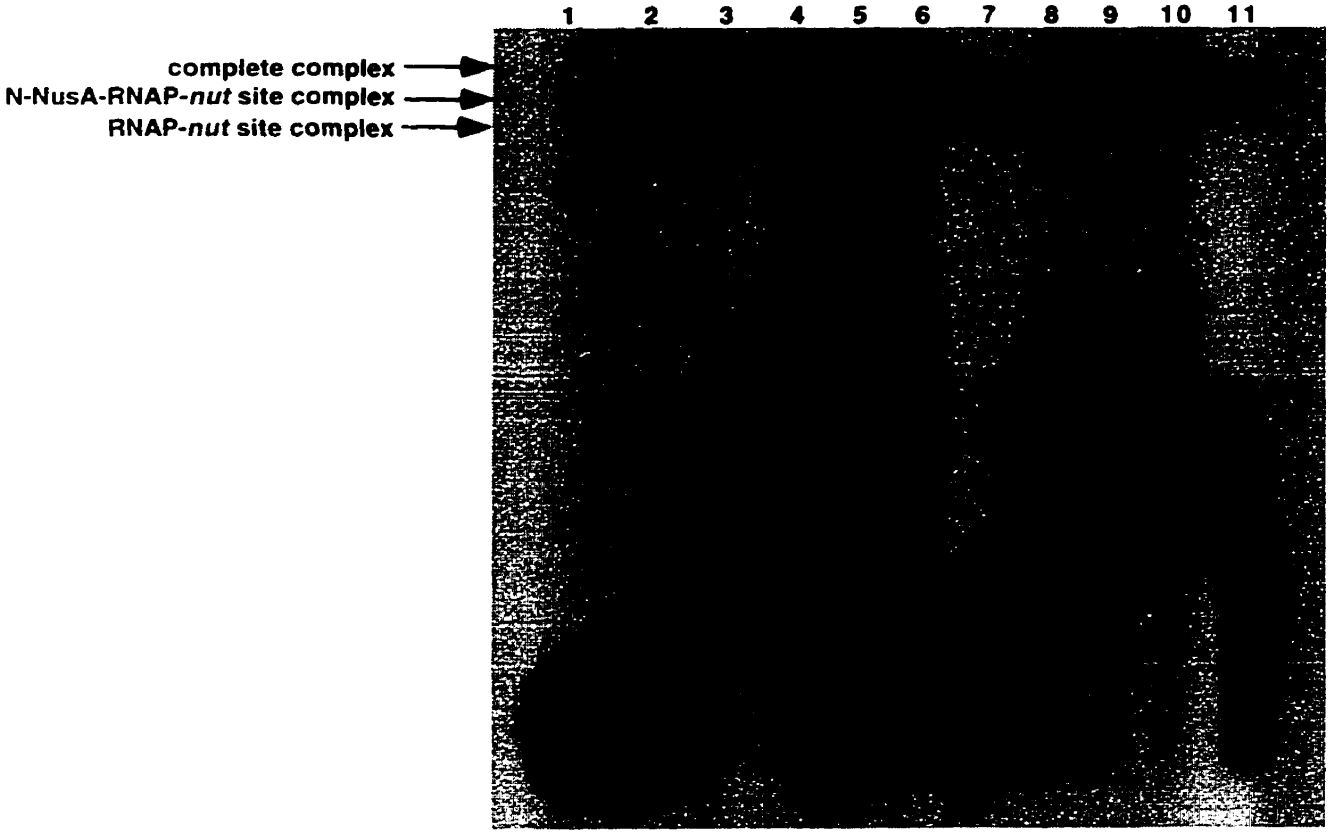


Figure 5. Mutations in the S1 homology region of NusA have differential effects on the formation of N-NusA-*nut* site complexes. (a) Mutation of position 199 in NusA prevents the formation of an N-NusA-*nut* site complex. (b) Alteration of the 944 region does not affect the formation of an N-NusA-*nut* site complex. Reactions containing ^{32}P-labelled wild type or consensus *nut* site RNAs and various combinations of 1500 nM N, and either 400 nM (+), 600 nM (++), or 800 nM (+++) NusA or NusA mutants (as indicated) were electrophoresed, dried and exposed to film.



RNA polymerase	-	+	+	+	+	+	+	+	+	+	+
NusA wt	-	-	+	+	-	-	-	-	+	+	-
NusA R199A	-	-	-	-	+	+	-	-	+	-	+
NusA 944	-	-	-	-	-	-	+	+	-	+	+
N	-	-	+	+	+	+	+	+	+	+	+
B/G /S	-	-	-	+	-	+	-	+	+	+	+

Figure 6. Mutations in the S1 homology region of NusA have differential effects on the formation of a complete complex containing RNA polymerase, N, NusA, NusB, NusG and S10. Reactions containing ^{32}P-labelled *nut* site RNA and various combinations of 500 nM N, 50 nM NusB, 50 nM NusG, 50 nM S10, 25 nM RNA polymerase and 100 nM NusA or NusA mutants (as indicated) were electrophoresed on 5% non-denaturing gels, dried and exposed to film.

supporting antitermination by N (Y. Zhou and D. Friedman, unpublished results). This is precisely the opposite of what might have been predicted on the basis of the effects of these mutations on the binding of NusA to N-*nut* site complexes (Fig. 5). To test whether these mutant proteins were able to assemble into complete complexes containing RNA polymerase, I performed gel mobility shift experiments using RNA polymerase, N, NusA, NusB, NusG and S10 (Fig. 6; Chapter 2; Mogridge *et al.*, 1995). RNA polymerase bound to RNA on its own and the addition of N and NusA resulted in the formation of an N-NusA-RNA polymerase-*nut* site complex (Fig. 6, lanes 2 and 3). Addition of NusB, NusG and S10, as well as N, NusA, and RNA polymerase, resulted in the formation of a complete complex of low mobility (Fig. 6, lane 4). Despite the fact that NusA R199A could not interact with an N-*nut* site complex (Fig. 5a) and only allowed formation of a small amount of an N-NusA-RNA polymerase-*nut* site complex (Fig. 6, lane 5), it could efficiently support the formation of a complete complex containing all of the added factors (Fig. 6, lane 6). This result was consistent with the observation that NusA R199A supports antitermination by N. However, NusA 944 was unable to support formation of the complete complex (Fig. 6, lane 8), even though it could associate with an N-*nut* site complex (Fig. 5b, lane 4), thus explaining why *nusA 944* does not support antitermination by N. Furthermore, in a reaction in which an equal amount of NusA R199A was mixed with NusA 944, NusA 944 was able to prevent NusA R199A from assembling into a complete complex (Fig. 6, lane 11). In contrast, NusA 944 allowed formation of a complete complex when wild type NusA was present (Fig. 6, lane 10). These results probably explain why the *nusA 944* mutation is dominant to *nusA R199A*, but recessive to wild type *nusA*.

Discussion

I have shown that mutation of nucleotides in the *interbox* or the 3' end of *boxA* reduce the association of NusA (1-416) with *nut* site RNA in the absence of N or the association of wild type NusA with an N-*nut* site complex. These results strongly suggest that NusA interacts directly with this region of the *nut* site RNA. A similar conclusion was reached on the basis of the observations that mutation in nucleotide 8 of *boxA* allows for the functioning of the *S.typhimurium* NusA with λ (Friedman and Olson, 1983) and that mutations in nucleotides 8 and 9 suppress the effect of the *nusA1* mutation on antitermination by N (Friedman *et al.*, 1990). However, I found that changes in *boxA* nucleotides 8 and 9 do not affect the binding of NusA (Fig. 5a) and Nodwell (Nodwell and Greenblatt, 1993) found that these mutations stabilize the binding to *rrn boxA* RNA of NusB and S10. Thus, NusA does interact with *boxA* and *interbox* RNA, but the interaction does not involve *boxA* nucleotides 8 and 9. Changes here most likely indirectly suppress the *nusA1* mutation by enhancing the binding of NusB and S10. Other studies have suggested that *boxA* is not required for antitermination over a short distance, but rather that it is part of a regulatory system that binds an inhibitor of antitermination (Zuber *et al.*, 1987; Patterson *et al.*, 1994). Nevertheless, it is likely that this region in *boxA* and the *interbox* defines a low affinity binding site for NusA, and that there is a second low affinity site in *boxB*.

The nucleotides in λ *boxA* and the *interbox* that are important for NusA interaction are similar but not identical to those that are important for S1 interaction with *rrn boxA* (Mogridge and Greenblatt, 1998). Furthermore, S1 binds to λ *boxA* (Mogridge and Greenblatt, 1998). Since S1 consists of six S1 domains, it is possible that the S1 domains in

both S1 and NusA recognize similar RNA sequences. Based on these observations, it would be reasonable to suggest that the S1 homology region in NusA binds *boxA*, whereas the KH homology regions in NusA recognize nucleotides in *boxB*.

Full length NusA does not bind *nut* site RNA on its own in gel mobility shift experiments (Mogridge *et al.*, 1995). The binding of NusA (1-416) to the wild type *nut* site RNA indicates that NusA is capable of interaction with RNA and suggests that removal of the carboxy-terminal region of NusA relieves inhibition of RNA-binding. Experiments in Chapter 3 of this thesis show that the binding of N or α to the carboxy-terminal region of NusA allows interaction of NusA with RNA. Together, my results strongly support the suggestion that the carboxy-terminal region of NusA inhibits RNA-binding.

The *boxA* and *interbox* nucleotides that are important for the association of NusA with an N-*nut* site complex are also important for the association of NusA (1-416) with *nut* site RNA in the absence of N. However, nucleotides in the loop of *boxB* differ in their importance for the formation of these complexes. These results suggest that these two different forms of NusA bind *boxB* RNA differently. It is possible that the RNA-binding regions of NusA that are exposed in NusA (1-416) are not in the same configuration as they are when NusA binds the *nut* site in an N-*nut* site complex. The binding of NusA (1-416) to the *nut* site RNA required the formation of the GNRA tetraloop-like structure, but the identities of the three internal loop nucleotides were not important. In contrast, wild type NusA could not associate with an N-*nut* site complex if nucleotide 2 or 4 of the loop was mutated (Mogridge *et al.*, 1995). It is likely that *boxB* RNA can exist in alternative conformations since it forms different GNRA tetraloop-like structures depending on whether the N protein from λ or P22 is bound (Cai *et al.*, 1998; Greenblatt *et al.*, 1998; Legault *et al.*,

1998). Thus, it is possible that N traps *boxB* in a conformation recognizable by NusA that can only be formed when the λ N protein is present. If this is correct, full length NusA would interact with the extruded loop nucleotide 4 in the N-*boxB* complex, whereas this nucleotide would have a different conformation in the absence of N.

My experiments with NusA variants in the S1 homology region suggested that at least the conserved arginine at position 199 of the protein is important for RNA-binding. This mutation has only a mild effect on antitermination by N *in vivo*, yet prevents NusA from interacting with an N-*nut* site complex. This apparent contradiction was resolved by an experiment in which RNA polymerase and the other Nus factors were added to the reaction. NusA R199A did allow the formation of a complete complex containing N, RNA polymerase, and the other Nus factors. Thus, other protein-protein and protein-RNA interactions within the N-modified antitermination complex can compensate for a weakened interaction of NusA with the *nut* site RNA.

These studies also provided an indication of the normal order of assembly of the antitermination complex. The NusA 944 mutant protein was able to associate with the N-*nut* site complex, but was unable subsequently to assemble into a complex containing RNA polymerase and the remaining Nus factors, NusB, NusG, and S10. Moreover, the NusA 944 protein prevented the formation of a complete complex containing NusA R199A, indicating that the N-NusA-*nut* site complex must normally assemble first, with the other factors joining the complex later. Since the *nusA 944* mutation is also dominant to *nusA R199A in vivo* (Y. Zhou and D. Friedman, unpublished data), it is likely that NusA also associates first with the N-*nut* site complex *in vivo* prior to the assocation with the other Nus factors and RNA polymerase.

Materials and Methods

Plasmids, strains and reagents

RNA polymerase, N, NusA, NusA (1-495), NusA (1-416), NusB, and NusG were purified as described previously (Chapter 2; Burgess and Jendrisak, 1975; Greenblatt *et al.*, 1980; Greenblatt and Li, 1981b; Swindle *et al.*, 1988; Li *et al.*, 1992). NusA R199A and NusA 944 were provided by Y. Zhou and D. Friedman. S10 was provided by V. Nowotny.

Gel mobility shift experiments

Gel mobility shift assays were performed as previously described (T. Mah, Masters Thesis, 1995; Mogridge *et al.*, 1995). Concentrations of proteins are indicated in the figure legends. The sequences of the wild type *nut* site and the mutant *nut* sites are as follows:

pNUT wild type: 5' CGCUCUUACACAUUCCAGCCCUGAAAAAGGGC 3'

pNUT 1-3: 5' GCGUCUUACACAUUCCAGCCCUGAAAAAGGGC 3'

pNUT 4-6: 5' CGCGGGUACACAUUCCAGCCCUGAAAAAGGGC 3'

pNUT 7-9: 5' CGCUCUGCGACAUUCCAGCCCUGAAAAAGGGC 3'

pNUT 10-12: 5' CGCUCUUACCGCUUCCAGCCCUGAAAAAGGGC 3'

pNUT 13-17: 5' CGCUCUUACACAGGGGCGCCCUGAAAAAGGGC 3'

pNUTrev: 5' ACCUUACACAUUCUCGCGCCCUGAAAAAGGGC 3'

pNUTcon: 5' CGCUCUUUAACAUUCCAGCCCUGAAAAAGGGC 3'

The RNA probes contain additional vector sequences and are 54 nucleotides long.

Chapter 5

Thesis Summary and Future Directions

In order to understand the ways in which the NusA elongation factor carries out its functions in elongation, termination and antitermination, I have analyzed protein-protein and protein-RNA interactions involving NusA. A mutational study of *boxA* and the *interbox* established that nucleotides at the 3' end of *boxA* and the *interbox* are most important for the interaction of the λ *nut* site RNA with NusA. A systematic deletion study on NusA allowed the identification of regions of interaction with RNA polymerase and the λ N protein, and also provided evidence that the S1 homology region in NusA is important for its interaction with *nut* site RNA. The two RNA polymerase-binding regions of NusA were localized to amino acids 1-137 and 232-495. A carboxy-terminal region of NusA, from amino acids 303-495, bound the α subunit of RNA polymerase, and this region of NusA also interacted with the λ N protein. The identification of a carboxy-terminally truncated form of NusA that could bind RNA on its own further supported the proposal that NusA interacts directly with RNA. Finally, I developed a model to explain certain aspects of NusA function (Chapter 3, Fig. 7). The following is a discussion of what additional experiments could be done to confirm and add to the model.

In Chapter 4, I showed that the 3' portion of *boxA* and the *interbox* are important for the association of NusA with *nut* site RNA. Ribosomal protein S1 binds to a similar region in *rrn boxA* and also interacts with λ *boxA*, although the exact region was not determined (Mogridge and Greenblatt, 1998). These observations suggest that the S1 domains in the S1 protein and NusA may interact selectively with a similar binding site in *boxA* RNA. However, S1 competes with the NusB/S10 heterodimer for binding to *rrn boxA* RNA, and has no effect on non-processive antitermination by N, which depends only on N and NusA.

or *rrn*-mediated antitermination (Mogridge and Greenblatt, 1998). Since NusA is required for both of these processes, S1 probably does not compete with NusA for binding to *boxA* when other factors are present to aid the association of NusA. However, it would be of interest to determine if NusA and S1 can bind to *boxA* at the same time by performing gel mobility shift experiments as well as RNA pull-down experiments. Since both proteins bind to *boxA*, it is possible that the S1 domains within the proteins are interchangeable. To determine whether this is true, the S1 domain of NusA could be replaced with one from the S1 protein and the resulting NusA mutant could be tested for function in termination and antitermination. However, only a positive result would be meaningful in this type of experiment.

Ample evidence to support the theory that NusA does interact directly with RNA and that *boxA* and the *interbox* are important for the interaction was provided by my observations that the ability of NusA to associate with an N-*nut* site complex is affected by mutations in *boxA* and the *interbox* and that NusA (1-416) can bind RNA on its own (i.e. in the absence of N). Whereas genetic experiments had suggested that *boxA* could be the binding site for NusA, other experiments had suggested that *boxA* is unimportant for NusA function (Olson *et al.*, 1982; Friedman and Olson, 1983; Olson *et al.*, 1984; Zuber *et al.*, 1987; Patterson *et al.*, 1994). Although I have shown that *boxA* and the *interbox* represent an important binding site for NusA, there are probably other important binding sites that promote NusA function when the *nut* site is not present. To identify other important sequences for NusA binding, one could do an experiment where the *nut* site RNA is randomized and the resulting pool of RNAs is subjected to selection for binding to NusA (1-416) (ie a SELEX-type experiment). In addition to identifying new binding sites, this type of experiment should confirm the

importance of the wild type *nut* site since the *nut* site should also be identified in the selection. A more direct approach could also be used in which RNA sequences from pause and termination sites that are influenced by NusA are used as RNA probes in gel mobility shift experiments. Once new sequences that interact with NusA are identified, interactions with intact NusA in the presence of the α subunit of RNA polymerase would provide evidence for the relevance of the sequences in transcription. As well, the isolated S1 and KH homology regions from NusA could be used in binding experiments to determine whether different binding sites could be identified for each type of RNA-binding domain within NusA.

The interaction of NusA (1-416) with *boxA* RNA was similar to that of full length NusA in the presence of N. Similarly, the non-Watson-Crick G-A base pair in the *boxB* loop is necessary for the interaction in both cases, but the extruded fourth nucleotide of the *boxB* loop was not necessary for the binding of NusA (1-416) in the absence of N. Therefore, the fourth nucleotide of the *boxB* loop may facilitate the interaction of a portion of N with NusA. Alternatively, the RNA-binding modules in NusA may be aligned differently in the two different NusA constructs. To determine if this were true, one could perform limited proteolysis on both of the proteins and compare the respective proteolysis patterns. Furthermore, addition of *nut* site RNA to the proteins followed by proteolysis would provide data on whether similar fragments are protected from cleavage by the presence of RNA. The proteolysis products could be analyzed by mass spectrometry to further document any differences between the two proteins.

Since the interaction between NusA (1-416) and the *nut* site RNA is weak, it is not clear if all of the RNA-binding regions within the truncated NusA protein are available for

RNA-binding. It is possible that the conformation of the protein is not optimal for full RNA-binding activity. By introducing mutations that prevent RNA-binding into one or other of the RNA-binding regions of NusA (1-416), it may be possible to determine which RNA-binding domain(s) is active. Two such mutations in the S1 domain of NusA have been characterized. The *nusA1* mutation has a substitution of a leucine for an arginine at position 183 and NusA R199A has a mutation of a conserved arginine to an alanine that prevents the association of NusA with an N-*nut* site complex (Mogridge *et al.*, 1995; Y. Zhou, T. Mah, N. Yu, J. Mogridge, E. Olsen, J. Greenblatt and D. Friedman, unpublished data). Thus, if this mutation were introduced into NusA (1-416) and it prevented the ability of the protein to bind the *nut* site RNA, it would suggest that the S1 homology region in NusA (1-416) is necessary for binding. Conversely, if NusA R199A (1-416) still bound *nut* site RNA, this would imply that the KH domains may be the only ones involved in RNA-binding. This could be confirmed by the use of a similar type of mutation in either of the KH domains of NusA.

In Chapter 2, I present results of a systematic deletion study on NusA. Regions that interact with RNA polymerase, N and RNA were identified. To delineate domain boundaries more precisely, limited proteolysis should be done. The resulting fragments of NusA should represent intact domains that are not accessible to limiting amounts of protease. This mixture could be passed over affinity columns containing immobilized RNA polymerase or N to more precisely map the protein-interacting domains. The RNA-binding domains could be studied using the binding assays outlined above. The exact domain boundaries could be determined by subjecting the interacting fragments to mass spectrometry. The information generated from this type of study could then be used to further study the interactions between the proteins either by NMR, or by mutational analysis (see below).

The interaction between NusA and the λ N protein was studied in Chapters 2 and 3. Results from these studies suggested that one of the purposes of the interaction between the carboxy-terminal region of NusA and N is to relieve the inhibition of RNA-binding by NusA mediated by the carboxy-terminal region of NusA. This concept would explain why NusA (1-348) was completely functional in antitermination even though it did not bind N (Chapter 2). The model presented in Chapter 3 proposes that, once inhibition of RNA-binding by NusA is relieved, either the *nut* site RNA or the nascent RNA will be bound by NusA. Thus, the other purpose of the N-NusA interaction would be to bring the *nut* site into close proximity with NusA so that the *nut* site, as opposed to nascent RNA, is bound to NusA. To test this idea, one could seek mutations in NusA that prevent the N-NusA interaction, but not the NusA-α or NusA-RNA interaction. Using information generated from the domain mapping experiments proposed above, it would be possible to mutate specific amino acids in the defined N-binding domain of NusA to create a mutant protein unable to interact with N. Once purified, this mutant protein could be used in gel mobility shift experiments to determine whether the loss of N-binding ability by NusA would prevent the association of NusA with an N-*nut* site complex. In terms of functional assays, the prediction is that a mutant NusA protein that does not bind N would be unable to enhance antitermination by N, but would still be able to enhance termination as a result of its interaction with α.

Thus far, the data has been generated with *in vitro* binding and transcription assays. However, structural data could add greatly to our understanding of the functions of NusA. Therefore, an NMR or X-ray diffraction study of a complex containing NusA (132-416) which binds N and RNA, N (1-47) which binds *boxB* and NusA, and *nut* site RNA would be of great value. Although the entire complex would be very large by NMR standards,

structure determination would be simplified by the possibility of introducing heavy isotopes individually into NusA (132-416), N (1-47), and the *nut* site RNA.

Finally, a key question for understanding NusA function in antitermination is whether N and α compete for binding to the carboxy-terminal region of NusA. First, the minimal binding sites for both proteins should be mapped by the method of limited proteolysis as described above. To determine if N and α can bind NusA at the same time, one could perform a gradient sedimentation experiment with mixtures containing various combinations of NusA, N and α.

References

Adhya, S., Gottesman, M. and De Crombrugghe, B. (1974) Release of polarity in *Escherichia coli* by gene *N* of phage lambda: termination and antitermination of transcription. *Proc Natl Acad Sci U S A* **71**, 2534-8.

Albrechtsen, B., Squires, C.L., Li, S. and Squires, C. (1990) Antitermination of characterized transcriptional terminators by the *Escherichia coli rrnG* leader region. *J Mol Biol* **213**, 123-34.

Alifano, P., Rivellini, F., Limauro, D., Bruni, C.B. and Carlomagno, M.S. (1991) A consensus motif common to all Rho-dependent prokaryotic transcription terminators. *Cell* **64**, 553-63.

Altmann, C.R., Solow-Cordero, D.E. and Chamberlin, M.J. (1994) RNA cleavage and chain elongation by *Escherichia coli* DNA-dependent RNA polymerase in a binary enzyme-RNA complex. *Proc Natl Acad Sci U S A* **91**, 3784-8.

Amster-Choder, O., Houman, F. and Wright, A. (1989) Protein phosphorylation regulates transcription of the beta-glucoside utilization operon in *E. coli*. *Cell* **58**, 847-55.

Amster-Choder, O. and Wright, A. (1990) Regulation of activity of a transcriptional anti-terminator in *E. coli* by phosphorylation *in vivo*. *Science* **249**, 540-2.

Amster-Choder, O. and Wright, A. (1992) Modulation of the dimerization of a transcriptional antiterminator protein by phosphorylation. *Science* **257**, 1395-8.

Arndt, K.M. and Chamberlin, M.J. (1990) RNA chain elongation by *Escherichia coli* RNA polymerase. Factors affecting the stability of elongating ternary complexes. *J Mol Biol* **213**, 79-108.

Artsimovitch, I. and Landick, R. (1998) Interaction of a nascent RNA structure with RNA polymerase is required for hairpin-dependent transcriptional pausing but not for transcript release. *Genes Dev* **12**, 3110-22.

Banik-Maiti, S., King, R.A. and Weisberg, R.A. (1997) The antiterminator RNA of phage HK022. *J Mol Biol* **272**, 677-87.

Barik, S., Ghosh, B., Whalen, W., Lazinski, D. and Das, A. (1987) An antitermination protein engages the elongating transcription apparatus at a promoter-proximal recognition site. *Cell* **50**, 885-99.

Baron, J. and Weisberg, R.A. (1992) Mutations of the phage lambda *nutL* region that prevent the action of Nun, a site-specific transcription termination factor. *J Bacteriol* **174**, 1983-9.

Baron, L.S., Penido, E., Ryman, I.R. and Falkow, S. (1970) Behavior of coliphage lambda in hybrids between *Escherichia coli and Salmonella*. *J Bacteriol* **102**, 221-33.

Battiste, J.L., Mao, H., Rao, N.S., Tan, R., Muhandiram, D.R., Kay, L.E., Frankel, A.D. and Williamson, J.R. (1996) Alpha helix-RNA major groove recognition in an HIV-1 rev peptide-RRE RNA complex. *Science* **273**, 1547-51.

Berg, K.L., Squires, C. and Squires, C.L. (1989) Ribosomal RNA operon anti-termination. Function of leader and spacer region *box B-box A* sequences and their conservation in diverse micro- organisms. *J Mol Biol* **209**, 345-58.

Blatter, E.E., Ross, W., Tang, H., Gourse, R. L. and Ebright, R. H. (1994) Domain organization of RNA polymerase alpha subunit: C-terminal 85 amino acids constitute a domain capable of dimerization and DNA binding. *Cell* **78**, 889-96.

Bokal, A.J., Ross, W., Gaal, T., Johnson, R.C. and Gourse, R.L. (1997) Molecular anatomy of a transcription activation patch: FIS-RNA polymerase interactions at the *Escherichia coli rrnB* P1 promoter. *Embo J* **16**, 154-62.

Boni, I.V., Isaeva, D.M., Musychenko, M.L. and Tzareva, N.V. (1991) Ribosome-messenger recognition: mRNA target sites for ribosomal protein S1. *Nucleic Acids Res* **19**, 155-62.

Borukhov, S., Polyakov, A., Nikiforov, V. and Goldfarb, A. (1992) GreA protein: a transcription elongation factor from *Escherichia coli*. *Proc Natl Acad Sci U S A* **89**, 8899-902.

Borukhov, S., Sagitov, V. and Goldfarb, A. (1993) Transcript cleavage factors from *E. coli*. *Cell* **72**, 459-66.

Brachet, P., Eisen, H. and Rambach, A. (1970) Mutations of coliphage lambda affecting the expression of replicative functions O and P. *Mol Gen Genet* **108**, 266-76.

Brennan, C.A., Dombroski, A.J. and Platt, T. (1987) Transcription termination factor rho is an RNA-DNA helicase. *Cell* **48**, 945-52.

Burgess, R.R. and Jendrisak, J.J. (1975) A procedure for the rapid, large-scale purification of *Escherichia coli* DNA-dependent RNA polymerase involving Polymin P precipitation and DNA- cellulose chromatography. *Biochemistry* **14**, 4634-8.

Burgess, R.R., Travers, A.A., Dunn, J.J. and Bautz, E.K. (1969) Factor stimulating transcription by RNA polymerase. *Nature* **221**, 43-6.

Burns, C.M. and Richardson, J.P. (1995) NusG is required to overcome a kinetic limitation to Rho function at an intragenic terminator. *Proc Natl Acad Sci U S A* **92**, 4738-42.

Burova, E., Hung, S.C., Sagitov, V., Stitt, B.L. and Gottesman, M.E. (1995) *Escherichia coli* NusG protein stimulates transcription elongation rates *in vivo* and *in vitro*. *J Bacteriol* **177**, 1388-92.

Bycroft, M., Hubbard, T.J., Proctor, M., Freund, S.M. and Murzin, A.G. (1997) The solution structure of the S1 RNA binding domain: a member of an ancient nucleic acid-binding fold. *Cell* **88**, 235-42.

Cai, Z., Gorin, A., Frederick, R., Ye, X., Hu, W., Majumdar, A., Kettani, A. and Patel, D.J. (1998) Solution structure of P22 transcriptional antitermination N peptide- *boxB* RNA complex. *Nat Struct Biol* **5**, 203-12.

Carpousis, A.J. and Gralla, J.D. (1980) Cycling of ribonucleic acid polymerase to produce oligonucleotides during initiation *in vitro* at the lac UV5 promoter. *Biochemistry* **19**, 3245-53.

Carpousis, A.J. and Gralla, J.D. (1985) Interaction of RNA polymerase with lac UV5 promoter DNA during mRNA initiation and elongation. Footprinting, methylation, and rifampicin- sensitivity changes accompanying transcription initiation. *J Mol Biol* **183**, 165-77.

Chadsey, M.S., Karlinsey, J.E. and Hughes, K.T. (1998) The flagellar anti-sigma factor FlgM actively dissociates *Salmonella typhimurium* sigma 28 RNA polymerase holoenzyme. *Genes Dev* **12**, 3123-36.

Chamberlin, M.J. (1992) New models for the mechanism of transcription elongation and its regulation. *Harvey Lect* **88**, 1-21.

Chan, C.L. and Landick, R. (1993) Dissection of the *his* leader pause site by base substitution reveals a multipartite signal that includes a pause RNA hairpin. *J Mol Biol* **233**, 25-42.

Chan, C.L. and Landick, R. (1994) New Perspectives on RNA Chain Elongation and Termination by *E.coli* RNA polymerase. In *Transcription: Mechanisms and Regulation* . Conaway, R. C. and Conaway, J. W. (eds.) Raven Press, Ltd., New York, pp. 297-321.

Chan, C.L., Wang, D. and Landick, R. (1997) Multiple interactions stabilize a single paused transcription intermediate in which hairpin to 3' end spacing distinguishes pause and termination pathways. *J Mol Biol* **268**, 54-68.

Chattopadhyay, S., Garcia-Mena, J., DeVito, J., Wolska, K. and Das, A. (1995a) Bipartite function of a small RNA hairpin in transcription antitermination in bacteriophage lambda. *Proc Natl Acad Sci U S A* **92**, 4061-5.

Chattopadhyay, S., Hung, S.C., Stuart, A.C., Palmer, A.G., Garcia-Mena, J., Das, A. and Gottesman, M.E. (1995b) Interaction between the phage HK022 Nun protein and the *nut* RNA of phage lambda. *Proc Natl Acad Sci U S A* **92**, 12131-5.

Chen, C.Y., Galluppi, G.R. and Richardson, J.P. (1986) Transcription termination at lambda tR1 is mediated by interaction of rho with specific single-stranded domains near the 3' end of *cro* mRNA. *Cell* **46**, 1023-8.

Chen, C.Y. and Richardson, J.P. (1987) Sequence elements essential for rho-dependent transcription termination at lambda tR1. *J Biol Chem* **262**, 11292-9.

Chen, Y.H., Yang, J.T. and Martinez, H.M. (1972) Determination of the secondary structures of proteins by circular dichroism and optical rotatory dispersion. *Biochemistry* **11**,4120-31.

Clerget, M., Jin, D.J. and Weisberg, R.A. (1995) A zinc-binding region in the beta' subunit of RNA polymerase is involved in antitermination of early transcription of phage HK022. *J Mol Biol* **248**, 768-80.

Court, D.L., Patterson, T.A., Baker, T., Costantino, N., Mao, X. and Friedman, D.I. (1995) Structural and functional analyses of the transcription-translation proteins NusB and NusE. *J Bacteriol* **177**, 2589-91.

Cowing, D.W., Mecsas, J., Record, M.T., Jr. and Gross, C.A. (1989) Intermediates in the formation of the open complex by RNA polymerase holoenzyme containing the sigma factor sigma 32 at the groE promoter. *J Mol Biol* **210**, 521-30.

Craven, M.G. and Friedman, D.I. (1991) Analysis of the *Escherichia coli nusA10*(Cs) allele: relating nucleotide changes to phenotypes. *J Bacteriol* **173**, 1485-91.

Craven, M.G., Granston, A.E., Schauer, A.T., Zheng, C., Gray, T.A. and Friedman, D.I. (1994) *Escherichia coli-Salmonella typhimurium* hybrid *nusA* genes: identification of a short motif required for action of the lambda N transcription antitermination protein. *J Bacteriol* **176**, 1394-404.

Dambly, C. and Couturier, M. (1971) A minor Q-independent pathway for the expression of the late genes in bacteriophage lambda. *Mol Gen Genet* **113**, 244-50.

Das, A., Court, D. and Adhya, S. (1976) Isolation and characterization of conditional lethal mutants of *Escherichia coli* defective in transcription termination factor rho. *Proc Natl Acad Sci U S A* **73**, 1959-63.

Das, A., Ghosh, B., Barik, S. and Wolska, K. (1985) Evidence that ribosomal protein S10 itself is a cellular component necessary for transcription antitermination by phage lambda N protein. *Proc Natl Acad Sci U S A* **82**, 4070-4.

Das, A. and Wolska, K. (1984) Transcription antitermination *in vitro* by lambda *N* gene product: requirement for a phage *nut* site and the products of host *nusA*, *nusB*, and *nusE* genes. *Cell* **38**, 165-73.

de Crombrugghe, B., Mudryj, M., DiLauro, R. and Gottesman, M. (1979) Specificity of the bacteriophage lambda *N* gene product (pN): *nut* sequences are necessary and sufficient for antitermination by pN. *Cell* **18**, 1145-51.

Dejgaard, K. and Leffers, H. (1996) Characterisation of the nucleic acid-binding activity of KH domains. Different properties of different domains. *Eur J Biochem* **241**, 425-31.

DeVito, J. and Das, A. (1994) Control of transcription processivity in phage lambda: Nus factors strengthen the termination-resistant state of RNA polymerase induced by N antiterminator. *Proc Natl Acad Sci U S A* **91**, 8660-4.

Dodson, R.E. and Shapiro, D.J. (1997) Vigilin, a ubiquitous protein with 14 K homology domains, is the estrogen-inducible vitellogenin mRNA 3'-untranslated region-binding protein. *J Biol Chem* **272**, 12249-52.

Doelling, J.H. and Franklin, N.C. (1989) Effects of all single base substitutions in the loop of *boxB* on antitermination of transcription by bacteriophage lambda's N protein. *Nucleic Acids Res* **17**, 5565-77.

Dombroski, A.J., Walter, W.A. and Gross, C.A. (1993) Amino-terminal amino acids modulate sigma-factor DNA-binding activity. *Genes Dev* **7**, 2446-55.

Dombroski, A.J., Walter, W.A., Record, M.T., Jr., Siegele, D.A. and Gross, C.A. (1992) Polypeptides containing highly conserved regions of transcription initiation factor sigma 70 exhibit specificity of binding to promoter DNA. *Cell* **70**, 501-12.

Ebright, R. H. and Busby, S. (1995) The *Escherichia coli* RNA polymerase alpha subunit: structure and function. *Curr Opin Genet Dev* **5**, 197-203.

Farnham, P.J., Greenblatt, J. and Platt, T. (1982) Effects of NusA protein on transcription termination in the *tryptophan* operon of *Escherichia coli*. *Cell* **29**, 945-51.

Faus, I., Chen, C.Y. and Richardson, J.P. (1988) Sequences in the 5' proximal segment of the paused transcript affect NusA-mediated enhancement of transcriptional pausing. *J Biol Chem* **263**, 10830-5.

Finger, L.R. and Richardson, J.P. (1982) Stabilization of the hexameric form of *Escherichia coli* protein rho under ATP hydrolysis conditions. *J Mol Biol* **156**, 203-19.

Franklin, N.C. (1985a) Conservation of genome form but not sequence in the transcription antitermination determinants of bacteriophages lambda, phi 21 and P22. *J Mol Biol* **181**, 75-84.

Franklin, N.C. (1985b) "N" transcription antitermination proteins of bacteriophages lambda, phi 21 and P22. *J Mol Biol* **181**, 85-91.

Franklin, N.C. (1993) Clustered arginine residues of bacteriophage lambda N protein are essential to antitermination of transcription, but their locale cannot compensate for *boxB* loop defects. *J Mol Biol* **231**, 343-60.

Friedman, D. (1971) A bacterial mutant affecting lambda development. In *The Bacteriophage Lambda*. A. D. Hershey (ed.), Cold Spring Harbour Laboratory, pp. 733-738.

Friedman, D.I. (1988) Regulation of Phage Gene Expression by Termination and Antitermination of Transcription. In *The Bacteriophages*. R. Calendar (ed.). Plenum Publishing Corporation, Vol. 2, pp. 263-319.

Friedman, D.I. and Baron, L.S. (1974) Genetic characterization of a bacterial locus involved in the activity of the *N* function of phage lambda. *Virology* **58**, 141-8.

Friedman, D.I., Baumann, M. and Baron, L.S. (1976) Cooperative effects of bacterial mutations affecting lambda N gene expression. I. Isolation and characterization of a *nusB* mutant. *Virology* **73**, 119-27.

Friedman, D.I., Jolly, C.T. and Mural, R.J. (1973) Interference with the expression of the *N* gene function of phage lambda in a mutant of *Escherichia coli*. *Virology* **51**, 216-26.

Friedman, D.I. and Olson, E.R. (1983) Evidence that a nucleotide sequence, "*boxA*," is involved in the action of the NusA protein. *Cell* **34**, 143-9.

Friedman, D.I., Olson, E.R., Johnson, L.L., Alessi, D. and Craven, M.G. (1990) Transcription-dependent competition for a host factor: the function and optimal sequence of the phage lambda *boxA* transcription antitermination signal. *Genes Dev* **4**, 2210-22.

Friedman, D.I. and Ponce-Campos, R. (1975) Differential effect of phage regulator functions on transcription from various promoters: evidence that the P22 gene *24* and the lambda gene *N* products distinguish three classes of promoters. *J Mol Biol* **98**, 537-49.

Friedman, D.I., Schauer, A.T., Baumann, M.R., Baron, L.S. and Adhya, S.L. (1981) Evidence that ribosomal protein S10 participates in control of transcription termination. *Proc Natl Acad Sci U S A* **78**, 1115-8.

Gamper, H.B. and Hearst, J.E. (1982) A topological model for transcription based on unwinding angle analysis of *E. coli* RNA polymerase binary, initiation and ternary complexes. *Cell* **29**, 81-90.

Gardella, T., Moyle, H. and Susskind, M.M. (1989) A mutant *Escherichia coli* sigma 70 subunit of RNA polymerase with altered promoter specificity. *J Mol Biol* **206**, 579-90.

Gentry, D., Xiao, H., Burgess, R. and Cashel, M. (1991) The omega subunit of *Escherichia coli* K-12 RNA polymerase is not required for stringent RNA control *in vivo*. *J Bacteriol* **173**, 3901-3.

Georgopoulos, C.P. (1971a) Bacterial mutants in which the gene *N* function of bacteriophage lambda is blocked have an altered RNA polymerase. *Proc Natl Acad Sci U S A* **68**, 2977-81.

Georgopoulos, C.P. (1971b) A bacterial mutation affecting *N* function. In *The Bacteriophage Lambda*. A. Hershey (ed.). Cold Spring Harbour Laboratory, Cold Spring Harbour, pp. 639-645.

Georgopoulos, C.P., Swindle, J., Keppel, F., Ballivet, M., Bisig, R. and Eisen, H. (1980) Studies on the *E. coli groNB (nusB)* gene which affects bacteriophage lambda *N* gene function. *Mol Gen Genet* **179**, 55-61.

Ghosh, B. and Das, A. (1984) NusB: a protein factor necessary for transcription antitermination *in vitro* by phage lambda *N* gene product. *Proc Natl Acad Sci U S A* **81**, 6305-9.

Ghysen, A. and Pironio, M. (1972) Relationship between the *N* function of bacteriophage lambda and host RNA polymerase. *J Mol Biol* **65**, 259-72.

Gibson, T.J., Thompson, J.D. and Heringa, J. (1993) The KH domain occurs in a diverse set of RNA-binding proteins that include the antiterminator NusA and is probably involved in binding to nucleic acid. *FEBS Lett* **324**, 361-6.

Goda, Y. and Greenblatt, J. (1985) Efficient modification of *E. coli* RNA polymerase *in vitro* by the *N* gene transcription antitermination protein of bacteriophage lambda. *Nucleic Acids Res* **13**, 2569-82.

Grayhack, E.J. and Roberts, J.W. (1982) The phage lambda *Q* gene product: activity of a transcription antiterminator *in vitro*. *Cell* **30**, 637-48.

Grayhack, E.J., Yang, X.J., Lau, L.F. and Roberts, J.W. (1985) Phage lambda gene *Q* antiterminator recognizes RNA polymerase near the promoter and accelerates it through a pause site. *Cell* **42**, 259-69.

Greenblatt, J. (1992) Protein-Protein Interactions as Critical Determinants of Regulated Initiation and Termination of Transcription. In *Transcriptional Regulation*. S. L. McKnight and K. R. Yamamoto (eds.). CSHL Press, Cold Spring Harbour, N.Y., pp. 203-226.

Greenblatt, J. and Li, J. (1981a) Interaction of the sigma factor and the *nusA* gene protein of *E. coli* with RNA polymerase in the initiation-termination cycle of transcription. *Cell* **24**, 421-8.

Greenblatt, J. and Li, J. (1981b) The *nusA* gene protein of *Escherichia coli*. Its identification and a demonstration that it interacts with the gene *N* transcription anti-termination protein of bacteriophage lambda. *J Mol Biol* **147**, 11-23.

Greenblatt, J., Li, J., Adhya, S., Friedman, D.I., Baron, L.S., Redfield, B., Kung, H.F. and Weissbach, H. (1980) L factor that is required for beta-galactosidase synthesis is the *nusA* gene product involved in transcription termination. *Proc Natl Acad Sci U S A* **77**, 1991-4.

Greenblatt, J., Mah, T., Legault, P., Mogridge, J., Li, J. and Kay, L.E. (1998) Structure and mechanism in transcriptional antitermination by the bacteriophage lambda N protein. In *Cold Spring Harbour Symposium on Quantitative Biology*. CSHL Press, Cold Spring Harbour, N.Y., Vol. 63, pp. 327-336.

Greenblatt, J., Malnoe, P. and Li, J. (1980) Purification of the gene *N* transcription anti-termination protein of bacteriophage lambda. *J Biol Chem* **255**, 1465-70.

Greenblatt, J., McLimont, M. and Hanly, S. (1981) Termination of transcription by *nusA* gene protein of *Escherichia coli*. *Nature* **292**, 215-20.

Greenblatt, J., Nodwell, J.R. and Mason, S.W. (1993) Transcriptional antitermination. *Nature* **364**, 401-6.

Gross, C.A., Lonetto, M. and Losick, R. (1992) Bacterial sigma factors. In *Transcriptional Regulation* . S. L. McKnight and K. Yamamoto (eds.), CSHL Press, Cold Spring Harbour, pp. 129-176.

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

Hanna, M.M. and Meares, C.F. (1983) Topography of transcription: path of the leading end of nascent RNA through the *Escherichia coli* transcription complex. *Proc Natl Acad Sci U S A* **80**, 4238–42.

Hansen, U.M. and McClure, W.R. (1980) Role of the sigma subunit of *Escherichia coli* RNA polymerase in initiation. II. Release of sigma from ternary complexes. *J Biol Chem* **255**, 9564-70.

Harley, C.B. and Reynolds, R.P. (1987) Analysis of *E. coli* promoter sequences. *Nucleic Acids Res* **15**, 2343-61.

Hawley, D.K., Johnson, A.D. and McClure, W.R. (1985) Functional and physical characterization of transcription initiation complexes in the bacteriophage lambda OR region. *J Biol Chem* **260**, 8618-26.

Hawley, D.K. and McClure, W.R. (1983) Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids Res* **11**, 2237-55.

Hofer, B., Muller, D. and Koster, H. (1985) The pathway of *E. coli* RNA polymerase-promoter complex formation as visualized by footprinting. *Nucleic Acids Res* **13**, 5995-6013.

Holben, W.E. and Morgan, E.A. (1984) Antitermination of transcription from an *Escherichia coli* ribosomal RNA promoter. *Proc Natl Acad Sci U S A* **81**, 6789-93.

Horwitz, R.J., Li, J. and Greenblatt, J. (1987) An elongation control particle containing the *N* gene transcriptional antitermination protein of bacteriophage lambda. *Cell* **51**, 631-41.

Houman, F., Diaz-Torres, M.R. and Wright, A. (1990) Transcriptional antitermination in the *bgl* operon of *E. coli* is modulated by a specific RNA binding protein. *Cell* **62**, 1153-63.

Huenges, M., Rolz, C., Gschwind, R., Peteranderl, R., Berglechner, F., Richter, G., Bacher, A., Kessler, H. and Gemmecker, G. (1998) Solution structure of the antitermination protein NusB of *Escherichia coli*: a novel all-helical fold for an RNA-binding protein. *Embo J* **17**, 4092-100.

Hung, S.C. and Gottesman, M.E. (1997) The Nun protein of bacteriophage HK022 inhibits translocation of *Escherichia coli* RNA polymerase without abolishing its catalytic activities. *Genes Dev* **11**, 2670-8.

Ito, K., Egawa, K. and Nakamura, Y. (1991) Genetic interaction between the beta' subunit of RNA polymerase and the arginine-rich domain of *Escherichia coli* NusA protein. *J Bacteriol* **173**, 1492-501.

Ito, K. and Nakamura, Y. (1993) Pleiotropic effects of the *rpoC10* mutation affecting the RNA polymerase beta' subunit of *Escherichia coli* on factor-dependent transcription termination and antitermination. ***Mol Microbiol* 9**, 285-93.

Ito, K. and Nakamura, Y. (1996) Localization of *nusA*-suppressing amino acid substitutions in the conserved regions of the beta' subunit of *Escherichia coli* RNA polymerase. *Mol Gen Genet* **251**, 699-706.

Jin, D.J., Burgess, R.R., Richardson, J.P. and Gross, C.A. (1992) Termination efficiency at rho-dependent terminators depends on kinetic coupling between RNA polymerase and rho. *Proc Natl Acad Sci U S A* **89**, 1453-7.

Jin, D.J., Walter, W.A. and Gross, C.A. (1988) Characterization of the termination phenotypes of rifampicin-resistant mutants. *J Mol Biol* **202**, 245-53.

Johnston, D.E. and McClure, W.R. (1976) Abortive initiation of *in vitro* RNA synthesis on bacteriophage lambda DNA. In *RNA Polymerase*. Losick, R. and Chamberlin M. (eds.), CSHL Press, Cold Spring Harbor Laboratory, pp. 413-28.

Kainz, M. and Roberts, J. (1992) Structure of transcription elongation complexes *in vivo*. *Science* **255**, 838-41.

Kassavetis, G.A. and Chamberlin, M.J. (1981) Pausing and termination of transcription within the early region of bacteriophage T7 DNA *in vitro*. *J Biol Chem* **256**, 2777-86.

Kay, L.E., Keifer, P., and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J Am Chem Soc* **114**, 10663-5.

Keppel, F., Georgopoulos, C.P. and Eisen, H. (1974) Host interference with expression of the lambda *N* gene product. *Biochimie* **56**, 1505-9.

Kimura, M., Fujita, N. and Ishihama, A. (1994) Functional map of the alpha subunit of *Escherichia coli* RNA polymerase. Deletion analysis of the amino-terminal assembly domain. *J Mol Biol* **242**, 107-15.

King, R.A., Banik-Maiti, S., Jin, D.J. and Weisberg, R.A. (1996) Transcripts that increase the processivity and elongation rate of RNA polymerase. *Cell* **87**, 893-903.

Kingston, R.E. and Chamberlin, M.J. (1981) Pausing and attenuation of *in vitro* transcription in the *rrnB* operon of *E. coli*. *Cell* **27**, 523-31.

Kingston, R.E., Nierman, W.C. and Chamberlin, M.J. (1981) A direct effect of guanosine tetraphosphate on pausing of *Escherichia coli* RNA polymerase during RNA chain elongation. *J Biol Chem* **256**: 2787-97.

Komissarova, N. and Kashlev, M. (1997a) RNA polymerase switches between inactivated and activated states by translocating back and forth along the DNA and the RNA. *J Biol Chem* **272**, 15329-38.

Komissarova, N. and Kashlev, M. (1997b) Transcriptional arrest: *Escherichia coli* RNA polymerase translocates backward, leaving the 3' end of the RNA intact and extruded. *Proc Natl Acad Sci U S A* **94**, 1755-60.

Komissarova, N. and Kashlev, M. (1998) Functional topography of nascent RNA in elongation intermediates of RNA polymerase. *Proc Natl Acad Sci U S A* **95**, 14699-704.

Krummel, B. and Chamberlin, M.J. (1992a) Structural analysis of ternary complexes of *Escherichia coli* RNA polymerase. Deoxyribonuclease I footprinting of defined complexes. *J Mol Biol* **225**, 239-50.

Krummel, B. and Chamberlin, M.J. (1992b) Structural analysis of ternary complexes of *Escherichia coli* RNA polymerase. Individual complexes halted along different

transcription units have distinct and unexpected biochemical properties. *J Mol Biol* **225**, 221-37.

Kumar, S.A. and Krakow, J.S. (1975) Studies on the product binding sites of the *Azotobacter vinelandii* ribonucleic acid polymerase. *J Biol Chem* **250**, 2878-84.

Kung, H., Spears, C. and Weissbach, H. (1975) Purification and properties of a soluble factor required for the deoxyribonucleic acid-directed *in vitro* synthesis of beta-galactosidase. *J Biol Chem* **250**, 1556-62.

Landick, R. and Turnbough, C.L. (1992) Transcriptional Attenuation. In *Transcriptional Regulation*. S. L. McKnight and K. R. Yamamoto (eds.), CSHL Press, Cold Spring Harbour, pp. 407-446.

Landick, R., Turnbough, C.L. and Yanofsky, C. (1996) Transcription Attenuation. In *Escherichia coli and Salmonella*. F. C. Neidhartdt (ed.), ASM Press, Washington, pp. 1263-1286.

Landick, R. and Yanofsky, C. (1984) Stability of an RNA secondary structure affects *in vitro* transcription pausing in the *trp* operon leader region. *J Biol Chem* **259**, 11550-5.

Landick, R. and Yanofsky, C. (1987) Isolation and structural analysis of the *Escherichia coli trp* leader paused transcription complex. *J Mol Biol* **196**, 363-77.

Lau, L.F., Roberts, J.W. and Wu, R. (1982) Transcription terminates at lambda tR1 in three clusters. *Proc Natl Acad Sci U S A* **79**, 6171-5.

Lau, L.F., Roberts, J.W. and Wu, R. (1983) RNA polymerase pausing and transcript release at the lambda tR1 terminator *in vitro*. *J Biol Chem* **258**, 9391-7.

Lazinski, D., Grzadzielska, E. and Das, A. (1989) Sequence-specific recognition of RNA hairpins by bacteriophage antiterminators requires a conserved arginine-rich motif. *Cell* **59**, 207-18.

Lee, J. and Goldfarb, A. (1991) lac repressor acts by modifying the initial transcribing complex so that it cannot leave the promoter. *Cell* **66**, 793-8.

Legault, P., Li, J., Mogridge, J., Kay, L.E. and Greenblatt, J. (1998) NMR structure of the bacteriophage lambda N peptide/*boxB* RNA complex: recognition of a GNRA fold by an arginine-rich motif. *Cell* **93**, 289-99.

Levin, J.R. and Chamberlin, M.J. (1987) Mapping and characterization of transcriptional pause sites in the early genetic region of bacteriophage T7. *J Mol Biol* **196**, 61-84.

Levin, J.R., Krummel, B. and Chamberlin, M.J. (1987) Isolation and properties of transcribing ternary complexes of *Escherichia coli* RNA polymerase positioned at a single template base. *J Mol Biol* **196**, 85-100.

Li, J., Horwitz, R., McCracken, S. and Greenblatt, J. (1992) NusG, a new *Escherichia coli* elongation factor involved in transcriptional antitermination by the N protein of phage lambda. *J Biol Chem* **267**, 6012-9.

Li, J., Mason, S.W. and Greenblatt, J. (1993) Elongation factor NusG interacts with termination factor rho to regulate termination and antitermination of transcription. *Genes Dev* **7**, 161-72.

Li, M., Moyle, H. and Susskind, M.M. (1994) Target of the transcriptional activation function of phage lambda cI protein. *Science* **263**, 75-7.

Lisser, S. and Margalit, H. (1993) Compilation of *E. coli* mRNA promoter sequences. *Nucleic Acids Res* **21**, 1507-16.

Liu, K. and Hanna, M.M. (1995a) NusA contacts nascent RNA in *Escherichia coli* transcription complexes. *J Mol Biol* **247**, 547-58.

Liu, K. and Hanna, M.M. (1995b) NusA interferes with interactions between the nascent RNA and the C- terminal domain of the alpha subunit of RNA polymerase in *Escherichia coli* transcription complexes. *Proc Natl Acad Sci U S A* **92**, 5012-6.

Liu, K., Zhang, Y., Severinov, K., Das, A. and Hanna, M.M. (1996) Role of *Escherichia coli* RNA polymerase alpha subunit in modulation of pausing, termination and anti-termination by the transcription elongation factor NusA. *Embo J* **15**, 150-61.

Losick, R. and Pero, J. (1981) Cascades of Sigma factors. *Cell* **25**, 582-4.

Lowery-Goldhammer, C. and Richardson, J.P. (1974) An RNA-dependent nucleoside triphosphate phosphohydrolase (ATPase) associated with rho termination factor. *Proc Natl Acad Sci U S A* **71**, 2003-7.

Mahadevan, S., Reynolds, A.E. and Wright, A. (1987) Positive and negative regulation of the *bgl* operon in *Escherichia coli*. *J Bacteriol* **169**, 2570-8.

Mahadevan, S. and Wright, A. (1987) A bacterial gene involved in transcription antitermination: regulation at a rho-independent terminator in the *bgl* operon of *E. coli*. *Cell* **50**, 485-94.

Malan, T.P., Kolb, A., Buc, H. and 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.

Malhotra, A., Severinova, E. and Darst, S. A. (1996) Crystal structure of a sigma 70 subunit fragment from *E.coli* RNA polymerase. *Cell* **87**, 127-36.

Mason, S.W. and Greenblatt, J. (1991) Assembly of transcription elongation complexes containing the N protein of phage lambda and the *Escherichia coli* elongation factors NusA, NusB, NusG, and S10. *Genes Dev* **5**, 1504-12.

Mason, S.W., Li, J. and Greenblatt, J. (1992a) Direct interaction between two *Escherichia coli* transcription antitermination factors, NusB and ribosomal protein S10. *J Mol Biol* **223**, 55-66.

Mason, S.W., Li, J. and Greenblatt, J. (1992b) Host factor requirements for processive antitermination of transcription and suppression of pausing by the N protein of bacteriophage lambda. *J Biol Chem* **267**, 19418-26.

Mecsas, J., Cowing, D.W. and Gross, C.A. (1991) Development of RNA polymerase-promoter contacts during open complex formation. *J Mol Biol* **220**, 585-97.

Mogridge, J. and Greenblatt, J. (1998) Specific binding of *Escherichia coli* ribosomal protein S1 to *boxA* transcriptional antiterminator RNA. *J Bacteriol* **180**, 2248-52.

Mogridge, J., Legault, P., Li, J., Van Oene, M.D., Kay, L.E. and Greenblatt, J. (1998a) Independent ligand-induced folding of the RNA-binding domain and two functionally distinct antitermination regions in the phage lambda N protein. *Mol Cell* **1**, 265-75.

Mogridge, J., Mah, T.F. and Greenblatt, J. (1995) A protein-RNA interaction network facilitates the template-independent cooperative assembly on RNA polymerase of a stable antitermination complex containing the lambda N protein. *Genes Dev* **9**, 2831-45.

Mogridge, J., Mah, T.F. and Greenblatt, J. (1998b) Involvement of *boxA* nucleotides in the formation of a stable ribonucleoprotein complex containing the bacteriophage lambda N protein. *J Biol Chem* **273**, 4143-8.

Morgan, E.A. (1980) Insertions of Tn 10 into an *E. coli* ribosomal RNA operon are incompletely polar. *Cell* **21**, 257-65.

Morgan, W.D., Bear, D.G., Litchman, B.L. and von Hippel, P.H. (1985) RNA sequence and secondary structure requirements for rho-dependent transcription termination. *Nucleic Acids Res* **13**, 3739-54.

Murakami, K., Fujita, N. and Ishihama, A. (1996) Transcription factor recognition surface on the RNA polymerase alpha subunit is involved in contact with the DNA enhancer element. *Embo J* **15**, 4358-67.

Nehrke, K.W. and Platt, T. (1994) A quaternary transcription termination complex. Reciprocal stabilization by Rho factor and NusG protein. *J Mol Biol* **243**, 830-9.

Nehrke, K.W., Zalatan, F. and Platt, T. (1993) NusG alters rho-dependent termination of transcription *in vitro* independent of kinetic coupling. *Gene Expr* **3**, 119-33.

Nodwell, J.R. and Greenblatt, J. (1991) The *nut* site of bacteriophage lambda is made of RNA and is bound by transcription antitermination factors on the surface of RNA polymerase. *Genes Dev* **5**, 2141-51.

Nodwell, J.R. and Greenblatt, J. (1993) Recognition of *boxA* antiterminator RNA by the *E. coli* antitermination factors NusB and ribosomal protein S10. *Cell* **72**, 261-8.

Nudler, E. (1999) Transcription elongation: structural basis and mechanisms. *J Mol Biol* **288**, 1-12.

Nudler, E., Avetissova, E., Markovtsov, V. and Goldfarb, A. (1996) Transcription processivity: protein-DNA interactions holding together the elongation complex. *Science* **273**, 211-7.

Nudler, E., Goldfarb, A. and Kashlev, M. (1994) Discontinuous mechanism of transcription elongation. *Science* **265**, 793-6.

Nudler, E., Kashlev, M., Nikiforov, V. and Goldfarb, A. (1995) Coupling between transcription termination and RNA polymerase inchworming. *Cell* **81**, 351-7.

Nudler, E., Mustaev, A., Lukhtanov, E. and Goldfarb, A. (1997) The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. *Cell* **89**, 33-41.

Oberto, J., Clerget, M., Ditto, M., Cam, K. and Weisberg, R.A. (1993) Antitermination of early transcription in phage HK022. Absence of a phage-encoded antitermination factor. *J Mol Biol* **229**, 368-81.

Olson, E.R., Flamm, E.L. and Friedman, D.I. (1982) Analysis of *nutR*: a region of phage lambda required for antitermination of transcription. *Cell* **31**, 61-70.

Olson, E.R., Tomich, C.S. and Friedman, D.I. (1984) The *nusA* recognition site. Alteration in its sequence or position relative to upstream translation interferes with the action of the N antitermination function of phage lambda. *J Mol Biol* **180**, 1053-63.

Orlova, M., Newlands, J., Das, A., Goldfarb, A. and Borukhov, S. (1995) Intrinsic transcript cleavage activity of RNA polymerase. *Proc Natl Acad Sci U S A* **92**, 4596-600.

Patterson, T.A., Zhang, Z., Baker, T., Johnson, L.L., Friedman, D.I. and Court, D.L. (1994) Bacteriophage lambda N-dependent transcription antitermination. Competition for an RNA site may regulate antitermination. *J Mol Biol* **236**, 217-28.

Peltz, S.W., Brown, A.L., Hasan, N., Podhajska, A.J. and Szybalski, W. (1985) Thermosensitivity of a DNA recognition site: activity of a truncated *nutL* antiterminator of coliphage lambda. *Science* **228**, 91-3.

Platt, T. and Richardson, J.P. (1992) Escherichia coli rho factor: protein and enzyme of transcription termination. In *Transcriptional Regulation*. McKnight, S. L. and Yamamoto, K. R. (eds.), CSHL Press, Cold Spring Harbour, pp. 129-176.

Prasad, I. and Schaefler, S. (1974) Regulation of the beta-glucoside system in *Escherichia coli* K-12. *J Bacteriol* **120**, 638-50.

Puglisi, J.D., Chen, L., Blanchard, S. and Frankel, A.D. (1995) Solution structure of a bovine immunodeficiency virus Tat-TAR peptide- RNA complex. *Science* **270**, 1200-3.

Rao, L., Ross, W., Appleman, J.A., Gaal, T., Leirmo, S., Schlax, P.J., Record, M.T., Jr. and Gourse, R.L. (1994) Factor independent activation of *rrnB* P1. An "extended" promoter with an upstream element that dramatically increases promoter strength. *J Mol Biol* **235**, 1421-35.

Rees, W.A., Weitzel, S.E., Yager, T.D., Das, A. and von Hippel, P.H. (1996) Bacteriophage lambda N protein alone can induce transcription antitermination *in vitro*. *Proc Natl Acad Sci U S A* **93**, 342-6.

Richardson, J. P. and Greenblatt, J. F. (1996) Control of RNA chain elongation and termination. In *Escherichia coli* and *Salmonella typhimurium* . ASM Press, Washington, pp 822-48

Richardson, L.V. and Richardson, J.P. (1996) Rho-dependent termination of transcription is governed primarily by the upstream Rho utilization (*rut*) sequences of a terminator. *J Biol Chem* **271**, 21597-603.

Ring, B.Z. and Roberts, J.W. (1994) Function of a nontranscribed DNA strand site in transcription elongation. *Cell* **78**, 317-24.

Ring, B.Z., Yarnell, W.S. and Roberts, J.W. (1996) Function of *E. coli* RNA polymerase sigma factor sigma 70 in promoter- proximal pausing. *Cell* **86**, 485-93.

Ringquist, S., Jones, T., Snyder, E.E., Gibson, T., Boni, I. and Gold, L. (1995) High-affinity RNA ligands to *Escherichia coli* ribosomes and ribosomal protein S1: comparison of natural and unnatural binding sites. *Biochemistry* **34**: 3640-8.

Robert, J., Sloan, S.B., Weisberg, R.A., Gottesman, M.E., Robledo, R. and Harbrecht, D. (1987) The remarkable specificity of a new transcription termination factor suggests that the mechanisms of termination and antitermination are similar. *Cell* **51**, 483-92.

Roberts, J.W. (1969) Termination factor for RNA synthesis. *Nature* **224**, 1168-1174.

Robledo, R., Atkinson, B.L. and Gottesman, M.E. (1991) *Escherichia coli* mutations that block transcription termination by phage HK022 Nun protein. *J Mol Biol* **220**, 613-9.

Robledo, R., Gottesman, M.E. and Weisberg, R.A. (1990) Lambda *nutR* mutations convert HK022 Nun protein from a transcription termination factor to a suppressor of termination. *J Mol Biol* **212**, 635-43.

Rosenberg, M., Court, D., Shimatake, H., Brady, C. and Wulff, D.L. (1978) The relationship between function and DNA sequence in an intercistronic regulatory region in phage lambda. *Nature* **272**, 414-23.

Ross, W., Gosink, K.K., Salomon, J., Igarashi, K., Zou, C., Ishihama, A., Severinov, K. and Gourse, R.L. (1993) A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science* **262**, 1407-13.

Ross, W., Thompson, J.F., Newlands, J.T. and Gourse, R.L. (1990) *E.coli* Fis protein activates ribosomal RNA transcription *in vitro* and *in vivo*. *Embo J* **9**, 3733-42.

Rost, B. and Sander, C. (1993) Prediction of protein secondary structure at better than 70% accuracy. *J Mol Biol* **232**, 584-599.

Rost, B. and Sander, C. (1994) Combining evolutionary information and neural networks to predict protein secondary structure. *Proteins* **19**, 55-77.

Ruteshouser, E.C. and Richardson, J.P. (1989) Identification and characterization of transcription termination sites in the *Escherichia coli lacZ* gene. *J Mol Biol* **208**, 23-43.

Salstrom, J.S. and Szybalski, W. (1978) Coliphage lambda *nutL*$^-$: a unique class of mutants defective in the site of gene N product utilization for antitermination of leftward transcription. *J Mol Biol* **124**, 195-221.

Schauer, A.T., Carver, D.L., Bigelow, B., Baron, L.S. and Friedman, D.I. (1987) lambda N antitermination system: functional analysis of phage interactions with the host NusA protein. *J Mol Biol* **194**, 679-90.

Schauer, A. T., Cheng, S. W., Zheng, C., St. Pierre, L., Alessi, D., Hidayetoglu, D. L., Costantino, N., Court, D. L. and Friedman, D. I. (1996) The alpha subunit of RNA polymerase and transcription antitermination. *Mol Micro* **21**, 839-51.

Schickor, P., Metzger, W., Werel, W., Lederer, H. and Heumann, H. (1990) Topography of intermediates in transcription initiation of *E.coli*. *Embo J* **9**, 2215-20.

Schmidt, M.C. and Chamberlin, M.J. (1984) Amplification and isolation of *Escherichia coli* NusA protein and studies of its effects on *in vitro* RNA chain elongation. *Biochemistry* **23**, 197-203.

Schmidt, M.C. and Chamberlin, M.J. (1987) NusA protein of *Escherichia coli* is an efficient transcription termination factor for certain terminator sites. *J Mol Biol* **195**, 809-18.

Schnetz, K. and Rak, B. (1990) Beta-glucoside permease represses the *bgl* operon of *Escherichia coli* by phosphorylation of the antiterminator protein and also interacts with glucose-specific enzyme III, the key element in catabolite control. *Proc Natl Acad Sci U S A* **87**, 5074-8.

Schnetz, K., Toloczyki, C. and Rak, B. (1987) Beta-glucoside (*bgl*) operon of *Escherichia coli* K-12: nucleotide sequence, genetic organization, and possible evolutionary relationship to regulatory components of two *Bacillus subtilis* genes. *J Bacteriol* **169**, 2579-90.

Severinova, E., Severinov, K. and Darst, S.A. (1998) Inhibition of *Escherichia coli* RNA polymerase by bacteriophage T4 AsiA. *J Mol Biol* **279**, 9-18.

Sharrock, R.A., Gourse, R.L. and Nomura, M. (1985a) Defective antitermination of rRNA transcription and derepression of rRNA and tRNA synthesis in the *nusB5* mutant of *Escherichia coli*. *Proc Natl Acad Sci U S A* **82**, 5275-9.

Sharrock, R.A., Gourse, R.L. and Nomura, M. (1985b) Inhibitory effect of high-level transcription of the bacteriophage lambda *nutL* region on transcription of rRNA in *Escherichia coli*. *J Bacteriol* **163**, 704-8.

Shih, M.C. and Gussin, G.N. (1984) Role of cII protein in stimulating transcription initiation at the lambda PRE promoter. Enhanced formation and stabilization of open complexes. *J Mol Biol* **172**, 489-506.

Siegele, D.A., Hu, J.C., Walter, W.A. and Gross, C.A. (1989) Altered promoter recognition by mutant forms of the sigma 70 subunit of *Escherichia coli* RNA polymerase. *J Mol Biol* **206**, 591-603.

Simon, L.D., Gottesman, M., Tomczak, K. and Gottesman, S. (1979) Hyperdegradation of proteins in *Escherichia coli rho* mutants. *Proc Natl Acad Sci U S A* **76**, 1623-7.

Simpson, R.B. (1979) The molecular topography of RNA polymerase-promoter interaction. *Cell* **18**, 277-85.

Siomi, H., Choi, M., Siomi, M.C., Nussbaum, R.L. and Dreyfuss, G. (1994) Essential role for KH domains in RNA binding: impaired RNA binding by a mutation in the KH domain of FMR1 that causes fragile X syndrome. *Cell* **77**, 33-9.

Siomi, H., Matunis, M.J., Michael, W.M. and Dreyfuss, G. (1993) The pre-mRNA binding K protein contains a novel evolutionarily conserved motif. *Nucleic Acids Res* **21**, 1193-8.

Sparkowski, J. and Das, A. (1990) The nucleotide sequence of *greA*, a suppressor gene that restores growth of an *Escherichia coli* RNA polymerase mutant at high temperature. *Nucleic Acids Res* **18**, 6443.

Squires, C.L., Greenblatt, J., Li, J. and Condon, C. (1993) Ribosomal RNA antitermination *in vitro*: requirement for Nus factors and one or more unidentified cellular components. *Proc Natl Acad Sci U S A* **90**, 970-4.

Straney, D.C. and Crothers, D.M. (1985) Intermediates in transcription initiation from the *E. coli* lac UV5 promoter. *Cell* **43**, 449-59.

Su, L., Radek, J.T., Hallenga, K., Hermanto, P., Chan, G., Labeots, L.A. and Weiss, M.A. (1997a) RNA recognition by a bent alpha-helix regulates transcriptional antitermination in phage lambda. *Biochemistry* **36**, 12722-32.

Su, L., Radek, J.T., Labeots, L.A., Hallenga, K., Hermanto, P., Chen, H., Nakagawa, S., Zhao, M., Kates, S. and Weiss, M.A. (1997b) An RNA enhancer in a phage

transcriptional antitermination complex functions as a structural switch. *Genes Dev* **11**, 2214-26.

Subramanian, A.R. (1983) Structure and functions of ribosomal protein S1. *Prog Nucleic Acid Res Mol Biol* **28**, 101-42.

Sullivan, S.L. and Gottesman, M.E. (1992) Requirement for *E. coli* NusG protein in factor-dependent transcription termination. *Cell* **68**, 989-94.

Sullivan, S.L., Ward, D.F. and Gottesman, M.E. (1992) Effect of *Escherichia coli nusG* function on lambda N-mediated transcription antitermination. *J Bacteriol* **174**, 1339-44.

Swindle, J., Zylicz, M., Georgopoulos, C., Li, J. and Greenblatt, J. (1988) Purification and properties of the NusB protein of *Escherichia coli*. *J Biol Chem* **263**, 10229-35.

Tan, R. and Frankel, A.D. (1995) Structural variety of arginine-rich RNA-binding peptides. *Proc Natl Acad Sci U S A* **92**, 5282-6.

Taura, T., Ueguchi, C., Shiba, K. and Ito, K. (1992) Insertional disruption of the *nusB* (*ssyB*) gene leads to cold-sensitive growth of *Escherichia coli* and suppression of the *secY24* mutation. *Mol Gen Genet* **234**, 429-32.

Tsugawa, A., Kurihara, T., Zuber, M., Court, D.L. and Nakamura, Y. (1985) *E. coli* NusA protein binds *in vitro* to an RNA sequence immediately upstream of the *boxA* signal of bacteriophage lambda. *Embo J* **4**, 2337-42.

Tsugawa, A., Saito, M., Court, D.L. and Nakamura, Y. (1988) *nusA* amber mutation that causes temperature-sensitive growth of *Escherichia coli*. *J Bacteriol* **170**, 908-15.

Van Gilst, M.R., Rees, W.A., Das, A. and von Hippel, P.H. (1997) Complexes of N antitermination protein of phage lambda with specific and nonspecific RNA target sites on the nascent transcript. *Biochemistry* **36**, 1514-24.

Vogel, U. and Jensen, K.F. (1997) NusA is required for ribosomal antitermination and for modulation of the transcription elongation rate of both antiterminated RNA and mRNA. *J Biol Chem* **272**, 12265-71.

von Hippel, P.H. and Yager, T.D. (1991) Transcript elongation and termination are competitive kinetic processes. *Proc Natl Acad Sci U S A* **88**, 2307-11.

Wang, D., Meier, T.I., Chan, C.L., Feng, G., Lee, D.N. and Landick, R. (1995) Discontinuous movements of DNA and RNA in RNA polymerase accompany formation of a paused transcription complex. *Cell* **81**, 341-50.

Wang, D., Severinov, K. and Landick, R. (1997) Preferential interaction of the his pause RNA hairpin with RNA polymerase beta subunit residues 904-950 correlates with strong transcriptional pausing. *Proc Natl Acad Sci U S A* **94**, 8433-8.

Ward, D.F. and Gottesman, M.E. (1981) The *nus* mutations affect transcription termination in *Escherichia coli*. *Nature* **292**, 212-5.

Warren, F. and Das, A. (1984) Formation of termination-resistant transcription complex at phage lambda *nut* locus: effects of altered translation and a ribosomal mutation. *Proc Natl Acad Sci U S A* **81**, 3612-6.

Whalen, W., Ghosh, B. and Das, A. (1988) NusA protein is necessary and sufficient *in vitro* for phage lambda N gene product to suppress a rho-independent terminator placed downstream of *nutL*. *Proc Natl Acad Sci U S A* **85**, 2494-8.

Whalen, W.A. and Das, A. (1990) Action of an RNA site at a distance: role of the *nut* genetic signal in transcription antitermination by phage-lambda N gene product. *New Biol* **2**, 975-91.

Yager, T.D. and von Hippel, P. (1987) Transcription elongation and termination in *Escherichia coli*. In *Escherichia coli and Salmonella typhimurium. Cellular and Molecular Biology*. ASM Press Washington, pp. 1241-1275.

Yager, T.D. and von Hippel, P.H. (1991) A thermodynamic analysis of RNA transcript elongation and termination in *Escherichia coli*. *Biochemistry* **30**, 1097-118.

Yang, X.J., Hart, C.M., Grayhack, E.J. and Roberts, J.W. (1987) Transcription antitermination by phage lambda gene Q protein requires a DNA segment spanning the RNA start site. *Genes Dev* **1**, 217-26.

Yang, X.J. and Roberts, J.W. (1989) Gene *Q* antiterminator proteins of *Escherichia coli* phages 82 and lambda suppress pausing by RNA polymerase at a rho-dependent terminator and at other sites. *Proc Natl Acad Sci U S A* **86**, 5301-5.

Yarnell, W.S. and Roberts, J.W. (1992) The phage lambda gene *Q* transcription antiterminator binds DNA in the late gene promoter as it modifies RNA polymerase. *Cell* **69**, 1181-9.

Yarnell, W.S. and Roberts, J.W. (1999) Mechanism of intrinsic transcription termination and antitermination. *Science* **284**, 611-5.

Zaychikov, E., Martin, E., Denissova, L., Kozlov, M., Markovtsov, V., Kashlev, M., Heumann, H., Nikiforov, V., Goldfarb, A. and Mustaev, A. (1996) Mapping of catalytic residues in the RNA polymerase active center. *Science* **273**, 107-9.

Zhang, O., Kay, L.E., Olivier, J.P., and Forman-Kay, J.D. (1994) Backbone ^{1}H and ^{15}N resonance assignments of the N-terminal SH3 domain of *drk* in folded and unfolded

states using enhanced-sensitivity pulsed field gradient NMR techniques. *J Biomol NMR* **4**, 845-58.

Zheng, C. and Friedman, D.I. (1994) Reduced Rho-dependent transcription termination permits NusA- independent growth of *Escherichia coli*. *Proc Natl Acad Sci U S A* **91**, 7543-7.

Zillig, W., Zechel, K., Rabussay, D., Schachner, M., Sethi, V.S., Palm, P., Heil, A. and Seifert, W. (1970) On the role of different subunits of DNA-dependent RNA polymerase from E.coli in the transcription process. In *Cold Spring Harbour Symposium Quant. Biol.*, Cold Spring Harbour, Vol. 35, pp. 47-58.

Zuber, M., Patterson, T.A. and Court, D.L. (1987) Analysis of *nutR*, a site required for transcription antitermination in phage lambda. *Proc Natl Acad Sci U S A* **84**, 4514-8.