Studies on the Mechanism of Transcription Antitermination by N Protein from Bacteriophage H-19B

Thesis submitted to

Manipal University

for the Degree of

Doctor of Philosophy

By

Anoop Cheeran
Registration number: 040100014

Centre for DNA Fingerprinting and Diagnostics
Hyderabad 500076

May 2007

DECLARATION

The research work embodied in this thesis entitled “Studies on the Mechanism of Transcription Antitermination by N protein from Bacteriophage H-19B”, has been carried out by me at the Centre for DNA Fingerprinting and Diagnostics, Hyderabad, under the guidance of Dr. Ranjan Sen. I hereby declare that this work is original and has not been submitted in part or full for any degree or diploma of any other university or institution.

Anoop Cheeran
Centre for DNA Fingerprinting
and Diagnostics, Hyderabad.

CERTIFICATE

This is to certify that this thesis entitled “Studies on the Mechanism of Transcription Antitermination by N protein from Bacteriophage H-19B”, submitted by Mr. Anoop Cheeran for the Degree of Doctor of Philosophy to Manipal University is based on the work carried out by him at the Centre for DNA Fingerprinting and Diagnostics, Hyderabad. This work is original and has not been submitted in part or full for any degree or diploma of any other university or institution.

Dr. Ranjan Sen
Thesis Supervisor
Centre for DNA Fingerprinting
and Diagnostics, Hyderabad.

Dr. Shekhar C Mande
Dean, Academic Affairs
Centre for DNA Fingerprinting
and Diagnostics, Hyderabad.

Table of Contents

Chapter II- Objectives

Chapter III- Screening for H-19B N specific RNA polymerase mutants, their suppressors in N and characterization of the mutants

Chapter IV- Biochemical probing of the binding surface of H-19B N on RNA polymerase

Chapter V- Mechanism of Transcription Antitermination

Chapter VI- Conclusions and Future perspectives

Chapter VII- References

Acknowledgements

This thesis is the result of work spanning over four and a half years, an exciting period of my life, full of new challenges and constant learning. I was ably accompanied and assisted by a whole lot of people in this long journey and it's a pleasant opportunity to express my gratitude to all of them.

Words will not express my heartfelt thanks to Dr. Ranjan Sen, my thesis advisor and mentor. He provided a motivating, enthusiastic and critical atmosphere through out the period of my work. It has been a pleasure and a privilege to conduct research under his able guidance. I am also indebted to him for guiding me into the exciting world of basic biology and its challenges.

The work in this thesis was carried out at Center for DNA Fingerprinting and Diagnostics (CDFD), Hyderabad. I am grateful to CDFD and to Drs. Seyed E Hasnain and J. Gowrishankar, who were the Directors of this institute during the different periods of my work here, for their support and constant encouragement.

I am grateful to the Vice Chancellor, Registrar and the Head, Dept of Biotechnology of the Manipal University for allowing me to register as a graduate student.

I would like to express my gratitude to Dr. J. Gowrishankar and his lab members at the Laboratory of Bacterial Genetics, especially Drs. Abhijit Sardesai and Harinarayanan for the many constructive discussions and inputs regarding my in vivo experiments and the bacteriophage stocks that they provided. Am also very grateful to Dr. Shekhar C. Mande, head Structural Biology lab and Dean, Academics section, CDFD. He has been a source of constant encouragement and support through out all this years and the many discussions that I have had with him has been very enjoyable and intellectually stimulating.

I would like to acknowledge Drs. David Freidman, Robert Weisberg, Irina Artsimovitch and Sergei Borukhov for providing us the H-19B phage, Y1090 (? Nin5) phage, *E.coli* NusA plasmid and ß S531Y HMK tagged RNAP plasmid, respectively.

I am grateful to the Council for Scientific and Industrial Research (CSIR), India for providing me with Junior and Senior research fellowships, during the period of my work. The research at our laboratory has been supported financially by research grants to Dr. Ranjan Sen from NIH, USA; Wellcome Trust, UK and CDFD. I would like to express my gratitude to all these institutions.

I would like to acknowledge Dr. H.A Nagarajaram and Sridhar for all the help they have provided us in computing the molecular distances. Sridhar has put in long hours doing this and has been a great teacher, helping me learn the basics of RASMOL. I would also like to thank the staff of sequencing facility of CDFD for putting up with impatient me, queuing up for results (often ahead of schedules) and for providing good quality sequencing data.

I have been fortunate to be a part of a very young and dynamic lab. My juniors, Bibhusita, Jisha and Ghazala have been wonderful friends and great co-workers. With all her enthusiasm and care, Bibhusita has been a great friend to have in the lab, always ensuring that there never was a dull moment. Both Bhibusita and Jisha have provided selfless help at all times. Am also indebted to Jisha for all the suggestions and help she provided in the formatting of the thesis. Deepak, who joined our lab recently, has been a good friend. I am also grateful to former lab members Sharmistha, Suganthan and Swapna for their friendship and support. I am grateful to Suganthan and Swapna for all the efforts they have put in while working with me during different points of time in this project and to Mahendra and Anagha, the summer trainees who worked along with me. I am indebted to Nanci and Gowresh for all the help provided by them in the lab, for the preparation of media, buffers, gels and for providing a constant supply of sterile plastic and glassware.

Many thanks to all my friends here at CDFD, who offered an enthusiastic environment in academia and otherwise, maintaining a healthy balance between work and play. Sudhish and Gokul have been great friends through out and along with Vaibhav and Jayendra made up the 'Pandavas, 'of the wonderful home that we have here in Hyderabad. Rohini has been as great a friend as any one could ever hope for. Her guidance through the initial periods of research, helping me understand how to cope up with small frustrations that creep up in research every now and then, has helped me enormously. Am also thankful to my present and past colleagues including Radha, Abira, Madhav, Subbu, Prabhat, Rajendra, Yashin, Sreenu, Senthil, Kauser, Sreeramana, Julie, KV, Prachee, Nitin, Akif, Yashaswini, Bharat, Shafiqque, Santhosh, Arvind, Devi, Tej, Sailu, Uma, Pankaj, Pavani, Vartul, Kshama, Debashree, Ritu and Sheeba. All of them and many more friends here at the CDFD family have enriched my life with their friendship and help. Am grateful to Dr. S. Mahalingam for the support he extended as well as for his encouraging words at all times during my stay here; and also to Drs. Sanjeev Khosla, Gayathri Ramakrishna, Murali Basyam and Sunil Manna for their constant encouragement. Punitha madam and Mr. Santhosh at the Academics section have always been very helpful. I would also like to thank Mr. Prakash, head of Accounts section and also Mr. Kashipathi, Mr. Ramayya and Mr. Neeli Sreenivas. I am also thankful to Drs. Sangita Sen Majee and Sharmila Mande, for their care and support as well as for their wonderful hospitality shown towards us graduate students, most of us far away from home and craving for some homely atmosphere.

I have not enough words to express my gratitude towards my father, mother, Anu and Biju. Appa and Amma have always put our needs and aspirations ahead of theirs and have ever been supportive and encouraging. It is their support and prayers that has helped me to accomplish whatever little I have so far achieved.

Anoop

Abbreviations

$A_{600\ nm}$	:	Absorbance at 600 nm
Å	:	Angstrom
a.a.	:	Amino acid
ADP	:	Adenosine tri phosphate
APS	:	Ammonium-per-Sulphate
ApU	:	Adenyly- (3', 5') Uridine
ATP	:	Adenosine tri phosphate
bp	:	Base pair
BPB	:	Bromophenol Blue
BSA	:	Bovine Serum Albumin
CBB	:	Coomassie brilliant blue
CnBr	:	Cyanogen Bromide
C-terminal	:	Carboxy terminal
CTD	:	C Terminal Deletant
CTP	:	Cytosine Tri Phosphate
DMF	:	DiMethyl Formamide
DNA	:	Deoxyribonucleic acid
dNTP	:	Deoxynucleotide
DTNB	:	5, 5-dithiobis (2-nitrobenzoic acid)
DTT	:	Di thiothreitol
EC	:	Elongation Complex
E. coli	:	*Escherichia coli*
EDTA	:	Ethylene diamine tetra acetic acid
EtBr	:	Ethidium Bromide
BABE	:	p-Bromoacetamidobenzyl-EDTA
GTP	:	Guanosine Tri Phosphate
IPTG	:	Isopropyl β-D thiogalactopyranoside
kb	:	Kilo base
KD	:	Kilo Dalton
L	:	Litre
LB	:	Luria Broth
M	:	Molar
MES	:	4-Morpholineethanesulfonic acid
mg	:	Milligram
µg/ml	:	Microgram per millilitre
µl	:	Microlitre
min	:	Minute
mg/ml	:	Milligram per millilitre
mM	:	MilliMolar
µM	:	Micromolar
MNNG	:	*N*-methyl-*N'*-nitro-*N*-nitrosoguanidine
MOI	:	Multiplicity of infection
Nm	:	Nano metre
nt	:	Nucleotide
NTCB	:	2-Nitro-5-thiocyanobenzoic acid
N-terminal	:	Amino terminal
NTP	:	Nucleoside triphosphate

OD	:	Optical density
O/N	:	Overnight
ONPG	:	O-Nitro phenyl-ß-D-Galactosidase
PAGE	:	Polyacrylamide Gel Electrophoresis
PCR	:	Polymerase Chain Reaction
pfu	:	Plaque forming units
P_i	:	Inorganic phosphate
PKA	:	Protein Kinase A
PMSF	:	Phenyl methyl sulphonyl fluoride
Ppi	:	Pyrophosphate
RB	:	Road Block
RBS	:	Ribosome binding site
rpm	:	Revolution per minute
RNA	:	Ribonucleic Acid
RNAP	:	RNA polymerase
rpm	:	Revolution per minute
SDS	:	Sodium Dodecyl Sulphate
s, sec	:	Seconds
TAE	:	Tris-acetate EDTA
TE	:	Tris-EDTA
TEMED	:	N,N,N',N'-Tetramethylethylenediamine
TEC	:	Ternary Elongation Complex
Tris	:	2-amino-2-(hydroxymethyl) propane-1,3-diol
Tris-Cl	:	Tris Chloride buffer
Tris-Glu	:	Tris Glutamate buffer
ts	:	Temperature sensitive
TTP	:	Thymidine tri phosphate
UTP	:	Uracil tri phosphate
UV	:	Ultra violate
Wt	:	Wild type
X Gal	:	5-Bromo 4-chloro 3-idolyl ß-D-galactose

Synopsis

Transcriptional antitermination is the process in which genes whose transcription is otherwise blocked by premature termination are expressed through termination suppression. Antiterminators like N in ? and in related lambdoid phages convert RNA polymerase into a termination resistant form, during early phases of transcription elongation. N protein binds to RNA polymerase (RNAP) and together with host factors called Nus factors, transforms it into a highly processive, termination resistant transcription apparatus. The N antitermination protein of phage H-19B, a lambdoid phage, requires only NusA for antitermination studies (Neely and Friedman, 2000). This makes it is a simpler and more attractive system than the N protein of phage ?, which requires additional host factors, to study the mechanism of antitermination. A functional antitermination system of H-19B can be reconstituted *in vitro* with a DNA template having a *nut* site, N and NusA proteins. H-19B N is a small 127 amino acid basic protein. The amino terminus of N protein contains an Arginine Rich Motif (ARM) that binds to a small stem loop structure in the *nut* RNA, known as BOXB and during transcription N binds on to it and gets tethered to the elongating RNA polymerase, converting the RNAP to a termination resistant form. The stable association of RNA polymerase to N-BOXB and to NusA modifies the ternary elongation complex (TEC) in such a way that it continues transcription through downstream terminators. The determination of the binding surface of N on RNAP is essential to understand the molecular mechanism of transcription antitermination. This work is devoted to the determination of binding surface of H-19B N on RNA polymerase and to understand the mechanism of transcription antitermination.

Recent crystal structures and structural models of free RNAP and the EC reveals that major RNAP-nucleic acid interactions of EC occur in the ß (*rpoB*) and ß' (*rpoC*) subunits (Zhang et.al 1999; Korzheva et al., 2000; Gnatt *et.al* 2001; Korzheva et al., 2001). We hypothesized that an antitermination process may involve alterations of RNAP-nucleic acid interactions and that the likely target of the antiterminator N protein could be in these two major subunits of RNAP. In Chapter III, we attempted the random mutagenesis of the ß and ß' subunits of RNAP and the subsequent screening and characterization of RNAP mutants with N-specific defects. We also attempted to find allele specific suppressors in H-19B N for the *rpoB/C* mutants that have been obtained and the characterization of these suppressor(s), both *in vivo* and *in vitro*. Chapter III gives a detailed account of the purification and *in vitro* characterization of viable RNA polymerase mutants with H-19B N specific mutations in the ß and ß' subunits and the localization

of these substitutions in a molecular model of TEC. Chapter IV describes the biochemical studies for the determination of binding surface of H-19B N on a stalled EC. This included footprinting of a functional 'roadblocked EC', which was modified/unmodified with H-19B N. Limited proteolysis by Trypsin, localized Fe^{2+} induced footprinting for active site conformational changes, Fe-EDTA protein footprinting and footprinting using a chemical nuclease called Fe-BABE (p-bromoacetamidobenzyl-EDTA), specifically conjugated to the single Cysteine residue at the C-terminus of H-19B N protein was carried out for identifying the N binding surface on the EC. We tried to shed light on the mechanism of N mediated transcription antitermination by various means, which are described in Chapter V. This included experiments to see if N protein does come near the active site in an N modified EC and whether the N modified EC can over come effects of Streptolydigin and Tagetoxin, two antibiotics whose binding sites are close to the active site (Temiakov et al., 2005; Tuske et al., 2005; Vassylyev et al., 2005). We assayed the effects of N modification on the stability of an elongation complex, stalled at a terminator. We also addressed the state of RNA hairpin in an N modified EC, to check if N modification disrupts hairpin formation or not. Assays using Class I (*his* pause) and Class II (*ops* pause) pause sequences (Artsimovitch and Landick, 2000) were carried out on appropriate DNA templates to see the effect of N modification on the paused EC, if any. We checked if N has any effect on the backtracking of RNAP on such paused templates and the effect of N modification on backtracked TEC complexes by checking for its sensitivity to GreB and PPi. Chapter VI deals with the major conclusions that we arrived at and perspectives for future, based on the insights obtained from this study.

The major results that we have achieved through this study are as follows. We obtained the following mutants in RNAP, which are specifically defective for N mediated transcription antitermination both *in vivo* and *in vitro*- P251S, P254L, R270C and G336S in the *rpoC* (ß') subunit, and G1045D in the *rpoB* (ß) subunit. We also obtained a suppressor in H-19B N, L108F (located in the RNAP binding C terminus) for P254L RNAP mutant. In the presence of this suppressor N protein, antitermination efficiency of ß' P254L mutant RNAP increased to the level of antitermination efficiency of Wt RNAP obtained in the presence of Wt N protein. This mutant N protein could also partially suppress the antitermination defects of P251S, P251S P254L double mutant and G336S *rpoC* mutants, where as it had no significant effects on R270C *rpoC* and G1045D *rpoB* mutant. Specific defects for N-mediated antitermination, increased concentration dependence for N, lack of processive antitermination of the mutant RNAPs; and the suppressor mutation in the RNAP binding domain of N (as discussed in Chapter III), together suggest that H-19B N exerts its effect on EC through the regions defined by the altered amino acids. These

mutants define a very crucial region of RNAP, which is involved in the interactions important for the stabilization of the EC (Darst, 2001; Korzheva et al., 2000). The footprinting results revealed that the C-terminal domain (the RNAP binding domain) of N from a lambdoid phage H-19B comes close to the active center of the EC. The cleavage pattern derived from Fe-BABE conjugated N bound to a stalled EC showed that the interaction surface must be close to the active site Mg^{2+} which was further corroborated by the N-induced localized conformation changes nearby to this region of EC. Our results also showed that H-19B N modified EC is significantly stable even when stalled at a terminator and that the folding of terminator hairpin in this modified EC is affected. In the presence of H-19B N, it is likely that proper formation of terminator hairpin is impaired and this in turn contributes to the stability of the EC at the terminator. The altered interactions at the beginning of the RNA exit channel and upstream part of RNA:DNA hybrid can directly prevent the formation of the base of the hairpin. Alternatively, this altered interactions in the modified EC can stabilize the RNA:DNA hybrid that will also prevent the completion of the hairpin folding as the melting of the hybrid is essential for hairpin to form (Komissarova et al., 2002). The concept that H-19B N affects the proper formation of terminator hairpin is consistent with earlier proposal of prevention or delayed folding of hairpin in presence of λ N (Gusarov and Nudler, 2001). Taking into account of all these results, we propose that the physical proximity of the C-terminal domain of N near the active site Mg^{2+} directly stabilizes the 3'-end of the weak U-rich RNA:DNA hybrid; and that the allosteric effects that this binding brings about destabilizes the RNA hairpin interactions in the distal RNA exit channel, preventing the RNAP pause at the U tract preceding the termination step; N could also prevent the proper formation of a terminator hairpin, thus enabling the TEC to go into a processive antitermination mode.

Publications

1. *Escherichia coli* RNA polymerase mutations located near the upstream edge of an RNA:DNA hybrid and the beginning of the RNA-exit channel are defective for transcription antitermination by the N protein from lambdoid phage H-19B. **Cheeran, A.,** Babu Suganthan, R., Swapna, G., Bandey, I., Achary, M. S., Nagarajaram, H. A., and Sen, R. (2005). **J Mol Biol** *352*, 28-43.

2. Site of action of the antiterminator N is close to the active site of RNA polymerase. **Cheeran, A., Kolli**, R.N., and Sen, R. (2007). *Manuscript submitted.*

CHAPTER I

Review of Literature

1.1 Transcription

Transcription is the process of synthesizing an RNA that is complimentary to the template strand of DNA and the information in this RNA (messenger RNA or mRNA) is in turn translated so as to assemble polypeptides, which are the building blocks of life. RNA Polymerase is the molecular machine that brings about transcription and is also a target of different regulatory mechanisms that control the three major stages of transcription: initiation, elongation and termination. Ternary Elongation Complexes (TEC) formed by the association of a multisubunit RNAP (such as one from *E. coli* comprising of a_2 ß ß' and ? subunits), DNA, nascent RNA and host elongation factors is an exceptionally stable structure and is usually resistant to dissociation under a variety of conditions and accounts for the high processivity of transcription (Arndt and Chamberlin, 1990).

1.1.1 Transcription cycle

The transcription cycle in bacterial cells can be divided into three main stages: Initiation, Elongation and Termination. Though catalytically active, the core enzyme is incapable of initiating transcription efficiently and with specificity. It must bind with an initiation factor called s, to form a holoenzyme that can recognize specific DNA sequences called promoters (Gross et al, 1992). During initiation, the holoenzyme specifically binds to two conserved hexamers in the promoter DNA at nucleotide positions ¯35 and ¯10 relative to the transcription start site (+1), to form a closed promoter complex. Then, it unwinds the double-stranded DNA (ds DNA) around the ¯10 region, forming the open promoter complex and initiates transcription in the presence of nucleoside triphosphate substrates (Record et al., 1996). After the synthesis of an RNA 9–12 nucleotides (nt) long, of which 8 or 9 are base-paired with the DNA template strand (DNA–RNA hybrid), the transcription complex passes from the initiation to elongation stage (von Hippel, 1998). Once the RNA chain becomes 10 nt or so, s subunit is released and the core RNAP elongates the RNA along the DNA template. The final step in transcript formation is termination, which occurs when the elongating complex moves into one or more terminator sites along DNA template that may serve as transcription regulators within genes or mark the end of a gene or an operon (von Hippel, 1998), where RNAP releases the RNA transcript and dissociates from the DNA. Transcription cycle in prokaryotes involves various phases, involving distinct TEC conformations; the combination of this conformations and the action of regulatory ligands (Intrinsic like discrete sequences of DNA like the promoters which direct transcription of adjacent segments of genes and DNA sequences that brings about the formation of secondary structures like stem loops in the nascent RNA forming a transcription terminator; and Extrinsic

like NTPs, host elongation proteins and transcript cleavage factors like GreA and B) confers enormous regulatory flexibility on the process of transcription (Mooney et al., 1998). The various phases of transcription are represented in the following cartoon.

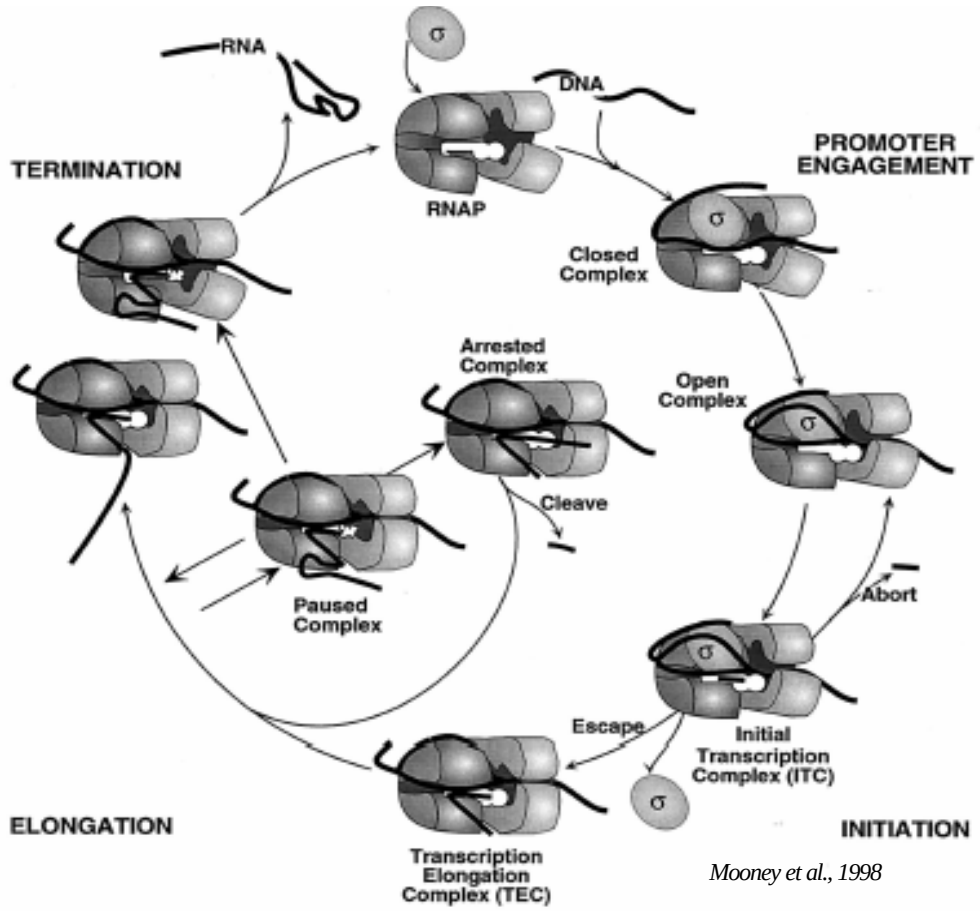


Figure 1.1: Phases of transcription
Diagrammatic representation of different stages of transcription like promoter engagement, promoter escape, transcript elongation and termination and the RNAP conformations in TEC. Intermediates in each step like abortive transcription before the formation of processive TEC, a paused complex as an intermediate before transcription arrest or termination are also shown.

1.2 The RNA polymerase

1.2.1 RNA Polymerase subunits and their roles

The DNA-dependent RNA polymerase (RNAP) is the principal enzyme of the transcription process, and is a target in many regulatory pathways that control gene expression in all living organisms. In bacteria, RNAP is responsible for the synthesis of all the RNAs (messenger RNA, ribosomal RNA, transfer RNA) in the cell. Bacterial RNAP exists in two forms: core and holoenzyme. The core enzyme has a molecular mass of about 400 KD and consists of five subunits: α dimer (α1, α2), ß, ß' and ω; *E. coli* RNA polymerase contains two large subunits, beta (ß) and beta prime (ß') with molecular masses around 150 and 155 KD respectively. The other subunits involve an alpha (a) dimer with molecular mass of 40 KD each and the ? subunit with a molecular weight of around 11 KD. Addition of another subunit called s with a molecular weight of 70 KD to the core RNAP forms the RNAP holoenzyme. The core enzyme contains the basic transcription machinery and the σ factor directs the core to transcribe specific genes, by tight binding between the RNAP holoenzyme and the promoter DNA.

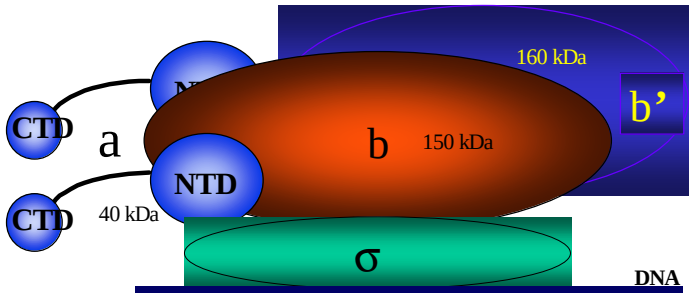


Figure 1.2: Cartoon representing a schematic representation of prokaryotic RNA Polymerase (Holo enzyme) structure The different subunits, ß, ß', α1, α2and σ subunits are colored red, dark blue, light blue and green respectively. ? subunit is not seen in this image.

The different subunits of RNAP are to a major part evolutionarily conserved in sequence, structure and function among bacteria. (Sweetser et al., 1987; Naryshkin et al., 2000; Figure 1.2). RNA polymerase from bacteria to humans contain an essential and highly conserved core of two large subunits called ß and ß' in bacteria and RPB1 and RPB2 in eukaryotic RNA polymerase II), and smaller subunits (a dimer and ? in bacteria; RPB3/11 and RPB 6 in RNA polymerase II) that stabilize association of the two large subunits. The ß and ß' subunits between themselves are the two largest subunits of *Escherichia coli* RNA polymerase and are homologous to the largest and

second largest subunits, respectively, of all multi subunit RNA polymerases (Allison et al., 1985; Biggs et al., 1985; Hudson et al., 1988; Leffers et al., 1989; Patel and Pickup, 1989).

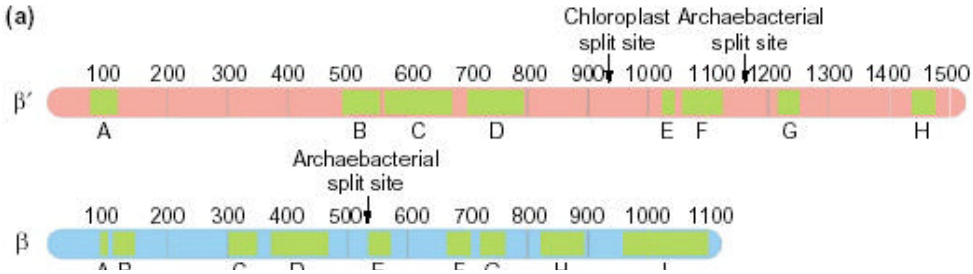


Figure 1.3: Evolutionarily conserved segments of the RNAP large subunits form the structural core.

The bars schematically represent the primary sequences of the *T. aquaticus* RNAP ß' (top, pink) and ß (bottom, blue) subunits, with amino acid numbering directly above the bars. The split sites of the large subunits, which occur as two separate peptides in chloroplast and archaebacterial RNAPs, are also indicated. Evolutionarily conserved regions among prokaryotes, chloroplasts, archaebacteria and eukaryotes are denoted as light green boxes and are labeled A–H for ß' A–I

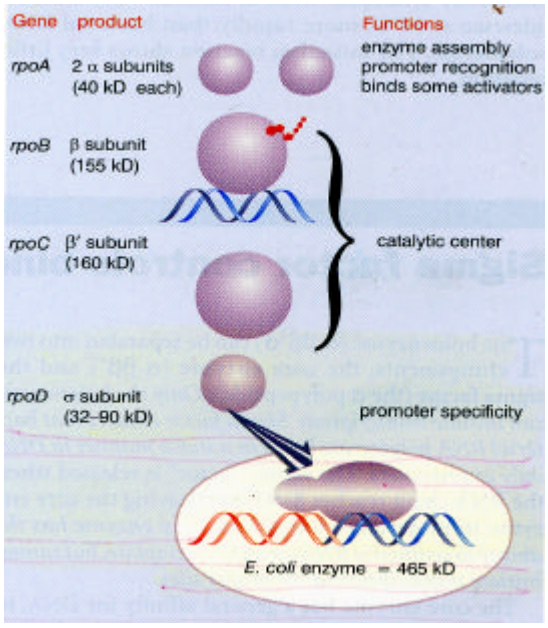


Figure 1.4: Cartoon representing the constituents of Eubacterial RNA polymerase and their functions

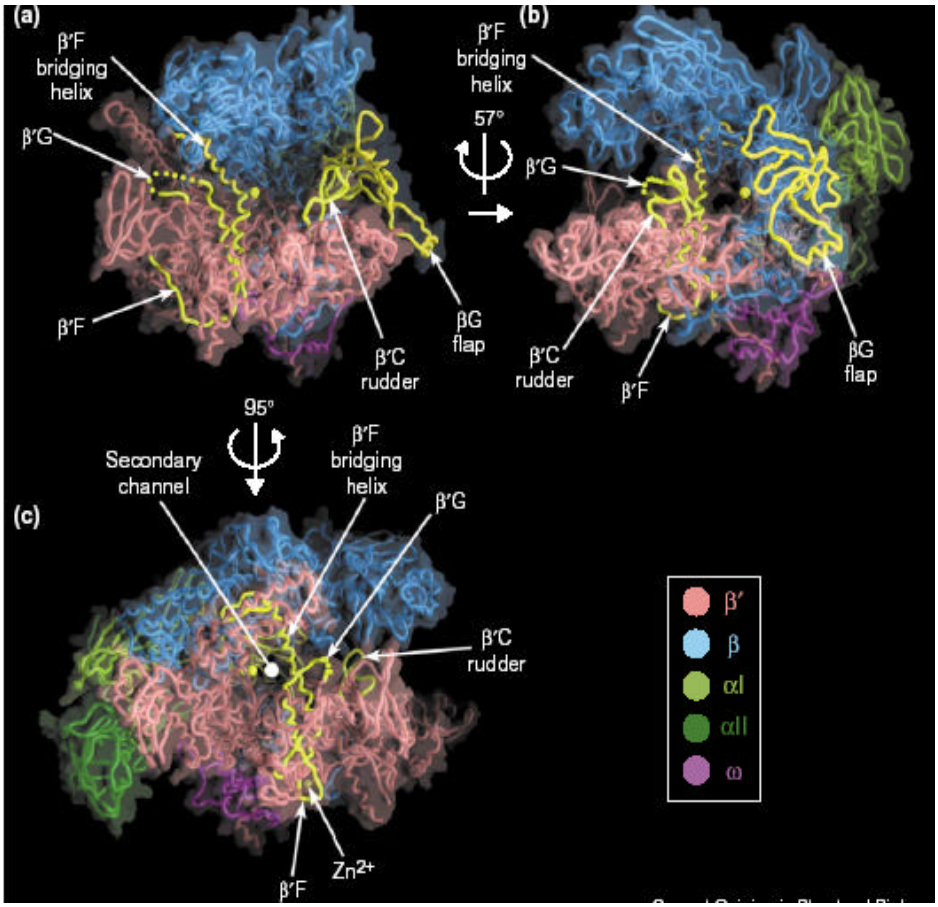
The genes encoding each of the subunits, their approximate molecular weights and correspomding functions are shown

Genes VII, Benjamin Lewin; Oxford University Press and Cell press, 2000.

The ß and ß' subunits are involved in DNA binding and ß subunit binds nucleotides at the active site of the RNAP where phosphodiester bonds are formed. Most catalytic and regulatory functions of RNA polymerase and the major RNA-Nucleic acid interactions occur in ß and ß' subunits (Gnatt et al., 2001; Korzheva et al., 2000; Nudler et al., 1998; Zhang et al., 1999). The alpha (α) subunit of RNA polymerase is critical for assembly of the Polymerase, for the positive control of initiation in *E.coli*, in promoter recognition and also plays a role in the interaction of RNAP with regulatory factors. The alpha subunit has two domains, a large N-terminal domain and an independently folded C-terminal domain that can recognize and bind to a promoter's UP element, which allows tight binding between polymerase and promoter (Figure 1.2) The α NTD serves to assemble the RNAP holoenzyme. The points of contact between the α NTD and the ß and ß' subunits lies in the opposite side of the α NTD from the dimer's C terminal domains. The Omega (?) subunit (coded by the *rpoZ* gene) promotes RNAP assembly and increases RNAP stability, through specific interactions with the ß' subunit of RNAP. The ? subunit function in RNAP assembly by "latching" the N- and C-terminal regions of the RNAP ß' subunit, thereby facilitating association of the RNAP largest subunit with the ß- α1- α2 subunit assembly intermediate in bacterial RNAP core enzyme (Minakhin et al., 2001). The Sigma (σ) subunit is concerned specifically with promoter recognition. There are various classes of sigma factors found in bacteria, each specific for a different class of promoter. Changes in the Sigma factors having different and distinct promoter specificities regulating discrete sets of genes appear in cases when a reorganization of transcription patterns is required, allowing cells to adjust rapidly to the changing external conditions and cellular signals. σ^{70} in *E.coli* is the class of Sigma factor that is typically present in high amounts in a cell undergoing normal growth and is responsible for transcription of the cell's general house keeping genes and is coded by the *rpoD* gene. Other classes of Sigma factors in *E.coli* include heat shock response σ^{32} (coded by the *rpoH* gene), which directs the synthesis of gene products that stabilize or refold proteins; σ^{54} (coded by the *rpoN* gene) which directs the synthesis of gene products for assimilation of nitrogen compounds; σ^{E} (coded by the *rpoE* gene) is involved in periplasmic stress response; σ^{F} (coded by the *filA* gene) helps in transcription of genes involved in chemotaxis and flagellar structure; and the stationary phase σ^{S} (coded by the *rpoS* gene) controls transcription of gene products required during stationary phase and some stresses (Mooney et al., 2005; Ishihama, 2000).

1.2.2 Crystal structure of core RNAP

The crystal structure of the *Thermus aquaticus* RNAP core (Zhang et al., 1999) contributed enormously towards a better molecular insight into the structure of RNAP core enzyme.



· *Darst S.A. 2001; Zhang et al., 1999*

Figure 1.4: The *Taq* core RNAP structure. The a-carbon backbone is shown as worms, along with the transparent molecular surface. Various structural features discussed in the text are labeled and highlighted in yellow. The individual subunits are color-coded as indicated in the key. A yellow sphere indicates the Mg2+ ion chelated at the active center. **(a)** View looking into the active site channel, perpendicular to the axis of the main channel. This view is rotated **(b)** 57° clockwise about the vertical axis and **(c)** 95° counter-clockwise about the vertical axis (this view looks directly down the axis of the secondary channel). Structural elements like ß' bridge helix and rudder are shown

Over all, the enzyme adopts a structure reminiscent of a crab claw; the claws consisting of the large ß and ß' subunits. The two claws define the enzyme's main channel, which accommodates DNA template and the RNA-DNA hybrid that forms during transcription and contains the determinants for interactions with DNA and with the RNA-DNA hybrid. The active center, which is marked by a stably bound Mg^{2+} ion, lies at the back of this main channel. The molecule was about 150 Å long (from the back to the tips of the claws), 115 Å tall and 110 Å wide (parallel with the channel). The channel has many internal features, but the overall width is approximately 27 Å. The active site of RNAP located on the floor of the active center cleft, contains a tightly bound catalytic Mg^{2+} ion, coordinated by three Asp residues, forming a conserved NADFDGD motif in the conserved region D of the ß' subunit; these region also contains binding sites for NTP substrates. A second, smaller channel links the external environment to the active site of the enzyme. From crystal structures of the RNAP holoenzyme obtained later, it was surmised that since DNA largely blocks access to the active center via the main channel during transcription, this secondary channel could be the entryway for substrate nucleotides (NTPs) to reach the site of catalysis. A third channel called RNA exit channel mediates the exit of nascent RNA from the active center cleft. (Campbell et al., 2001; Cramer et al., 2000; Korzheva et al., 2000; Magyar et al., 1999). Distinct structural features or 'parts' important for function were also determined. These include a flap-like structure that appears to be flexibly connected to the rest of the RNAP (the ß G flap). Region F of ß' is most remarkable; it begins in the upper domain of ß', where it interacts with ß, and then forms a helical segment that traverses the middle of the main channel, bridging the top and bottom of the main channel and is called the ß' F bridge helix. The active center Mg^{2+} is positioned on the back wall of the main channel, directly opposite the bridging helix. The ß' F bridging helix conspires with ß' G, which extends into the main channel, to form a structure that bifurcates the main RNAP channel into two separate channels, dividing the active center cleft from the secondary channel. The smaller of these channels (the secondary channel) is roughly 12 Å in diameter; a coiled-coil-like structure extends from the main channel and supports a loop-like structure (comprising part of ß' C) that protrudes into the main channel, reminiscent of an upside-down rudder. Another structural element called the 'trigger loop' interacts with the bridge helix also forms the downstream face of the active center (Bar-Nahum et al., 2005; Cramer et al., 2000; Gnatt et al., 2001; Yuzenkova et al., 2002). The core enzyme contains all the catalytic machinery required for the synthesis of RNA from nucleotides, but is incapable of initiating transcription efficiently and with specificity. For this, it must bind an initiation factor, σ, to form a holoenzyme that can recognize specific DNA sequences called promoters (Gross et al., 1998; Olins et al., 1983; Gross et al., 1996).

1.2.3. Structure of the s subunit

Housekeeping σ subunit is divided into four conserved sequence regions that fold into four independent structural domains (σ 1.1, σ 2, σ 3, and σ 4; Figures 1.6 A and B). The domains are separated by flexible linkers; two of which can span significant distances (the σ 1.1–2 linker and the σ 3–4 linker, which contains conserved sequence region 3.2). In holoenzyme, these domains and linkers contact an extensive surface of RNAP (Murakami and Darst 2003). The σ factor has a central role in initiation, being directly involved in promoter recognition, DNA melting and promoter escape and clearance (Gross et al., 1996; Gruber and Gross, 2003a; Gruber and Gross, 2003b). Among several distinct σfactors (Gross et al, 1992; Ishihama, 2000), σ 70 is primarily responsible for the transcription of most 'housekeeping' genes in the cell during exponential growth. The σ subunit can be divided into four regions based on sequence conservations the various σ factors in prokaryotes (Helmann and Chamberlin 1988). The N terminal Region 1 (σ1) is found mainly in primary σ subunit (σ^{70} in E coli and σ^{43} in Bacillus) and its role appears to preventing σ binding by itself to DNA. This auto inhibition is relieved when σ binds to core RNAP. Region2 is found in all the s factors and is the most highly conserved region and is involved in binding to core RNAP, and the σ2 region also plays a role in recognizing the ⁻10 region of promoter. The σ3 region plays a role in recognition of the ⁻10 region and in the stabilization of the open complex. σ4 plays key roles in promoter recognition and also contains helix-turn-helix DNA binding domains, which plays key roles in RNAP-DNA binding. The binding of one of the several specificity sigma (σ) factors results in the formation of a holoenzyme that can recognize promoters. After the initial and reversible binding of RNAP to the duplex promoter DNA (called as a Closed complex formation), approximately 12 bp of DNA including the transcription start site is melted, exposing the transcription start site and forms an initiation competent open complex after a few conformational changes in the holoenzyme; the σ subunit dissociates from the core enzyme once a stable elongation complex is formed. This process was originally described as a ''σ cycle'' in which s associates with RNAP to orchestrate initiation and then dissociates after the transition to a stable elongation complex (EC) is complete (Travers and Burgess, 1969; Chamberlin, 1976). Once RNAP finish a transcription cycle and release DNA and RNA, it is free to be bound anew by s and begin another cycle of transcription. The key feature of the σ cycle is the ability of RNAP to be reprogrammed rapidly by different σ subunit in each new round of transcription. Thus, the σ cycle allows cells to adjust transcription patterns rapidly to optimize cellular metabolism in response to changing external conditions and

cellular signals and to orchestrate developmental programs by using different s subunit with different promoter specificities to regulate discrete sets of genes (Mooney et al., 2005).

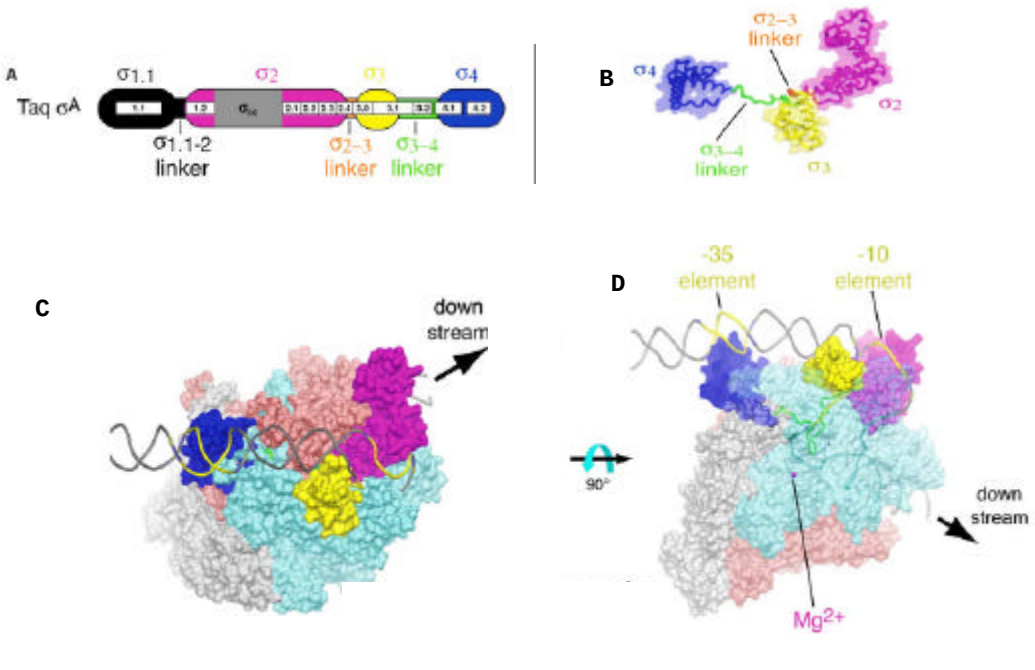


Figure 1.6: σstructure and contacts to RNAP

A. s **regions** Conserved σ sequences are shown (σ 1, black; σ 2, magenta; σ 3, yellow; and σ 4, blue) and linkers (σ 1.1–1.2 linker, black; σ 2–3 linker, orange; and σ 3–4 linker, green) between the domains (Campbell et al., 2002; Murakami et al., 2002; Vassylyev et al., 2002). σ1.2 and σ 2.1 are separated by a non-conserved sequence.

B. **Structure of s in holoenzyme**. The arrangement of σ domains and linkers observed in the RNAP holoenzyme (Murakami et al., 2002) is shown, colored as in (A). Transparent molecular surfaces of the structured domains σ 2, σ3, and σ4 are shown along with the backbone worm.

C and **D**. s **contacts on RNAP Holoenzyme**. RNAP is shown as a molecular surface, colored as follows: aI, a II, and ? , gray; ß, cyan; and ß', pink. RNAP surfaces contacted by s are colored according to the s structural element they contact (colored as in A). The ß flap is shown as a backbone worm, with its surface displaced to reveal s contacts behind it.

1.3. Transcription initiation

1.3.1 Promoter elements

A promoter is a conserved sequence of DNA whose function is to be recognized by a protein and the information for promoter function is provided directly by the structure of DNA sequence and acts as a *cis*-acting site. They are also sites where RNA polymerase binds, melts the DNA, and initiates transcription. Different sequences modulate the promoter engagement of RNAP. Prokaryotic promoters contain two regions centered at ¯10 and ¯35 bp upstream to the

transcription start site, and in *E.coli* usually has a resemblance to two hexametric consensus sequences, TATAAT and TTGACA. The DNA sequence between these two elements is called the Spacer region. A few strong promoters also contain an extra element, the UP element, 40 to 60 bp upstream to the core promoter, which makes these promoters more attractive for RNAP engagement. Thus there are four conserved features in a bacterial promoter: the start point, the ¯10 sequence, the ¯35 sequence and the separation between the ¯10 and the ¯35 sequences called a spacer element, spanning 16-19 bp.

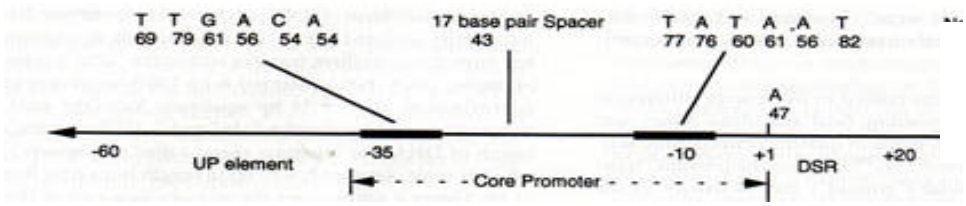
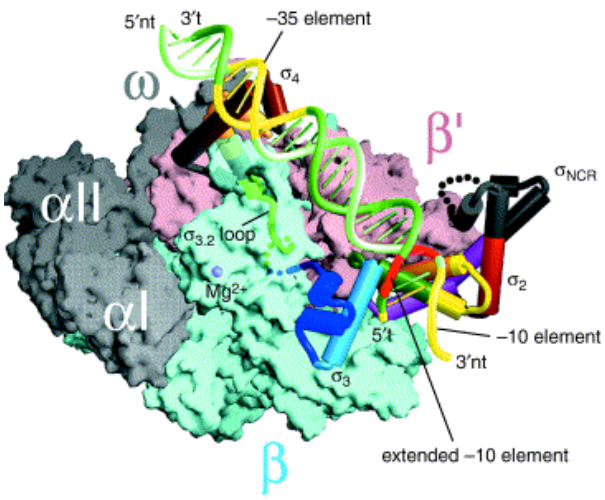


Figure 1.7: Cartoon representing the structure of a consensus s ^{70}promoter
The relative percentages of the NTPs at the –35 and –10 elements are also given.

1.3.2 Structure of RNAP-promoter

The core Polymerase has a structure configuration whereby there is a closed channel with a diameter of about 25Å, enough to accommodate a double helical DNA, whereas in holoenzyme, this channel is open (Figure 1.10). The promoter engagement by RNAP can be divided into different steps: the promoter searching and recognition by an RNAP holoenzyme (the s subunit complexed to the RNAP core plays an important role in selecting the promoters to which RNAP will bind tightly), initial and reversible binding of RNAP to the duplex promoter DNA forming a closed complex (RP_c) and then the formation of an open complex (Rp_o) in which approximately 12 bp of DNA including the transcription start site is melted (Mooney et al., 1998).



Murakami & Darst., 2003; Murakami et al., 2002.

Figure 1.8: Structure of bacterial RNAP holoenzyme with promoter DNA associated
All sequence-specific contacts with promoter DNA (at –10 and –35, colored yellow, the extended –10 element colored red) mediated by s subunit are shown as cylinders. a1, a 2, ? subunits are colored gray, ß subunit is colored cyan ß' subunit pink, Template DNA is dark green, while non-template DNA is light green.

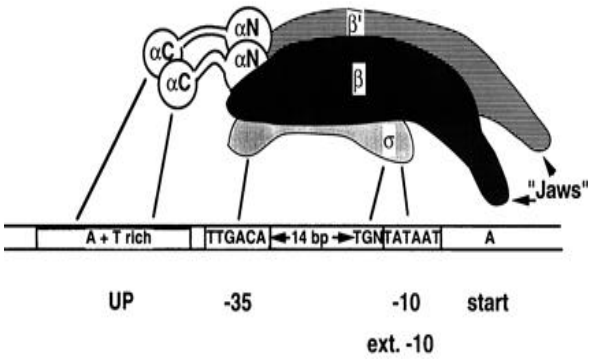


Figure 1.9: Summary of RNA polymerase-promoter interactions A promoter with consensus sequences for the ¯10 and ¯35 regions (boxed) is shown; The jaws region of the RNA polymerase which would grasp the DNA downstream of the catalytic site of RNA polymerase are shown on the right of the molecule. The solid lines show contacts between RNA polymerase and promoter DNA.

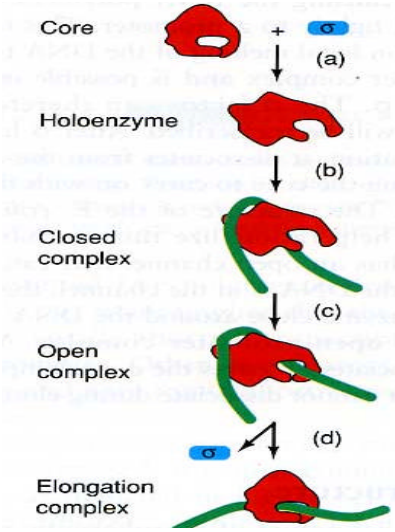


Figure 1.10: Model of interaction between *E.coli* RNAP holoenzyme and a promoter

a. Core RNAP (colored red) with its closed channel bind to s subunit (colored blue) to form holoenzyme with an open channel. **b.** The holoenzyme binds loosely to the DNA (green) to form a closed complex. **c**. The s factor allows initiation of transcription by causing the RNAP to bind tightly to the promoter, local melting of DNA occurs, forming an open promoter complex. **d.** The s subunit dissociates, leaving core firmly clamped around the DNA and form a stable complex, which can transcribe DNA processively.

Robert F Weaver, Molecular Biology, IInd edition, 2003; McGraw Hill

1.3.3 Initiation of transcription

Initiation of transcription is a process involving several different phases. The first stage in initiation is the promoter location by RNA polymerase, brought on by the translocation or 'hopping' of RNAP along the DNA. The next sage involves the formation of a competent initiation complex. The initial contacts between RNA polymerase and the promoter forms the first double-stranded, or closed, intermediate called RP_{c1} and an intermediate called RP_{c2}. Two intermediates (designated I_1 and I_2 in Fig 1.11) are kinetically significant as their interconversion is the rate-limiting step in open-complex formation and the conversion of I_1 to I_2 is the step at which a jaw like structure in RNA polymerase closes around the DNA. Conformational changes occur in the DNA too at this stage, the most important being the strand separation extending from

the -10 region to past the start site. Completion of the process of RNA polymerase-induced strand opening is dependent on Mg^{2+} The complex formed in the absence of Mg^{2+} has been designated RP_{o1} (Fig 1.10) and has a base pair opening extending from the middle of the -10 region to bp -1, just upstream of the start site. It is thought to be an intermediate in the formation of the functional open complex RP_{o2} (Fig 1.11), formed in the presence of Mg^{2+} where the opening extends from -12 to +2. Once formed, the open complex is stabilized by interactions between single-stranded DNA and the RNA polymerase. After strand opening, the template strand becomes accessible to the NTPs so that it can (by Watson-Crick base pairing) specify the sequence of the RNA to be synthesized. The kinetic stability of the open complex is be enhanced by the binding of the initiating NTP and the synthesis of the initial phosphodiester bonds. Though the ternary complexes of promoter, RNA polymerase, and NTP(s) are poised to initiate productive RNA synthesis after the formation of the first phosphodiester bond, several rounds of abortive transcription in which short (<10 nt) RNA transcripts (productive and non productive transcription initiation complexes) are made and released in a repetitive manner until a critical length (about 10 nt) is reached for the nascent mRNA transcript, whereupon the complex became committed to productive RNA synthesis (de Haseth et al., 1998; Vo et al., 2003). Once an RNA chain of eight or nine nucleotide residues has been synthesized, s factor is released; RNA polymerase leaves the promoter and becomes committed to productive chain elongation.

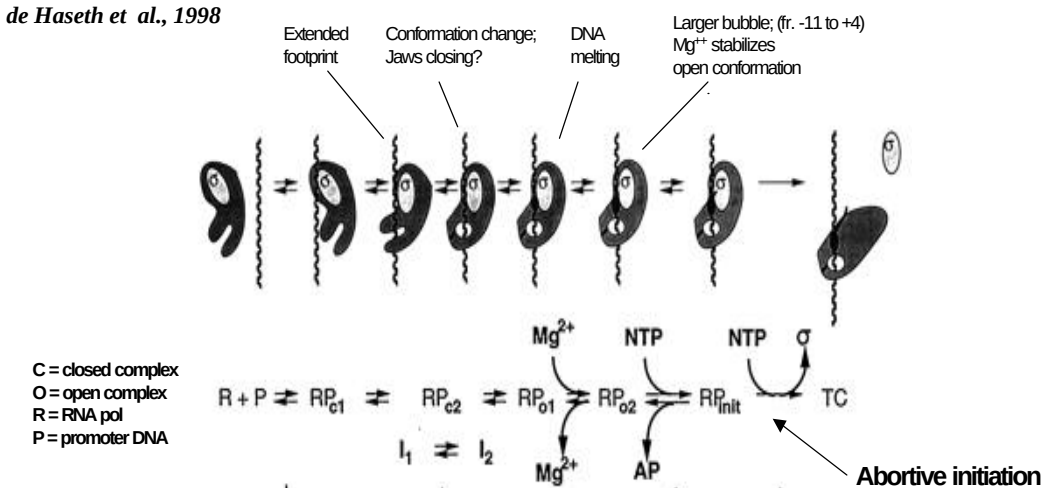


Figure 1.11: Cartoon summarizing the intermediates in the process of initiation of RNA synthesis The complexes are shown in cartoon form at the top with a descriptive notation below. In the designation of structurally characterized complexes, R stands for RNA polymerase, P stands for promoter DNA, and c and o indicate closed and open complexes, respectively (where strand separation has not or has occurred). I_1 and I_2 are RP_{c2}-like complexes, and are kinetically significant intermediates. AP stands for abortive RNA product, and TC stands for the transcribing complex in elongation mode.

During the transition from promoter binding, initiation and to elongation, RNAP undergoes a significant conformational change (Record *et al.,* 1996; Craig et al., 1998; Naryshkin et al., 2000; Polyakov et al., 1995). Once the open complex has bound to the initiating NTPs, it becomes an initial transcription complex and can follow several alternate reaction pathways (Figure 1.1). After the synthesis of RNA is initiated from the promoter and after a few abortive cycles of RNA synthesis in which the short nascent RNA is released, the RNAP breaks its interactions with the promoter DNA (this process is called promoter clearance) and form a very stable elongation complex which translocates along the DNA adding ribonucleotide triphosphates (NTPs) successively to the growing chain of RNA. The basic reaction cycle here consists of binding the appropriate NTP, incorporation of the associated nucleotide monophosphate into the RNA and the release of pyrophosphate.

1.3.4 Transcription factors affecting Initiation

There are many examples of agents or conditions that can affect (inhibit or activate) the interconversion of intermediates during transcription initiation; some of these have different effects, depending on the promoter or RNA polymerase. In some cases, modification of an activator leads to DNA binding that enables initiation of transcription by an RNA polymerase. In others, interaction with an inducer relieves inhibition by a DNA-binding repressor. Proteins like the *lac* repressor, CRP and elongation factors GreA and GreB (which can affect promoter escape *in vitro*) are examples of *E. coli* proteins which can affect various intermediates in transcription initiation. Phage-encoded proteins like cI (from bacteriophage lambda) display polymerase or promoter dependent effects. Rifampicin is an antibiotic which targets the ß subunit of RNA polymerase affecting transcription initiation. The *E. coli* cAMP receptor protein (CRP, also referred to as catabolite activator protein or CAP) is an example for a transcriptional activator, which activates genes in response to glucose starvation and other stresses. The activity of CRP is triggered by binding of cAMP. CRP functions as a homodimer and, at target promoters, binds and sharply bends a 22-bp two fold-symmetric DNA site. CRP activates transcription initiation at most target promoters by making direct protein–protein interactions with a CTD that recruits a CTD, and hence the rest of RNAP, to promoter DNA. Simple CRP-dependent promoters (i.e., promoters that have only one DNA site for CRP) can be grouped into two classes based on the position of the DNA sites for CRP (Busby and Ebright 1999). At class I CRP dependent promoters, the DNA site for CRP is located upstream of the core promoter (i.e., centered near positions -61, -71, -81 or -92 in the DNA), and CRP recruits a CTD to the DNA segment

immediately downstream of the DNA site for CRP (Fig 1.12). At class II CRP-dependent promoters, the DNA site for CRP overlaps the core promoter (i.e., centered near position -41), and CRP recruits a CTD to the DNA segment immediately upstream of the DNA site for CRP. Transcription activation involves both direct protein–protein interaction (black filled circle in Fig 1.11) between AR1 of the upstream subunit of CRP and the 287 determinant of a CTD, and direct protein–protein interaction (gray filled circle in Fig 1.12) between AR2 in the downstream subunit of CRP and a subunit N-terminal domain (Lloyd et al., 2002).

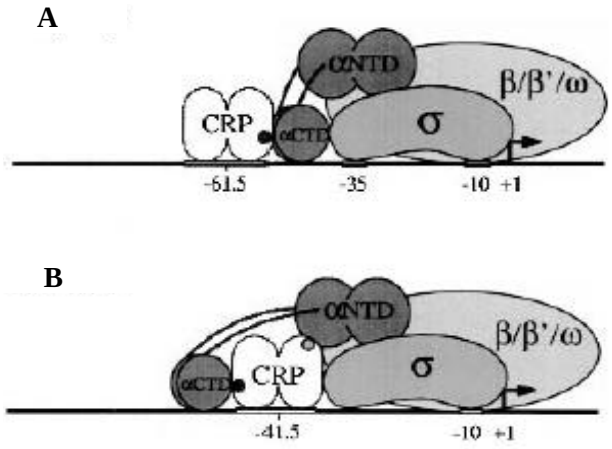


Figure 1.12: Transcription activation at class I and class II cyclic AMP receptor protein (CRP)-dependent promoters
A. Ternary complex of CRP-*Escherichia coli* RNA polymerase holoenzyme and a class I CRP-dependent promoter
B. Ternary complex of CRP, RNAP, and a class II CRP-dependent promoter

Lloyd et al., 2002

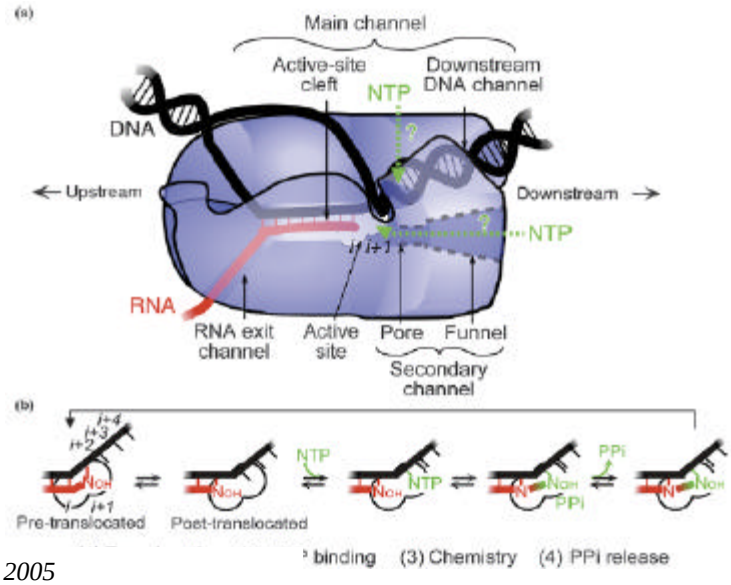
1.4 Transcription Elongation

In most cases, once initiation of transcription is accomplished, the EC loses s factor associated with it and normally continues to elongate the RNA transcript, until it reaches a termination site.

Many elongation factors and other regulatory proteins bind to the RNAP as well as to DNA, forming a stable complex called as Ternary Elongation Complex (TEC) and can elongate the RNA chain at a rate of 30 to 100 nt/s *in vivo*. During its translocation along the DNA, the TEC encounters pause sites and potential arrest sites, at which the nascent RNA remains stably bound to the enzyme (Cramer et al., 2000; Korzheva et al., 2000; Nickels and Hochschild, 2004). Two classes of models have been posed regarding the mechanism of translocation during transcriptional elongation. The first class postulates that a power stroke tightly coupled to pyrophosphate release drives the motion (Yin and Steiz; 2004). The second class proposes that the reversible diffusion of the enzyme along the DNA template between its pre- and post translocated states is directionally rectified through the binding of the incoming NTP in a Brownian ratchet mechanism (Bar-Nahum et al, 2005; Bai et al., 2004; Guajardo and Sousa, 1997).

1.4.1 The Chemistry of polymerization

During RNA synthesis, the TEC can undergo two types of movements-translocation, occurs during nucleotide addition and moves RNAP one step forward, melting one bp downstream, allowing one bp upstream to re-anneal and moving the RNA 3' nucleotide from the NTP binding site (i+1 site) to the product site (i site). NTP can pair to the template strand base at the i+1 site and react to form a new 3'neucleotide (RNA 3' end translocation, Figure 1.13 b). The polymerization reaction occurs via a pentavalent phosphate intermediate, with the help of the two Mg^{2+} ions in active site, which directs the 3' OH and 5' phosphate of the participating NTPs into proper alignment. At pause, arrest and termination sites, changes in the location of the 3' nt relative to active site and changes in the conformations of active site and the RNA exit channel leads to transcriptional arrest or reverse translocation of RNAP (Mooney et al., 1998; Landick, 2001; Li et al., 2005; Murakami and Darst, 2003; Toulokhonov et al., 2001).



Landick, R. 2005

Figure 1. 13 Catalysis by RNAP

a. Diagram of a TEC. The two possible routes of NTP entry (main channel and secondary channel) into RNAP (blue) are indicated (green). The features of RNAP that is responsible for binding to DNA and RNA (the RNA exit channel, the main channel subdivided into the downstream DNA channel and active-site cleft), the active site (subdivided into the i and i +1 sub sites) and the secondary channel (subdivided into the funnel and pore) are labeled.

b. The nucleotide-addition cycle. The four main steps of nucleotide addition are depicted:

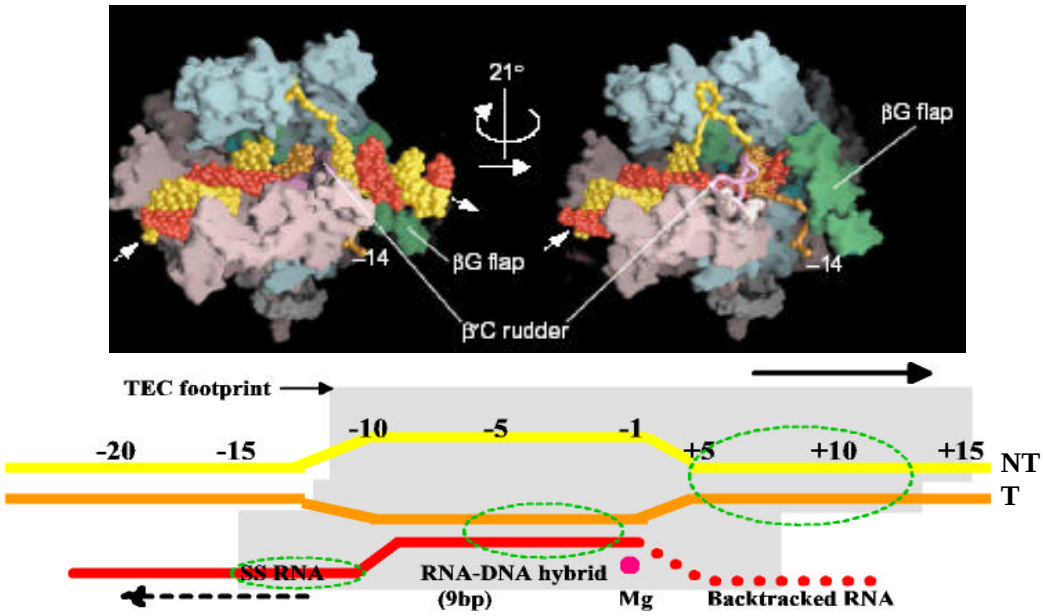
1) Translocation of the RNA 30 nucleotide from the i +1 sub-site it occupies in the pre-translocated register to the i sub-site it occupies in the post-translocated register;

2) Binding of NTP either at the main channel or the secondary channel; **3)** Nucleophilic attack of the RNA 3'OH on the NTP a phosphorus atom and polymerization; **4)** Release of pyrophosphate (PPi). It is possible, although unproven, that translocation could be coupled to either NTP binding or PPi release.

1.4.2 Protein-Nucleic acid contacts in EC

During RNA synthesis, DNA strands are separated in a short area around the RNAP catalytic center, forming a transcription bubble and an RNA-DNA hybrid between the nascent RNA and the single stranded template DNA is formed inside the transcription bubble. This hybrid is transient and is displaced after termination from the DNA template and the DNA strands re-anneal to form the double stranded up-stream DNA. Three nucleic acid binding sites in the RNAP that stabilize the TEC have been characterized through biochemical and structural means. They are, a DNA-binding clamp that holds the downstream DNA (DBS), a hybrid binding domain consisting of a 8-9 nucleotide RNA: DNA hybrid, 3' proximal region of which interacts directly with regions in and around the active site (HBS) and a single stranded RNA binding channel which interacts with the exiting nascent RNA (RBS) (Nudler et al., 1996; Sidorenkov et al., 1998; Korzheva et al., 2000; Darst, 2001; Gnatt et al., 2001). DBS contacts the ~9 bp of DNA duplex, distal to active site and to the amino terminal parts of the ß subunit and the amino and carboxyl-terminal parts of ß'subunit. The HBS has close contacts between the ß' and ß subunits and to the 8-9 bp of RNA-DNA hybrid, located within the 17 nucleotide melted DNA duplex or "bubble" (Korzheva et al., 2000; Nudler et al., 1996). The RBS of the RNAP consists of regions interacting with the single stranded nascent RNA -7 to -14 relative to its 3' terminus (Kandel and Nudler, 2002). Protein-nucleic acid interactions at these sites contribute to overall stability of TEC. Many structural elements present in RNAP around the active site bring about the spatial organization required during these steps. A structural element comprised of the conserved segment C of the ß' subunit called the 'Rudder' (Gnatt, 2002; Gnatt et al., 2001; Zhang et al., 1999) is thought to play a role in maintaining the correct length of the RNA-DNA hybrid, in stabilizing the nascent RNA in complex and in promoting and maintaining the localized DNA melting at the upstream edge of the bubble. Another structurally important element in the TEC is the ß flap, which constitutes the RNA exit channel along with a mobile clamp domain which includes several loops that cover RNAP's main channel and RNA exit channel comprising of structural elements like the rudder, zipper, lid and the Zinc Binding Domain or ZBD. The flap tip can modulate intrinsic termination and is required for pause RNA hairpin inhibition of catalysis by RNAP (Toulokhonov and Landick, 2003). The ß' Lid is another structural element that protrudes from the clamp domain into the main channel. The ß' Lid is composed of a 14 amino acid residue loop, in the conserved segment B of the ß' subunit. Based on its location in the transcription complex, the ß' lid could determine the length of RNA-DNA hybrid by peeling the RNA away from the hybrid, stabilize the interaction of the nascent RNA with RNAP and help maintain the localized DNA melting at the upstream boundary of transcription bubble by interacting with DNA and other structural

elements and also participates in transcription termination (Naryshkina et al., 2006; Darst, 2001; Toulokhonov and Landick, 2006)



TEC is extremely stable yet surprisingly dynamic *Korzheva et al., 2000*

Figure 1. 14 Structure of a Ternary Elongation Complex

Thermus aquaticus core RNAP is shown as color-coded molecular surfaces. ß subunit is colored cyan, ß' subunit pink a1, a 2, ? subunits are colored white and Mg $^{2+}$ at active site is colored magenta. DNA template is colored red, non-template is colored yellow and RNA is colored gold. Directions of entering downstream DNA and exiting upstream DNA are shown as white colored arrows. Structural elements like ß flap and ß' rudder is also shown.

1.4.3 Transcriptional pausing

Pausing can be defined as a temporary delay in RNA chain elongation and it synchronizes transcription and translation in prokaryotes. It also slows RNAP to allow timely interaction of regulatory factors and acts as precursors for both transcriptional arrest (complete halting of the EC, but without dissociation of the nascent RNA) as well as factor dependent and independent transcription termination. Pause sequences could produce different TEC configurations (Artsimovitch and Landick, 2000), especially with respect to the growing RNA chain and altered active site mechanizations. Pausing can act to regulate the rates of polymerization, synchronize transcription and translation in prokaryotes, bind cofactors to modify transcription and facilitate

the co transcriptional folding of transcripts (Artsimovitch and Landick, 2002; Bailey et al., 2000; Dalal et al., 2006; Marr and Roberts, 2000).Two classes of defined pauses with regulatory functions have been identified. Class I pauses are one where nascent RNA hairpin-RNAP interaction inhibits nucleotide addition by stabilizing the RNA 3'OH in a frayed or hypertranslocated position (Artsimovitch and Landick, 1998; Artsimovitch and Landick, 2000; Chan et al., 1997; Dalal et al., 2006; Toulokhonov et al., 2001; Toulokhonov and Landick, 2003). Backtracking pauses (Class II pauses) occur when the enzyme encounters a weak RNA-DNA hybrid, inducing RNAP to move upstream on the DNA template there by extruding the nascent RNA 5' end into the nucleotide entry channel (Dalal et al., 2006; Komissarova and Kashlev, 1997; Reeder and Hawley, 1996). Both Class I and Class II paused TECs can form an "unactivated intermediate" conformation. (Erie, 2002; Neuman et al., 2003; Palangat and Landick, 2001) and also can form distinct TEC conformations (Artsimovitch and Landick, 2000), involving a complex network of protein-protein and protein -nucleic acid interactions in the TEC (Toulokhonov et al., 2001).

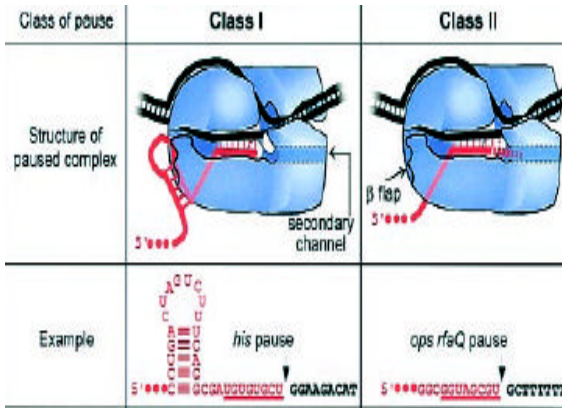


Figure 1.15: Structures of two classes of pause signals

Structures of class I and class II paused TECs are depicted with DNA (black) entering RNAP (blue). RNA (red) pairs with template DNA in an eight-bp hybrid (vertical red lines), then exits under the ß flap domain. RNA nt in the hybrid are underlined. At a class I pause, interaction of the pause hairpin with the ß flap domain displaces the RNA 3'OH away from the catalytic center. At a class II pause, RNAP enters pretranslocated or backtracked conformations as dashed red line).

Artsimovitch and Landick 2000

1.4.4 Host protein factors in regulation of elongation

Transcription is a tightly regulated event and involves a high level of complexity. Many transcription factors acts on DNA, RNAP and RNA during gene expression, in a temporally and spatially co-coordinated manner (Borukhov et al., 2005). Though most of the *E.coli* transcription factors that are known are DNA binding proteins (Ishihama, 2000) that regulates transcription initiation, many proteins are known to act during the elongation and termination stages of transcription by direct modifications of RNAP properties. Proteins like the Nus factors (NusA,

NusB, NusG and NusE (S10), RfaH, Gre-factors (GreA, GreB), Mfd and Rho (Bailey et al., 1997; Bailey et al., 2000; Squires and Zaporojets, 2000; Fish and Kane, 2002; Kandel and Nudler, 2002; Roberts and Park, 2004), bacteriophage proteins like N and Q from ? phage, Nun of HK022, Alc of T4 (Friedman and Court, 1995; Kandel and Nudler, 2002; Weisberg and Gottesman, 1999; Yarnell and Roberts, 1992), P7 of Xp10 (Nechaev et al., 2002, Psu from P4 (Linderoth and Calendar, 1991; Pani et al., 2006) also comes under proteins which regulate transcription during its different stages.

1.4.4.1 NusA host elongation factor

NusA is an essential multifunctional transcription elongation factor that is universally conserved among eubacteria and archaea (Kandel and Nudler, 2002). NusA can elicit different effects on transcription, based on the DNA/RNA sequence and in the presence or absence of other transcription factors. NusA by itself increases the rate of termination at many intrinsic terminators (Farnham et al., 1982; Liu et al., 1996; Nudler et al., 1998; Schmidt and Chamberlin, 1987) and induces type I pausing, as in *his*- or *trp*- pause hairpins and factor independent transcription termination. However, in complex with other Nus factors and phage proteins like N and Q, it can stimulate N mediated transcription antitermination, at both Rho dependent and independent terminators. The multiple functionality of NusA could be reflected by its multidomain structural organization. NusA is an elongated protein composed of several discrete domains (Gopal et al., 1999; Shin et al., 2003; Worbs et al., 2001). The N terminal RNAP binding domain (NTD) is connected through a flexible hinge to three globular domains, S1, KH1 and KH2 (Figure 1.16). *E.coli* NusA has an additional C terminal extension consisting of two auto-inhibitory domains, AR1 and AR2, which acts by preventing S1 and KH1 domains from binding RNA by free NusA (Mah et al., 2000; Mah et al., 1999; Zhou et al., 2002).

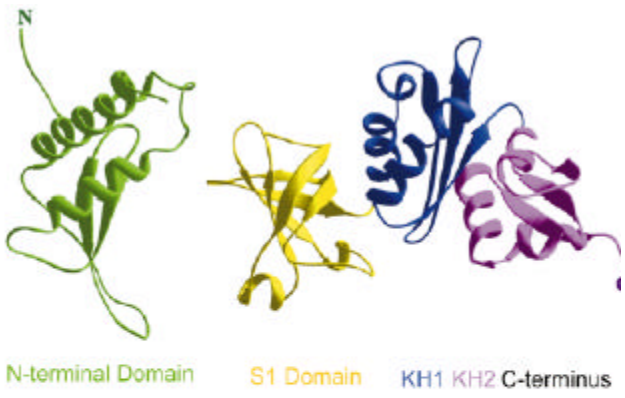
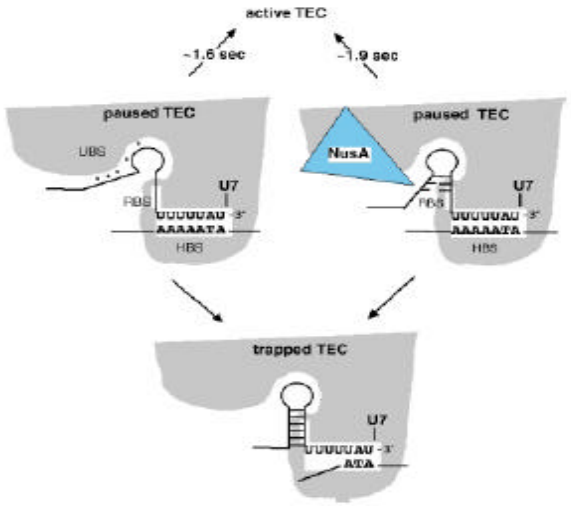


Figure 1.16: The structure of *M. tuberculosis* NusA

The N-terminal domain, the S1 domain and the

During transcription antitermination, the S1 and KH1 domains of NusA may interact with RNAP ß and ß' that form the RNA exit channel and the ß' coil-coil regions and could indirectly facilitate the formation of termination hairpin by disrupting the different RNA-RNAP interactions (Borukhov et al., 2005; Figure 1.17). NusA S1 and KH1 domains could also be actively involved in the strengthening of its specific interactions with structural elements surrounding the RNA exit channel, especially the ß' flap tip helix (Toulokhonov et al., 2001) modulating class I transcriptional pausing. The mechanism by which NusA stimulates termination involves both kinetic and mechanistic components. It is known that NusA decreases the overall rate of elongation and hence may stimulate termination by prolonging the pause at the downstream portion of the U stretch, at the termination point, thereby providing additional time for the hairpin to form. NusA also has a direct mechanistic effect on intrinsic termination by weakening the RNAP-RNA transcript contacts and can also suppress the steric hindrances to hairpin formation imposed by RNAP (Nudler and Gottesman 2002; Gusarov and Nudler 2001).



Gusarov and Nulder, 2002

Figure 1.17: A model for mechanism of Termination control by NusA
RNAP is colored gray, with RNAP nucleic acid binding sites like HBS (Hybrid Binding Site), RBS (tight RNA binding site) and UBS (weak Upstream RNA binding site) shown. NusA (colored blue) interacts with the RNAP and weakens the contacts between the UBS of RNAP and the upstream shoulder of the unfolded hairpin, thus stimulating the formation of a hairpin structure in the nascent transcript, resulting in transcription termination. NusA also prolongs pausing of the RNAP at the termination point.

1.4.4.2 Host elongation factor NusG

NusG is another important bacterial transcription elongation factor, and can accelerate the rates of RNA chain growth by RNA polymerase (Burova et al., 1995). NusG also plays key roles in termination along with Rho protein. It is a 21 KD essential protein in *E. coli*, (Burova et al., 1999; Downing and Dennis, 1991; Mason and Greenblatt, 1991) and modulates Rho-dependent transcription termination (Pasman and von Hippel, 2000; Sullivan and Gottesman, 1992; Sullivan et al., 1992). *Invitro* (Li et al., 1992; Mason and Greenblatt, 1991) and *in vivo* studies (Burova et

al., 1999) has shown that *E.coli* NusG influences and is required for efficient N mediated transcription antitermination. NusG facilitates the termination activity of Rho protein both *in vivo* and *in vitro*; binds to Rho protein (Li et al., 1993) and is the bridge between Rho and RNAP (Nehrke and Platt; 1994). Although it was initially identified as a component of the N antitermination complex, *E. coli* NusG was subsequently shown to be a regulator of Rho-dependent transcription termination (Pasman and von Hippel; 2000), phage HK022 Nun termination (Sullivan and Gottesman; 1992) and rRNA transcription antitermination (Li et al., 1993, Squires and Zaporojets; 2000). *Aquifex aeolicus* NusG is comprised of three discrete domains (D1- D3, given as I,II, III in Figure 1.18). D1 and D3 are derived from conserved sequences with counterparts in all NusG proteins. D2, on the other hand, is composed of residues corresponding to the variable insertion sequence (Figure 1.18). Another highly flexible region is a six-residue linker between D1 and D3. NusG stimulates the escape of RNAP from Class II pauses, while it has no specific effect on hairpin dependent Class I pauses (Artsimovitch and Landick, 2000).

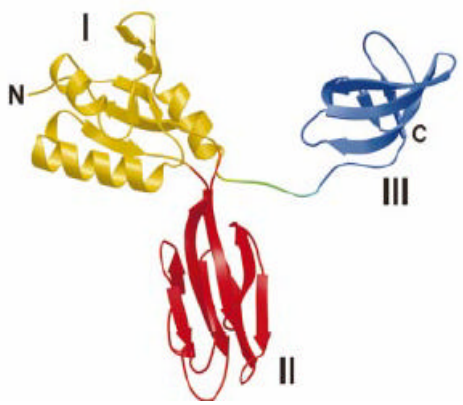


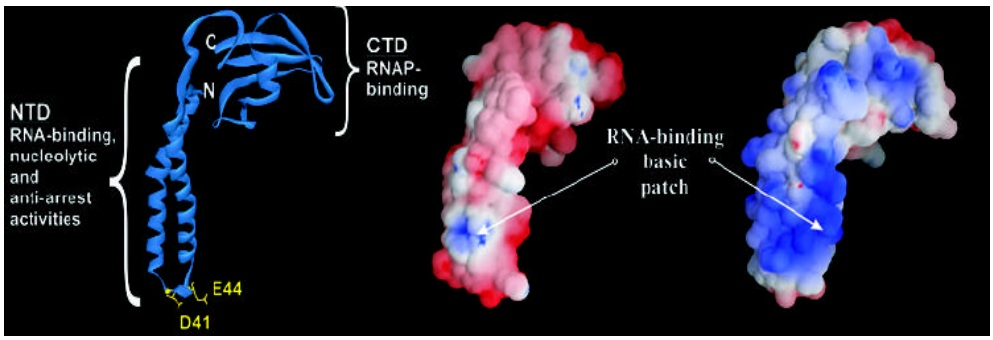
Figure 1.18: Crystal structure of NusG from *Aquifex aeolicus.*
The three domains and the six amino acid linker of NusG protein are shown.

Steiner et al., 2002

1.4.4.3 Transcript clevage factors GreA and GreB

Transcript clevage factors GreA and GreB suppress RNAP pausing and arrest *in vitro* and *in vivo* by stimulating the intrinsic nucleolytic activity of RNAP. During transcription, RNAP occasionally backtracks on the DNA template, displacing the 3' tail of the transcript into the secondary channel. An endonucleolytic clevage reaction generates a new RNA 3'-OH that is correctly positioned for nucleotide addition at the active site of the enzyme and the reaction though catalyzed by the RNAP active site, is enhanced by the presence of GreA and GreB (Borukhov et al., 2001; Borukhov et al., 2005; Opalka et al., 2003; Sosunova et al., 2003). The CTD of Gre binds to the ß' coiled-coiled region that forms the rim of the secondary channel

opening. The N terminal coiled-coiled domain (NTD) of Gre factors contain a cluster of positively charged residues that define a surface exposed basic patch, well positioned to contact the extruded RNA in backtracked transcription complexes. The interaction between this basic patch and the extruded RNA also stabilizes the binding of GreA/B to the elongation complex. The NTD of the Gre factors protrudes into the RNAP secondary channel and this protrusion is in such a way that this positions a pair of conserved residues D41 and E44 located at the tip of NTD very near to the RNAP active site and can stimulate transcript cleavage by coordinating the active site Mg^{2+}ion and the water molecule required for catalysis of RNA hydrolysis. The mechanistic model of Gre-induced hydrolysis is based on the two-metal-ion mechanism of phosphoryl transfer reactions. Here three invariant Asp residues in the RNAP catalytic site chelate one of the two catalytic Mg^{2+} ions, while the second Mg^{2+} is weakly bound by RNAP. The carboxyls of D41 and E44 from Gre directly participate in the co ordination of the second Mg^{2+} and also co-ordinates the attacking water molecule (or hydroxide ion) to stabilize the pentavalent transition state of the phosphate (Laptenko et al., 2003; Borukhov et al., 2005).



Borukhov et al., 2005

Figure 1.19: Structural organization of Gre proteins

A. Ribbon structure of *E. Coli* GreA with side chains of the essential E44 and D41. Charge distribution on the solvent accessible surface of GreA (middle) and GreB (right) are also shown. The surface is colored by the electrostatic potential: white, uncharged; red, negative (Asp, Glu); blue, positive (Arg, Lys). Only one side of the proteins is shown.

1.5 Transcription termination

TECs become extremely unstable and dissociate from the template DNA in response to specific signals called terminators. About 50% of protein-coding transcription unit of *E.coli* ends with this signal (Lesnik et al., 2001). The final step in transcript formation is termination, which occurs

when the elongating complex moves into one or more terminator sites along DNA template that may serve as transcription regulators within genes or mark the end of a gene or an operon (von Hippel, 1998), where RNAP releases the RNA transcript and dissociates from the DNA. Specific termination sites on the DNA fall into two classes – factor dependent, where termination factors like Rho (Richardson and Richardson, 1996; Roberts, 1969) acts; and factor independent intrinsic terminators.

1.5.1 Factor independent (hairpin dependent) transcription termination

A canonical intrinsic termination signal in DNA is composed of a GC-rich palindromic element followed by an oligo (T) sequence ('T-stretch'). In this multipartite signals, the G/C rich nascent RNA hairpin is located 7-9 nucleotides from the point of release (d'Aubenton Carafa et al., 1990).

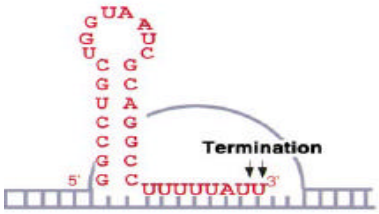
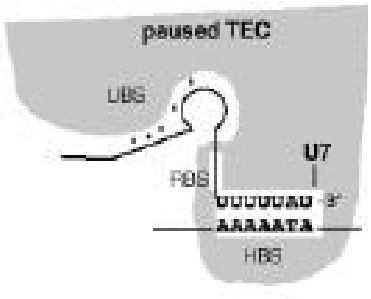


Figure 1.20: Cartoon representing an intrinsic terminator formed on a DNA template,
The terminator RNA having a stem loop structure (GC-rich palindromic element) followed by the U rich sequence is shown in Red. DNA is colored

The inverted repeat allows a hairpin to form at the end of the RNA transcript and the T-rich region that follows immediately in the non-template strand of the gene results in a string of weak rU-dA base pairs holding the transcript to the template strand in the RNA-DNA hybrid. The 3' proximal RNA is U rich and invariably contains three U's immediately after the hairpin, the downstream portion of which induces a transient pausing of RNAP (Gusarov and Nudler 1999). The pause at the terminator modulates hairpin formation and subsequent dissociation of RNA from the EC (Artsimovitch and Landick, 2000), as it provides more time for the RNA hairpin to form. The formation of a terminator hairpin closer to the RNA 3' end and the weak rU. dA base pairing favors transcripts release, resulting in transcription termination. Three nucleic acid binding sites in RNAP that stabilize the TEC have been characterized. They are, the double stranded DNA binding site (DBS) that contacts the DNA duplex just promoter distal to the catalytic site and consists of the NH_2 terminus of ß subunit and the NH_2 and COOH terminal parts of the ß' subunit; the RNA-DNA heteroduplex binding site (HBS), characterized by contacts between the ß and ß' subunits, the RNA-DNA hybrid located within the melted DNA or the 'bubble'; and the single stranded RNA binding site (RBS), consisting of a region close to the –7 to –14 of the nascent RNA 3' terminus. The RBS includes the rudder region and NH_2 terminus of ß' subunit as well as the COOH terminus of the ß subunit.

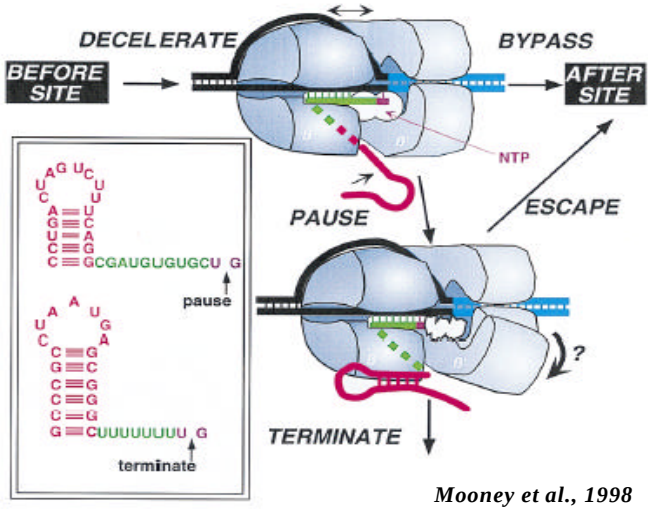


Gusarov and Nudler 2001

Figure1.21: Cartoon representing the nucleic acid binding sites in RNAP

Three nucleic acid binding sites in RNAP: the DBS, HBS and RBS are shown. At the moment of termination, the U-stretch occupies HBS, generating an unstable A:U hybrid, while the hairpin occupies the RBS, destabilizing RNA binding to RBS. The T stretch in the DNA also induces pausing of RNAP, providing additional time for hairpin to form as well as causing backtracking, displacing the 3' end of RNA from the catalytic center, temporarily inactivating the TEC.

The terminator destroys the interactions at all the stable nucleic acid binding sites of RNAP, as RNA and DNA are rapidly dissociated during intrinsic termination. The canonical intrinsic terminator signal is designed so that, at the moment of termination, the U stretch occupies the entire HBS, generating an unusually unstable A:U DNA:RNA hybrid; while the potential hairpin occupies the single strand specific RBS. The hairpin formed displaces RNA from the RBS and can also unwind the 4-5 bp of the A:U hybrid adjacent to the hairpin stem. Thus this simultaneous disruption of key interactions in both HBS and RBS (Gusarov and Nudler, 1999; Yarnell and Roberts, 1999), and the loss of these vital contacts propagate instability to the DBS and destabilize the entire TEC (Nudler and Gottesman 2002). A combination of these events brings about the termination process and the release of RNA occurs usually at the 7th and 8th U downstream of the hairpin (d'Aubenton Carafa, 1990).



Mooney et al., 1998

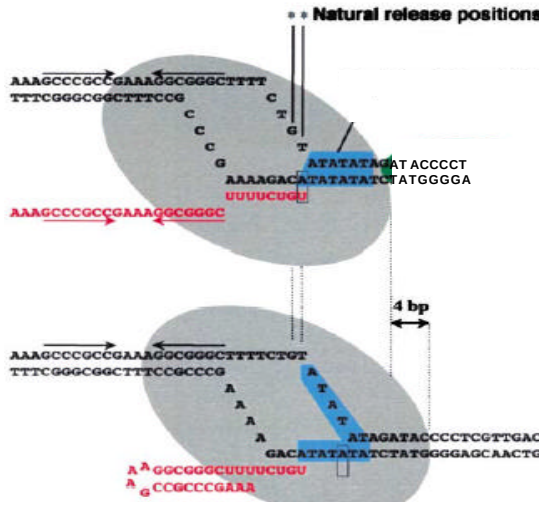
Figure 1.22: Cartoon showing factor independent transcription termination

The terminator or pause hairpin is colored magenta, 3' proximal

Most termination models for factor independent of hairpin dependent transcription termination are evolved from the works of several groups, and involve either of the following steps. The U-

rich RNA: DNA hybrid induces pausing at the termination point (Sharp et al., 1986; McDowell et al., 1994; Nudler et al., 1998) which is believed to allow the hairpin to form at the RBS.

The rigid body model or the forward translocation model of termination, suggests that a terminator hairpin begins forming when its loop and upper stem has emerged out of the RNA exit channel, which pulls RNA through the channel and away from the active site to avoid a steric clash with a rigid RNA polymerase as the lower stem pairs). This in turn partially unwinds the RNA-DNA hybrid and moves RNAP forward without nucleotide addition (Yarnell and Roberts, 1999; Santangelo and Roberts, 2004; Figure 1.23). RNAP and the associated transcription bubble both translocates downstream in the absence of RNA synthesis, thus mediating the release of RNA transcript by freeing it from the enclosure within the enzyme (bubble) as an RNA-DNA hybrid.

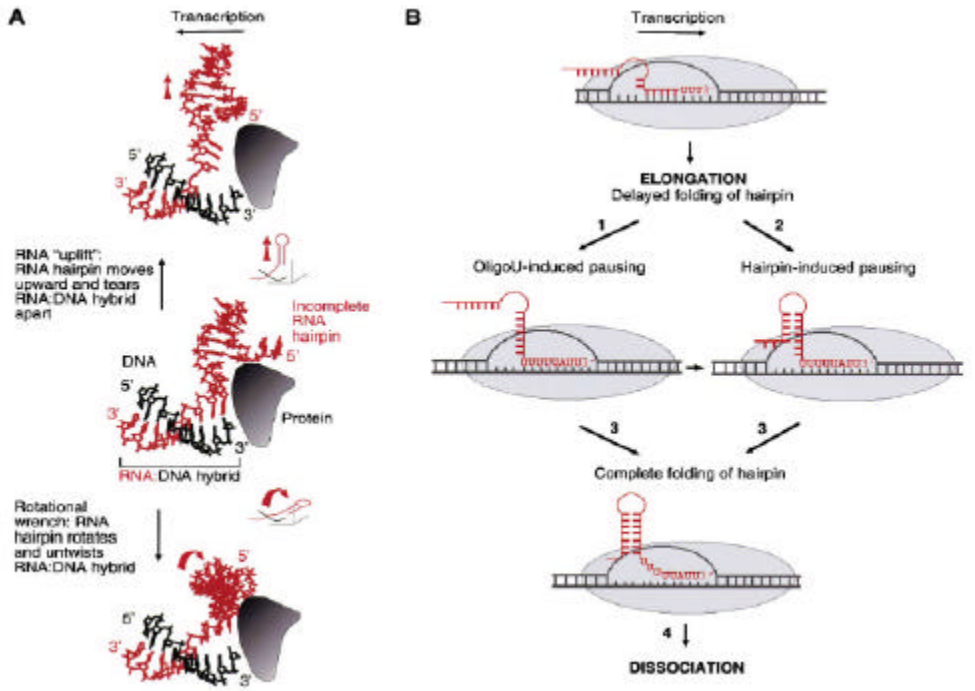


Santangelo and Roberts; 2004

Figure 1.23: Forward Translocation Model of Intrinsic Termination
Hairpin formation drives RNAP and the transcription bubble forward in the absence of further synthesis, unwinding DNA ahead of the catalytic site and rewinding DNA at the upstream edge of the transcription bubble. After translocation of 4–6 nucleotides, only a minimal, weak RNA/DNA hybrid remains. The RNA 3' terminus is removed from the active site, and the ssRNA/RNAP interactions are lost; this combination ultimately results in transcript release. RNAP is shown in the pretranslocated position with the RNA 3' terminus occupying the active site (boxed). The RNA is shown in red, the DNA is shown in black, and RNAP is a gray oval. Hairpin sequences are overscored (DNA) or underscored (RNA).

The allosteric model of termination suggests that once hairpin starts to form, it triggers an alteration in the conformation of TEC in and around the active site, inhibiting nucleotide addition in the active site and reduces affinity for product RNA (Arndt and Chamberlin, 1990; Toulokhonov et al., 2001). Hairpin formation can also disrupt the upstream part of the RNA: DNA hybrid (Kashlev and Komissarova, 2002). Upon hairpin formation, a bulky double stranded RNA has to be accommodated within the confines of the RBS, in a narrow cleft between the ß and ß' subunits, which is usually enough for the single stranded RNA transcript. A change in the arrangement of the nucleic acids within the EC will be required to accommodate the bulky RNA hairpin, as it forms in the EC. These changes involves either the uplifting of hairpin stem away

from the protein and from the long axis of the hybrid or the rotation of the hairpin stem around the long axis of the hybrid to allow hairpin folding, which may result in the unwinding of the hybrid (Figure 1.24). Either of these events may alter the RA:DNA hybrid conformations or its position in the catalytic cleft, affecting its binding to the RNAP. The terminator hairpin induced alteration in the conformation of TEC in and around the active site and the destabilization to the EC brought on by the RNA hairpin and the melting of RNA:DNA hybrid combines to form a series of events leading to the dissociation of elongation complex and the release of nucleic acids at the intrinsic terminator.



Komissarova et al., 2002

Figure 1.24: Hairpin Action during Termination
A. Coupling between the hairpin formation and the melting of the RNA:DNA hybrid in the EC. The RNA and transcribed DNA strand are shown in red and black, respectively. The gray irregular shape symbolizes the RNAP domain(s) proposed to interfere with the folding of the hairpin. Red arrows indicate direction of the movement of the RNA hairpin upon completion of the stem. The motions of the hairpin are also shown schematically. **B**.The consecutive steps in intrinsic transcription termination in *E. coli*. 1.The oligo T tract in the DNA pause transcription in the absence of hairpin. 2. Pausing of the enzyme at the termination point. 3.Hairpin-mediated melting of the RNA:DNA. Hybrid is coupled with the disruption of up to three adjacent RNA:DNA bp. 4.The complex spontaneously dissociates from the unstable 4–5 bp rU:dA hybrid..

1.5.2 Factor dependent (Rho mediated) transcription termination

Rho is an RNA dependent ATPase that binds to cytosine rich unstructured regions in nascent RNA and terminates elongation complexes that are paused at nearby downstream sites (Banerjee et al., 2006; Kato et al., 2003; Richardson, 2002). Rho protein is a 419 amino acid polypeptide, which functions as a homohexamer. Unlike the conserved and highly localized intrinsic termination signals, Rho dependent terminators are bipartite sites that are present over 200 bp of DNA (Richardson and Richardson, 1996). The proximal Rho binding site called *rut* (for Rho utilization), which spans a region of around 70-80 nucleotides, is rich in Cytosine and devoid of secondary structures (Del Giudice et al., 1990; Platt, 1994) and a distal sequence comprising the termination zone. Rho posses RNA-DNA helicase activity and hence can disrupt the RNA-DNA hybrid interactions as well as RANP-hybrid interactions. N modified complexes could also prevent Rho-dependent transcription termination by suppressing RNAP pausing and accelerating Polymerases from Rho bound at upstream RNA sites (Weisberg and Gottesman, 1999).

1.5.2.1 Properties of Rho as a reflection of its structure

The crystal structure of Rho complexed with nucleic acids (Skordalakes and Berger 2006; Skordalakes and Berger 2003) and mutational analysis have specified the following functional domains in a Rho protomer (Figure 1.25). A core RNA binding domain called the primary RNA binding site of Rho that can bind to a single stranded DNA molecule as well as a single stranded RNA molecule held responsible for tethering of Rho to the Rho loading or Rho utilization (*rut*) site extends from residues 22-116 (Modrak and Richardson 1994). P-loop, associated with ATP binding and ATPase activity of Rho spanning from 179-183 is highly conserved among RecA family of ATPases (Opperman and Richardson 1994; Wei and Richardson 2001). Q-loop and R-loop comprises the secondary RNA binding site of Rho. Q-loop is formed by an 8-residue segment within the 278-290 residues (Wei and Richardson 2001a,b; Xu *et al.*, 2002; Skordalakes and Berger 2003). R-loop, a less defined and characterized part of Rho protein, is thought to span between 322-326 amino acid residues in *E. coli* (Skordalakes and Berger 2003). Rho can self assemble in a solution with a variety of assembly states, the homohexamer being the most predominant (Geiselmann *et al.*, 1992). The major factors that influence Rho subunit assembly are the ionic environment, concentration of the Rho protein and the presence of cofactors. Rho protomer assemble into stable dimeric or tetrameric forms under a high concentration of salt, low protein concentration and in the absence of cofactors. The hexameric unit of Rho in the absence of RNA can exist in either an open ‘key washer’ like pentameric form or in closed ‘ring like’ hexameric form. The combinatorial analysis through electron microscopic, hydrodynamic, X-ray

and neutron scattering techniques, the geometry of the Rho hexamer was found to be a regular hexagon (Gogol *et al.*, 1991; Geiselmann *et al.*, 1992a; Geiselmann *et al.*, 1992b).

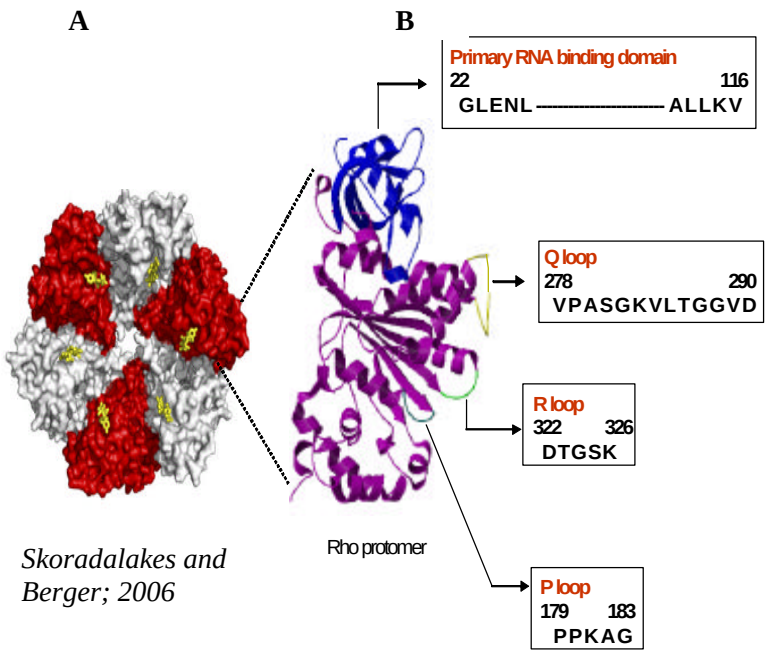


Figure 1.25: Rho structure and functional domains
A. Surface representation of the closed ring Rho hexamer. Alternate Rho subunits are colored red and white, RNA bound to primary RNA binding site is colored yellow.
B. Functional domains of *E.coli* Rho. Cartoon represents an open 'key washer' like pentameric form of Rho, depicting the four defined functional domains: primary RNA binding domain, P loop, Q loop and R loop.

1.5.2.2 Sequence of events in Rho-dependent termination

Transcription termination mechanism of Rho can be expressed grossly in the following steps: Rho binds to naked, non-translating RNA at a C-rich site called Rho utilization site (*rut* site). The *rut* site binding of Rho is achieved by the N-terminal or primary binding domain. The bound RNA, passing through primary binding site reaches the secondary binding site that resides on C-terminal and the RNA bound secondary site sensitize Rho for ATP binding and hydrolysis. The energy derived from ATP hydrolysis powers the translocase / helicase activity of Rho to unwind RNA/ DNA duplex, leading to the release of the RNA transcript from TEC (Banerjee et al., 2006; Figure 1.26).

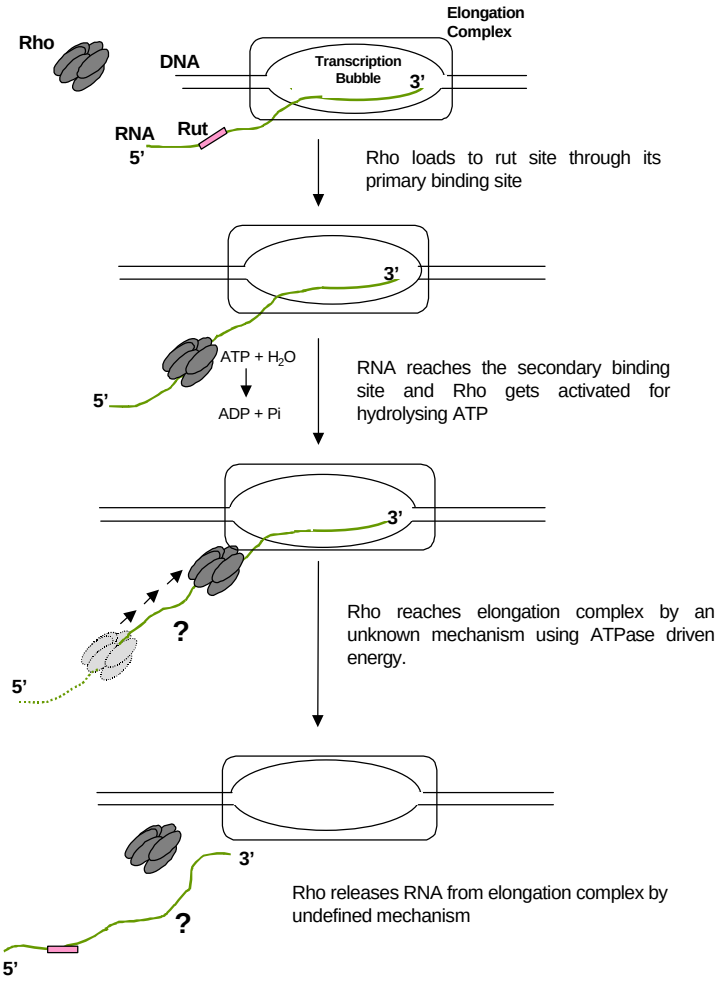


Figure 1.26: **An overview of the steps involved in transcription termination by Rho**.

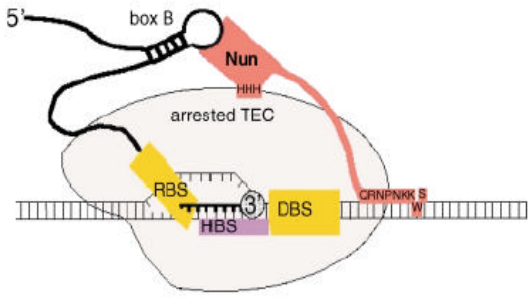
Cartoon representing various steps involved in Rho mediated transcription termination: Rho translocation; interaction with elongation complex; and the unwinding and eventual release of RNA.

Banerjee et al., 2006

1.5.3 HK022 bacteriophage Nun protein and its role in transcription termination

Lambdoid phage HK022 encodes a 109 amino acid protein called Nun, which brings about transcription termination with the help of NUT RNA and the Nus factors (Robert et al., 1987; Hung and Gottesman 1997). The N terminus of Nun protein posses an Arginine Rich Motif (ARM) like that of N protein and binds to ? BOXB (Chattopadyay et al 1995), where as its C-terminal domain (CTD) is different from that of N protein and interacts with NusA, DNA and RNAP. Nun CTD contains three Histidine residues at its C terminus, which forms part of a Zn^{2+} binding motif required for Nun activity (Watnick and Goettesman, 1998). CTD of Nun also posses a Trp residue at position 108, which interacts with DNA. The interaction of the CTD of

Nun to RNAP and DNA blocks the translocation of RNAP, by anchoring it to DNA and brings about transcription arrest and termination. It is believed that the amino terminal ARM motif binds to the BOXB RNA in the nascent transcript, tethering Nun in proximity to the TEC. The Histidine residues in the CTD contacts RNAP in a Zinc dependent manner. The CTD then contacts DNA, possibly by intercalation of the penultimate Tryptophan residue into the DNA template.



Nudler and Gottesman 2002

Figure 1.27: Model for transcription arrest by Nun

The NH2-terminal ARM motif of Nun interacts with BOXB, tethering Nun in proximity to the TEC. The carboxyl-terminus contacts DNA, possibly by intercalation of the penultimate Tryptophan residue into the DNA template

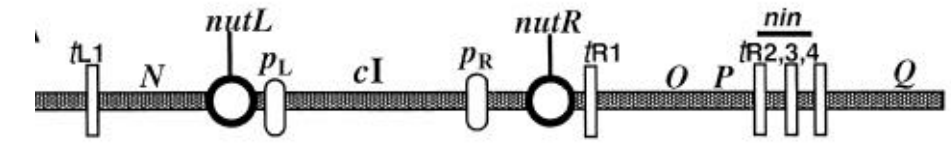
1.6 Transcription antitermination

Transcription antitermination can be defined as the process in which genes whose transcription is otherwise blocked by premature termination are expressed through termination suppression. Antitermination in phage ? and related bacteriophages is brought on by two distinct mechanisms. The first is the result of interaction between ? N protein and its targets in the early phage transcripts and the second is the interaction between the ? Q protein and its target in the late phage promoter pR'. Antiterminator proteins like N binds to specific sequences called NUT in the nascent RNA and along with host elongation factors like Nus factors from *E.coli*, bind to the EC and converts it into a termination resistant form (Greenblatt et al., 1993; Mogridge et al., 1998; Friedman and Court, 1995; Weisberg and Gottesman, 1999).

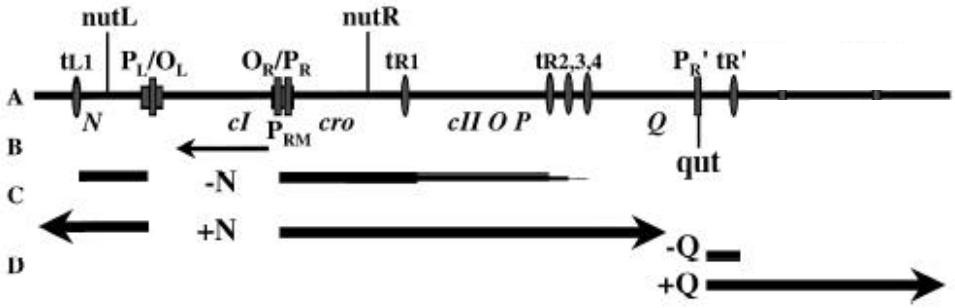
1.6.1 The Lytic- Lysogenic cascade of phage ? and its antitermination machinery

Bacteriophage ? is a temperate phage, which infects *E.coli* and can choose between two lifecycles from then on: A Lytic mode, whereby the phage DNA, after its entry into the host *E.coli* takes over the host cell machinery, synthesize phage proteins using the host resources, assemble them into progeny phage particles and eventually release the progeny phages by the lysis of host *E.coli* cell. In the lysogenic mode, phage DNA enters the host cell, the early genes are transcribed and

translated, and then a ? repressor protein binds to the phage operator regions, shutting down the transcription of all genes except for *cI* (gene coding the ? repressor protein). Here, the phage DNA integrates to the host bacterial chromosome, and exists as a prophage in the host cell, which becomes a lysogen. The phage DNA can replicate along with the host DNA. Under certain conditions, the lysogeny can be broken and the phage can enter the lytic mode, as in the case of prophage induction by SOS response to host bacterial DNA damage (Little, J.W. 1995). The ? genome is divided into three classes called immediate early, delayed early and late genes, which are sequentially arranged in the phage DNA. The interplay between these gene products, according the need of the hour brings about the regulation of the lytic-lysogenic cascade. After the entry of the phage DNA into the host, the *E.coli* RNA polymerase holoenzyme transcribes the immediate early genes, *cro* and *N* present immediately downstream of the rightward and leftward promoters, P_R and P_L, respectively. At this stage in the lytic cycle, no repressor is bound to the operators that govern this promoters (O_R and O_L respectively), hence transcription proceed unimpeded till the TEC reaches the end of early genes and encounters terminators, stopping short of delayed early genes. The *cro* gene product is a repressor that blocks transcription of ? repressor gene *cI*, preventing the synthesis of ? repressor protein. The *N* gene codes for N protein, a transcriptional antiterminator, which helps RNAP to ignore downstream terminators and read through to the delayed early genes (Friedman et al., 1984; Greenblatt et al., 1993; Van Gilst and von Hippel, 1997). One of the delayed gene products, Q, codes for another antiterminator Q protein, that permits the transcription of the late genes from the late promoter $P_{R'}$ and allows transcription into the late genes. The action of these gene products serves as an important switch for the phage to advance into the lytic mode.



Neely and Friedman., 2000



Shi and Friedman 2001

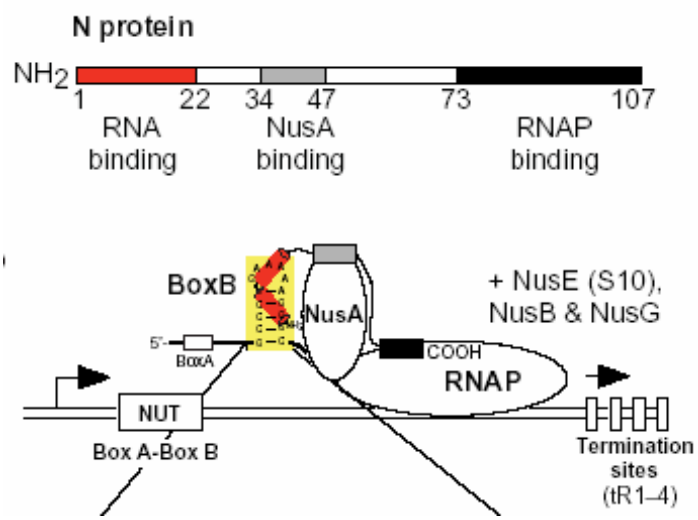
Figure 1.28: Arrangement of the regulatory signals, early genes and transcription patterns in early region of a lambdoid phage H-19 B
The regulatory genes responsible for the lytic-lysogenic cascade like *N, cI, cro Q* and regulatory DNA sequences like *nut*, operators, promoters and terminators are shown.
Arrows represent transcription from early P_L and P_R promoters in the presence and absence of antiterminators N and Q; change in thickness of lines indicates different levels of read through of transcription terminators.

1.6.2 N mediated transcription antitermination

N protein from Lambda and related phages is an 'antiterminator' protein, which can convert RNA Polymerase to a termination resistant form at both intrinsic and Rho-dependant terminators (Friedman and Court, 1995; Weisberg and Gottesman, 1999) during early phases of transcription elongation. Denis Luzzati first showed that N regulates transcription elongation, whereby she demonstrated that the function of N gene product would be to allow transcription to proceed past some sort of a stop signal (Luzzati, 1970). Bacteriophage ? N dependent antitermination is an important model system for the *cis*-regulatory interactions in transcription (Conant et al., 2005; Friedman et al., 1984; Greenblatt et al., 1993; Van Gilst and von Hippel, 1997). TECs contain factors that modify the rate of transcription elongation and the efficiency of transcription termination. The association of some of these factors with RNAP is promoted by sequences in the

template or in nascent transcript. An RNA sequence called NUT in lambdoid phages located close to the P_L and P_R operons promotes the formation of an anti-termination complex consisting of *E. coli* host transcription factors like the Nus proteins and phage protein N that binds specifically to the lambda NUT RNA (Greenblatt et al., 1993; Mogridge et al., 1998a).

N is a small basic protein of the Arginine Rich Motif (ARM) family of RNA binding proteins and is an example of intrinsically unstructured proteins regulating transcriptional modulation of RNA Polymerase (Fink A.L., 2005; Van Gilst and von Hippel, 1997). ? N is largely unfolded in solution (Mogridge et al., 1998a; Van Gilst and von Hippel, 1997) and consists of three functionally distinct regions, binding to different interaction partners. The Amino terminus of N binds to *nut* RNA, the middle portion binds to NusA and the C terminus binds to RNA Polymerase (Greenblatt et al., 1993; Mogridge and Greenblatt, 1998; Prasch et al., 2006).



Xia et al., 2003

Figure 1.29: Cartoon showing structure and function of N protein from bacteriophage ?
A. Domain structure of Lambda N. Regions for *boxB* RNA binding (1-22), NusA binding (residues 34-47) and RNA polymerase binding (residues 73-107) (Mogridge et al., 1998a) are shown.
B The Lambda antitermination complex with N bound to the *boxB* mRNA, NusA and to RNA polymerase, moving along DNA with intrinsic termination sites.

The N protein binds to a 15-nucleotide (nt) stem-loop RNA called BOXB (Barik et al., 1987; Chattopadhyay et al., 1995) and acts via an RNA looping mechanism to contact the elongating RNAP at downstream terminators (Barik et al., 1987; Nodwell and Greenblatt, 1991) .The NTD of N protein is conserved in various N proteins (Figure 1.29), with the Arginine Rich Motif (ARM) composing of amino acids RRR or RR, being universally present. The ARM binds to the *nut* RNA, present in the nascent transcript. The CTD of N protein, which binds to RNA polymerase, is poorly conserved among N proteins from different phages, compared to its N terminus, though there is a remarkable similarity between the CTDs of H-19B N, HK97, Ni12 and P22 (Figure 1.30). The terminator specificity of N protein from Lamdoid phages

depends on the *nut* (N-utilization) DNA sequences (Rosenberg and Court, 1979; Salstrom and Szybalski, 1978; Warren and Das, 1984) which are located close to the early promoters. Transcription of *nut* produces the RNA regulatory elements BOXA and BOXB, which function as protein binding sites within the nascent transcript (Chattopadhyay et al., 1995; Nodwell and Greenblatt, 1991).

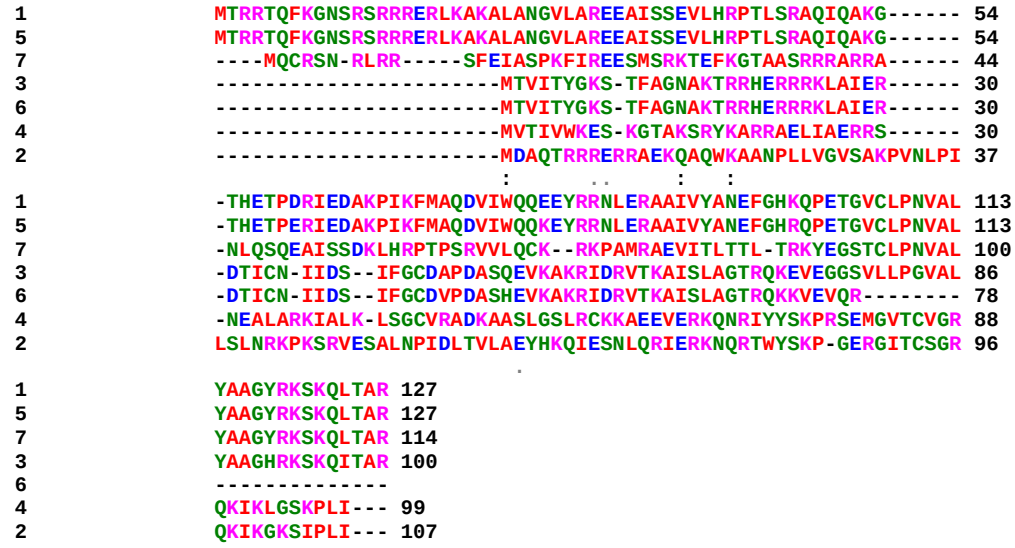


Figure 1.30: CLUSTAL W multiple sequence alignment of N proteins from different bacteriophages

The numbers given denotes N proteins from bacteriophages as follows. H19-B>**1**; Lambda>**2**; P22 >**3**; Phi-21> **4**; HK97>**5**; Phage L>**6**; Bacteriophage Nil2 >**7**

1.6.2.1 *Nut* Sequences coded by Lamdoid phages

The BOXA and BOXB RNA sequences together constitute the *nut* (for N utilization) site, containing all the *cis* acting elements required for N-dependent antitermination. The ability of newly transcribed *nut* RNA to bind to N protein-RNAP-host elongation factor complex and change the expression of downstream terminators, are an essential component of the genetic switch from lysogeny to lytic modes of development in the lambdoid phages. The interaction of N with BOXB is also phage-specific, as N proteins of related lambdoid phages, which are homologous, recognize only their cognate BOXB sequences (Barik et al., 1987; Harada et al., 1997). Recognition of BOXB by N thus ensures the correct location, timing, and phage specificity of lambda development. N protein interacts with free RNAP or RNAP in elongation complexes through its CTD (Mogridge et al., 1998a). ? N scans the nascent RNA for the *nut* BOXB signal and binds to it, forming a processive antitermination complex (Barik et al., 1987;

Chattopadhyay et al., 1995; Mason and Greenblatt, 1991; Mogridge et al., 1995; Mogridge et al., 1998b). Although the *boxB* sequence is not well conserved in other bacteriophages in lambdoid family, most of the phages encodes proteins that are analogous to ? N and have sequences capable of forming *boxB* like structures in their pL and pR operons, which are recognized by their cognate N analogs, which accounts for the phage specificity of N-mediated antitermination (Weisberg and Gottesman, 1999). *In vivo*, antitermination appears to be sensitive to cellular N protein levels as well, as the phage specificity of antitermination is overcome by over expression of N protein (Franklin and Doelling, 1989).

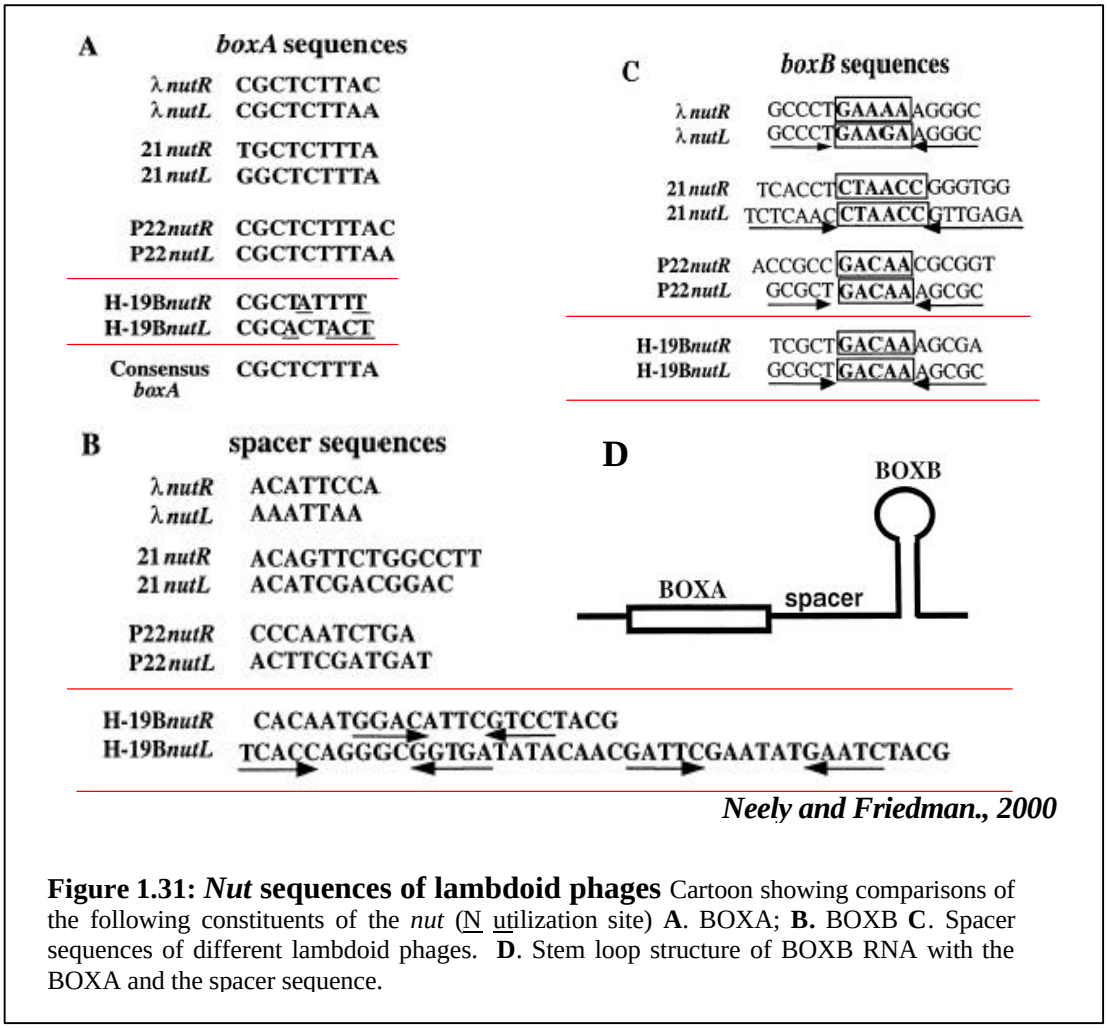


Figure 1.31: *Nut* sequences of lambdoid phages Cartoon showing comparisons of the following constituents of the *nut* (N utilization site) **A**. BOXA; **B.** BOXB **C**. Spacer sequences of different lambdoid phages. **D**. Stem loop structure of BOXB RNA with the BOXA and the spacer sequence.

1.6.2.2 Host elongation factors and their roles in N mediated transcription antitermination

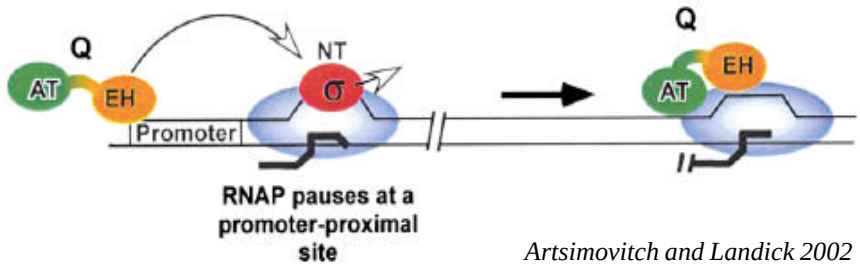
Along with ? N, four other proteins from the *E. coli* host- NusA, NusB, NusG and S10 collaborate in antitermination at the ? immediate early terminators (Friedman, 1971; Greenblatt

and Li, 1981a; Friedman et al., 1984; Li et al., 1992). These proteins do not possess antitermination activity in the absence of N, but they do form stable complexes with N and RNAP (Mason and Greenblatt, 1991; Mogridge et al., 1995) and increase the activity of N at terminators far downstream from *nut* (DeVito and Das, 1994; Mason and Greenblatt, 1991). The NusA protein binds both N and RNAP (Nie et al., 2003; Van Gilst and von Hippel, 1997).The first evidence that ? N mediated antitermination has specific host factor requirements derived from experiments characterizing *E.coli* variants with mutations that cause failure of the bacterium to support N action, which came to be called as Nus (N utilization substance) phenotype (Friedman and Court, 1995; Friedman et al., 1984). ? N protein interacts with the sequence element within the *nut* RNA called BOXB (Chattopadhyay et al., 1995) and the NusB-NusE complex interacts with the BOXA RNA while NusA and S10 bind to RNA polymerase (Nodwell and Greenblatt, 1991; Nodwell and Greenblatt, 1993; Rees et al., 1993). Pair wise interactions between these proteins like N and NusA (Greenblatt and Li, 1981b), NusA and the core Polymerase (Gill et al., 1991; Greenblatt and Li, 1981a), NusG and RNAP (Li et al., 1992), NusB and NusE (Mason and Greenblatt, 1991) and NusE and RNAP (Mason and Greenblatt, 1991) has been demonstrated *in vitro*. The host elongation factor proteins do not possess antitermination activity in the absence of N, but forms stable complexes with N and RNAP and facilitate antitermination at terminators downstream from the *nut* sequence (processive antitermination). N protein is the active factor in an antiterminating TEC, as it has been shown that it in sufficient concentrations, N suppresses transcription termination in vitro in the absence of *nut* or the Nus factors (DeVito and Das, 1994; Mason and Greenblatt, 1991).

1.6.3 Q mediated antitermination

Q protein is a ? phage encoded factor, which brings about antitermination in the ? late operon. The ? phage encoded Q antiterminator protein as well as the Q proteins of related phages including F 21 and F 82 are stably incorporated into RNAP during a promoter proximal s^{70} dependent pause (Ko et al., 1998; Ring et al., 1996; Yarnell and Roberts, 1992). Q modified complexes have increased elongation rates, diminished pausing and resistance to all downstream intrinsic and Rho dependent termination sites. Q protein binds specifically to two DNA elements immediately upstream of $P_{R'}$ in the ? genome- a Q binding element (QBE) located between $^{-}10$ and the $^{-}35$ elements of $P_{R'}$ and a sequence located just downstream of the transcription start site, which brings about a pause to the transcribing RNAP (pause inducing sequence; Ko et al., 1998) RNAP initiates transcription at the $P_{R'}$ and pauses after synthesizing a short transcript (16 or 17 nucleotides long) at the pause site, allowing promoter clearance and Q protein binds to the QBE.

The ?Q bound to the DNA interacts with the paused elongation complex and joins the TEC, helping it to exit the pause more rapidly and making it resistant to downstream terminators. The pausing of RNAP at +16 is induced by the interactions between s^{70} and a downstream repeat of a $^{-}10$ promoter element (Ring et al., 1996). Transcript clevage factors GreA and GreB helps the TEC to pass through the pause, by clevage and resynthesis of the 3' nucleotides of the paused +16 transcript .The presence of s^{70} subunit of RNAP holoenzyme is essential for pausing and Q mediated antitermination; it has been shown that ?Q - s^{70} interaction is required for Q-TEC interaction and engagement (Nickels et al., 2002). Once Q protein interacts with the TEC, s^{70} is no longer required for stable association of Q with the elongation complex. It is believed that Q protein suppresses termination by directly interfering with hairpin formation at intrinsic terminators.



Artsimovitch and Landick 2002

Figure 1.32: Cartoon representing a Q modification of an EC Blue sphere is RNAP, The two function of Q protein is represented as AT (antitermination module) and EH (elongation enhancement module).Q transfers from its DNA binding site adjacent to the promoter to promoter proximal paused TEC.s^{70} then dissociates and Q establishes a tight interaction with RNAP and helps it antiterminate past terminators.

RNA hairpin of an intrinsic terminator disrupts the 5' terminal bases of the RNA:DNA hybrid and removes the 3' end of RNA from active site. Q modification is postulated to stabilize the interactions between RNA and the RNA-DNA hybrid, preventing the disruption of the upstream portions of the hybrid required for the formation of the terminator hairpin. Modification of the EC by Q also helps in maintaining an optimal alignment of the active site, by maintaining the 3' end of RNA in the active site (Santangelo et al., 2003). This along with reduction in pausing at the terminator by the stabilization of the elongation competent configuration helps the Q modified EC to read through terminators and synthesize late genes.

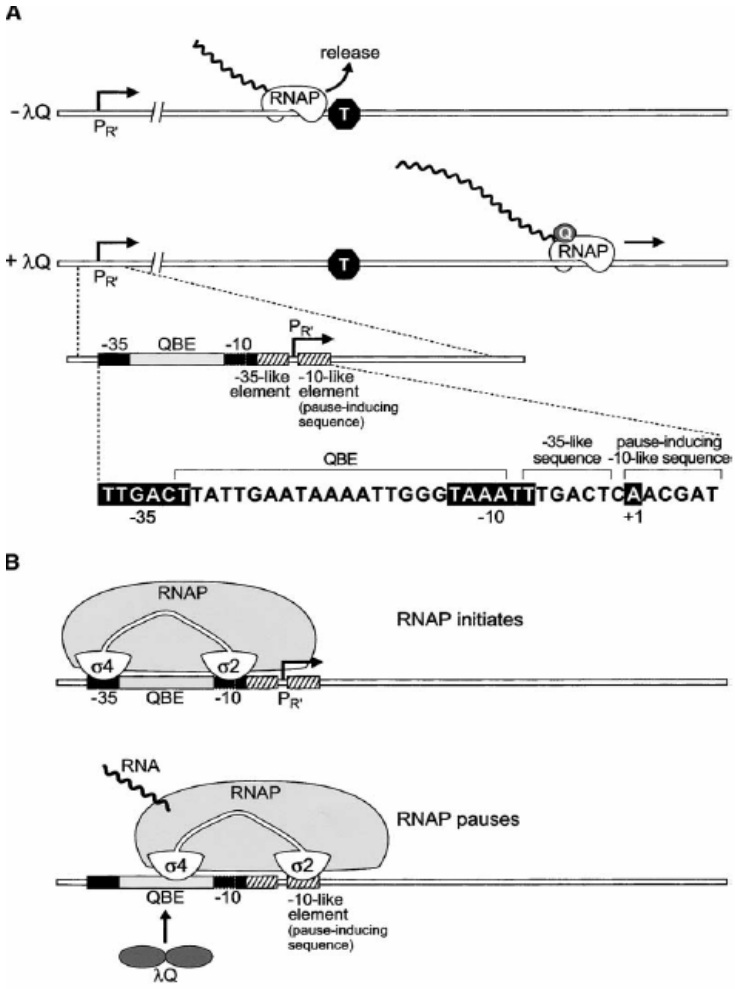


Figure 1.33: Pictorial representation of Q interactions with EC

A. Functionally important elements at $P_{R'}$. Promoter ¯10 and ¯35 elements, the ?Q binding element (QBE), the –35 like element and the pause inducing –10 like element are depicted.
B.Q mediated antitermination. The paused TEC at $P_{R'}$, is formed by the interactions between s 70 and the ¯10 like element. Q protein (shown as a dimer), binds to QBE and engages the pause Elongation Complex

1.6.4 Transcription antitermination by PUT RNA

Phage HK022 prophage expresses a protein similar to N, called Nun, which binds lambda NUT and induces transcription termination (Robert et al., 1987). It can also regulate gene expression by suppression of transcription termination. The early gene expression relies on a unique mechanism by which RNAP can associate with a non-protein, auxiliary factor like an RNA moiety, leading to transcription antitermination. In this case, a highly structured RNA, called PUT (for polymerase utilization), interacts with the ß'subunit of RNAP, enabling it to read through terminators in an analogous fashion to the lambda N antitermination (Banik-Maiti et al., 1997; King et al., 1996).

Cis acting structured RNA sequences in the nascent transcripts from the HK022 phage early promoters pL and pR operons called PUTL and PUTR respectively interact with RNAP and modify it into a termination resistant form (Banik-Maiti et al., 1997; King et al., 1996; Sen et al., 2001). Like the ?N and Q proteins, the HK022 PUT RNA can also suppress Rho dependent terminators (King et al., 1996; Yang et al., 1989; Yang and Roberts, 1989). PUTL and PUTR form stem loop structures separated by an unpaired nucleotide and interacts with the Zn-finger motif in the ß' subunit of RNAP. The 3' PUT stem interacts with RNAP in TEC, while the 5' stem loop is believed to promote this interaction. An RNA-RNAP interaction in the form of a persistent interaction between stem-loop 2 of *put* RNA and the transcript elongation complex is required for antitermination (King et al., 2004; King et al., 2003; Sen et al., 2002; Sen et al., 2001).

1.7 Bacterial antitermination

1.7.1 RfaH protein mediated antitermination

RfaH is a transcriptional regulator protein that increases the expression of several bacterial operons, especially for the expression of certain extracytoplasmic macromolecules, including pathogenicity factors like hemolysin. RfaH (a NusG homologue) and a *cis* acting promoter proximal sequence element, *ops* (for operon polarity suppressor) increase the expression of several bacterial operons, by suppressing termination. RfaH gets recruited to the TEC at the *ops* site, where it delays the escape of a fraction of RNAPs (RfaH interacts with the *ops* sequence in the non-template DNA strand and the core of RNAP) and can act as an antiterminator and as an elongation enhancer and can affect transcription up to 20kb downstream (Artsimovitch and Landick, 2002; Bailey et al., 1996; Bailey et al., 2000; Leeds and Welch, 1996; Stevens et al., 1997). Although RfaH delays escape from the *ops* pause, once the escape occurs, RfaH enhances elongation by suppressing pausing and inhibiting Rho dependent as well as intrinsic termination, without apparent involvement of other accessory proteins.

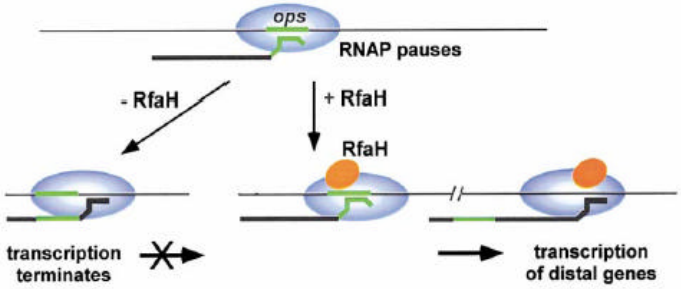


Figure 1.34: Model of RfaH action.

RfaH modifies TEC to a termination resistant state, whereas in the absence of RfaH, the transcript is released at the terminator downstream of *ops* site.

1.7.2 Antitermination in ribosomal operons

Examples of processive antitermination are also found in the bacterial operons. During synthesis of rRNA in *E.coli*, RNA Polymerase undergoes several *rrn* operon specific modifications that increases transcription elongation rates and allow read through of Rho dependent terminators (Albrechtsen et al., 1991; Condon et al., 1995; Vogel and Jensen, 1995; Zellars and Squires, 1999) The modifications require specific nucleotide sequences in the leader and spacer regions of the *rrn* operons as well as different host elongation factors and are regulated by an antitermination mechanism. The rRNA operons of many bacteria have a consensus *boxA* sequence and a stem-loop structure that structurally resembles the *boxB* sequence of lambda and a *boxC* sequence in the leader sequence (Berg et al., 1989; Link, 1984) and *boxA* and *boxB* located promoter proximal in the spacer region between the 16S and 23S structural genes in each operon (Aksoy et al., 1984; Condon et al., 1995). These sequences are related to those used in bacteriophages lambda N protein antitermination system and also include most of the host factors involved in lambda transcription antitermination (Greenblatt et al., 1993). Unlike in the N mediated antitermination system, in the rRNA system only factor-dependent (Rho protein) terminators are efficiently neutralized (Albrechtsen et al., 1991; Albrechtsen et al., 1990; Berg et al., 1989). The fact that intrinsic or factor-independent terminators remain immune to rRNA antiterminated transcription (Aksoy et al., 1984), points to differences between the rRNA and bacteriophage lambda mediated transcription antitermination systems. The sequences of the *rrn* BOXA (a conserved 12 nucleotide sequence) are more similar to the bacteriophage consensus than that of ? and they can bind NusB and S10 more efficiently (Nodwell and Greenblatt, 1993). Point mutations in *boxA* induce premature transcription termination. BOXA in conjunction with host elongation factors like NusA, NusB, NusE (ribosomal protein S 10), NusG and ribosomal protein S4 brings about antitermination in *rrn* operons and increases the rate of elongation by RNA Polymerase (Quan et al., 2005). *BoxB* is a conserved stem loop structure and is situated promoter proximal to the *boxA* sequences and was found not to be essential for antitermination.

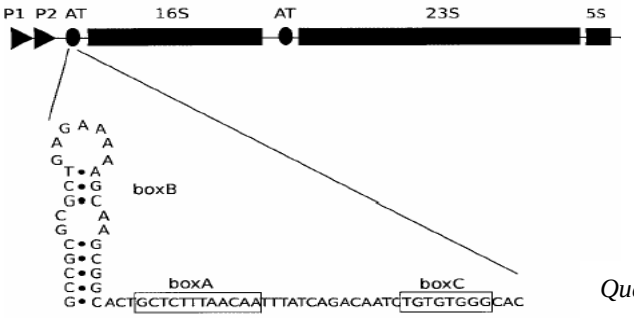


Figure 1.35: DNA regulatory sequences in rRNA operon
DNA sequences of boxA and boxC sequences are shown inside the boxes in the cartoon; the stem loop structure of boxB is also shown.

Quan et al., 2005

CHAPTER II

Objectives

2.1 Introduction

Transcriptional antitermination is the process in which genes whose transcription otherwise blocked by premature termination are expressed through termination suppression. Antiterminators like N in ? and in related lambdoid phages convert RNA polymerase into highly processive, termination resistant form during early phases of transcription elongation. Though there has been a tremendous increase in the knowledge by way of biochemical and genetic information over the past three decades on the subject of N mediated transcription antitermination in lambdoid phages, the exact molecular mechanisms of it remains unknown. Question like where N protein binds to RNA polymerase, what is it that it does to RNAP in consort with transcription factors which makes the RNAP to ignore and read through the strong terminators (both factor independent and dependent) downstream forming a very stable and processive antitermination complex, remains unanswered. In this work, we have attempted to try and address these questions. We have used the N protein mediated transcription antitermination system from a Lambdoid bacteriophage, H-19B in this study.

2.1.1 H-19B bacteriophage

Shiga toxin producing *Esherichia coli* (STEC) are pathogens that can cause Diarrhea (Smith et al., 1983, 1984; Scotland et al., 1983; Huang et al., 1987; Head et al., 1988). Stx encoding lamdoid bacteriophage H-19B was first isolated from *E. coli* 026:H11 STEC strain H19, obtained in 1967 from an infant with diarrhea (Smith and Linggwood, 1971). It was demonstrated (Huang et al., 1986) that structural genes encoding the Shiga like toxin were encoded in the H-19B genome. H-19B was identified as a relative of bacteriophage ? (lambdoid phages) through restriction mapping and hybridization studies with ? DNA (Huang et al., 1987). Analysis of H-19B DNA sequence suggests a transcription antitermination system resembling the N-mediated systems of most other lambdoid phages (Franklin, 1985a; Franklin, 1985b). The H-19B N gene encodes a 127 amino acid protein that contains an Arginine rich motif, as found in N protein of other lambdoid phages (Neely and Friedman, 1998a; Neely and Friedman, 1998b).

2.1.2 Why chose H-19B for transcription antitermination study?

The H-19B N mediated antitermination system is characterized by its reduced need for host Nus factors, compared to ? N antitermination. H-19B phage was found to have no growth problems in *E.coli* strains having *nus* mutations in *nusB5*, *nusB:IS 10* and *nusE71* mutants as well as in the *nusB*/*nusE71* double mutant, whereas ? growth was not supported in any of these mutants, under restrictive temperatures. This led to the conclusion that H-19B does not require ribosomal protein S-10 or NusB for N mediated antitermination. It has also been shown in the study that H-19B failed to grow in the *NusA1* mutant, which points out that NusA is essential for antitermination

and in turn, lytic growth of H-19B phage (Neely and Friedman, 2000). From all these, it can be concluded that H-19B N requires NusA whereas it does not require NusB or ribosomal protein S10 for processive antitermination. Comparisons of *nut* sites (as depicted earlier in Figure 1.26; Neely and Friedman 2000) shows that the H-19B *nut* sites have similarities as well as differences from *nut* sites of other characterized lambdoid phages (Freidman and Gottesman, 1983; Franklin, 1985a). They include variations in the *boxA* consensus: H-19B *nutR* has a C to A nucleotide substitution in the highly conserved fifth position in *boxA* (Figure 2.1). This position is functionally significant, as N mediated antitermination was reduced when the ? *nutR boxA* sequence had this change (Lorenzi et al., 2006; Patterson et al., 1994). The H-19B *nutL boxA* has very less homology to the consensus *boxA* sequence of other lambdoid phages and also is different from the *boxA* sequence of its own *nutR* (Neely and Friedman, 1998a). Another important difference found in the H-19B *nut* regions is the unusually long spacing between *boxA* and *boxB* sequences compared with that seen in *nut* regions of other phages. The spacer regions are also important for antitermination as changes in the spacer sequence of ? nut regions influences N-mediated antitermination (Doelling and Franklin, 1989; Peltz et al., 1985; Zuber et al., 1987).

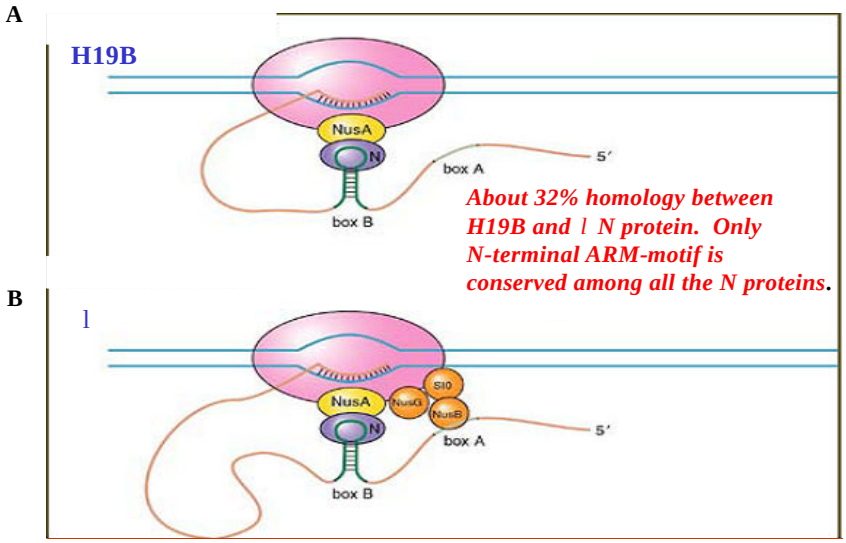


Figure 2.1: Differences between H-19B and ? phage transcription antitermination
A. H-19B N mediated transcription antitermination. The TEC consists of RNAP, H-19BN, NusA and BOXB RNA. **B**. ? N mediated transcription antitermination. TEC in addition to RNAP, ?N, NusA, BOXB also consists of other host proteins and BOXA.

The spacers in H-19B are significantly longer than other Lambdoid spacer regions, having 22 nucleotides in *nutR* and 42 nucleotides in *nutL* compared with the 8-14 nucleotides found in characterized *nut* sites (Neely and Friedman, 2000). The H-19B *boxB* sequences are typical and

have hyphenated dyad symmetries with potential of forming stem loop structures in the RNA. The *nutL* and *nutR* sequences are identical in the potential loop region. Mutational studies were carried out whereby the nucleotide sequences in the highly conserved, functionally important consensus sequences in *boxA* were changed and the effects of this as well as the effects of a *boxA* deletion and its effects on antitermination *in vivo* were studied. None of these mutations, including the *boxA* deletion had any significantly adverse effect on transcription antitermination rates in H-19B. Similar studies in *boxB* with nucleotides in the ascending arm of *boxB* dyad symmetry were changed (eliminating the formation of a stem-loop in the BOXB element in transcript RNA) showed that H-19 B *nut* regions with such changes in *boxB* do not support antitermination (Neely and Friedman, 2000). In H-19B, the spacer stem loop through its sequence and structure may substitute for the Nus factors that are not required while the *nutR* BOXB stem loop structure is very much necessary, for H-19B N mediated transcription antitermination (Friedman and Court, 2001; Juhala et al., 2000; Neely and Friedman, 1998a; Neely and Friedman, 2000; Yang et al., 2003). The amino terminus of H-19B N protein contains an Arginine Rich Motif (ARM) similar to the ARM found in its ? N counterpart, which binds to a small stem loop structure in the *nut* RNA, known as BOX B. During transcription of the H-19B early genes; as the BOX B stem loop mRNA emerges, N binds on to it and becomes tethered to the elongating RNA polymerase. It is believed that the BOXB bound N protein contacts RNA polymerase through RNA looping, converting the RNA polymerase to a termination resistant form. This stable association of RNA polymerase to N-BOXB- NusA complex makes it to form a stable Ternary Elongation Complex (TEC) which can read through downstream terminators. The reduced requirement of host factors for the antitermination system, which helps one to design simpler experiments both *in vivo* and *in vitro*, made us choose H-19B as our model in this study.

2.1.3 Questions addressed and experimental approach

The two major questions that we have tried to address in this study are:

1. What is the binding site of N protein on EC?
2. What is the mechanism of N mediated transcription antitermination?

Recent crystal structures and structural models of free RNAP and the EC reveals that major RNAP-nucleic acid interactions of EC occur in the ß (*rpoB*) and ß' (*rpoC*) subunits (Zhang et.al 1999; Korzheva et al., 2000; Gnatt *et.al* 2001; Korzheva et al., 2001). We hypothesized that an antitermination process may involve alterations of RNAP-nucleic acid interactions and that the likely target of the antiterminator N protein could be in these two major subunits of RNAP. For the determination of N interaction site on RNAP, we attempted to obtain mutations in ß and ß' subunits of RNAP, which makes the EC specifically defective for N mediated antitermination.

For this, we carried out random mutagenesis of the ß and ß' subunits of RNAP and screened and characterized of such N specific *rpoB/C* mutants, both *in vivo* and *in vitro*. The location of such mutations in RNAP would give an insight of the interacting surface of H-19B N on EC. To confirm the specificity of these RNAP mutants for N mediated antitermination, we also attempted to generate allele specific suppressors in H-19B N for the *rpoB/C* mutants obtained and characterize them. We tried biochemical studies for the determination of binding surface of N on RNAP; which included footprinting of a functional 'roadblocked EC', which was modified/unmodified with H-19B N. Limited proteolysis by Trypsin, localized Fe^{2+} induced footprinting for active site conformational changes, Fe-EDTA protein footprinting and footprinting using a chemical nuclease called Fe-BABE (p-bromoacetamidobenzyl-EDTA), specifically conjugated to the single Cysteine residue at the C-terminus of H-19B N protein were carried out for identifying the N binding surface on the EC.

The stability of an elongating complex can be attributed to the interactions between RNAP and the RNA-DNA hybrid, as well as between polymerase and downstream DNA template (Nudler et al., 1996; Sidorenkov et al., 1998). During elongation while the 3' end of the nascent RNA is engaged in the active site of RNAP, the following 8-9 nucleotides are hybridized to the to the template DNA (HBS) and the next ~9 nucleotides remain closely associated with the exit channel of RNAP (RBS). The strength of the RNA-DNA hybrid assures the lateral stability of the TEC. Intrinsic terminators consist of a GC rich RNA hairpin stem immediately followed by a short Uracil rich segment within which termination can occur. The weakness of base pairing in the RNA-DNA hybrid between rU and dA destabilizes the RNA-DNA hybrid in the U rich segment. This along with the pausing of RNAP at the terminator and the destabilization of the protein-nucleic acid interactions due to the formation of the terminator hairpin contributes to transcription termination. Antiterminator proteins like N could prevent these destabilizations in the TEC, especially that of the RNA-DNA hybrid and reduce pausing at terminator sites due to increased elongation rates, which helps the TEC to overcome the terminator signal. We tried to shed light on the molecular mechanism of N mediated transcription antitermination by designing experiments to see if N protein does come near the RNAP active site in an N modified EC and whether the N modified EC can over come effects of Streptolydigin and Tagetoxin, two antibiotics whose binding sites are close to the active site (Temiakov et al., 2005; Tuske et al., 2005; Vassylyev et al., 2005). We assayed the effects of N modification on the stability of an elongation complex, stalled at a terminator. We also addressed the state of RNA hairpin in an N modified EC, to check if N modification disrupts hairpin formation or not. Assays using Class I (*his* pause) and Class II (*ops* pause) pause sequences (Artsimovitch and Landick, 2000) were

carried out on appropriate DNA templates to see the effect of N modification on the paused EC, if any. We checked if N has any effect on the backtracking of RNAP on such paused templates and the effect of N modification on such backtracked TEC complexes by checking for its sensitivity to GreB and PPi.

CHAPTER III

Screening for H-19B N specific RNA polymerase mutants, their suppressors in N and characterization of the mutants

3.1 Introduction

RNA Polymerase mutations in the ß subunit (Georgopoulos, 1971; Ghysen and Pironio, 1972; Jin et al., 1988a; Sternberg, 1976; Jin et al., 1988; Jin et al., 1988b) and the a subunits of RNA Polymerase (Obuchowski et al., 1997; Schauer et al., 1996; Szalewska-Palasz et al., 2003; Weilbaecher et al., 1994), defective in supporting the growth of ? phage and as well as impaired for N mediated transcription antitermination has been reported. In spite of these reports, a concerted effort to find the interacting surface of N on RNAP by the isolation of N mediated antitermination defective RNAP mutants have never been attempted. N interacts with the core RNAP subunits (a, ß and ß‘) and locks it into an antitermination mode. The NUT-bound N, together with different Nus factors of *E.coli* binds to EC and converts it to a termination resistant state (Greenblatt et al., 1993; Das., 1993; Mogridge et al., 1998). The C-terminal half of N protein is believed to interact with RNAP, while the N-terminal ARM motif recognizes the *nut* site. C terminal domain (CTD) of N is required for transcription antitermination (Gusarov and Nudler, 2001). The RNA binding motif, called the ARM motif, is well conserved among different N proteins. The C-terminal half is poorly conserved and the N protein-binding domain of RNAP has not been well characterized (Mogridge et al., 1998; Weisberg and Gottesman, 1999). In order to understand the mechanism of transcription antitermination, it is essential to define the binding surface of N on RNAP, which has surprisingly been ignored so far and the exact binding surface of N protein on the EC is not yet known. Recent crystal structures and structural models of free RNAP and the EC reveals that major RNAP-nucleic acid interactions occur in the ß (*rpoB*) and ß’ (*rpoC*) subunits (Zhang et.al 1999; Korzheva et al., 2000; Gnatt *et.al* 2001; Korzheva et al., 2001). We hypothesized that an antitermination process may involve alterations of the nucleic acid interactions with the EC and that the likely target of the antiterminator N protein could be in these two subunits of RNAP. We devised a genetic screen for isolating the N specific mutants in RNAP as well as to characterize them, in an effort to find the interacting partners on N in RNAP. We performed random mutagenesis of the *rpoB* (ß) and *rpoC* (ß’) genes of *E.coli* and screened for down mutations specific to N function. We have used the H-19B N mediated N antitermination system for its reduced requirement of host Nus factors as described early, whereby the complexity of *the in vivo* and *in vitro* characterization of the mutants obtained were reduced considerably.

3.2. Materials and Methods

3.2.1 Construction of the background strains for mutant screening

The first objective in the design of the strain for filtering out N specific mutants was to incorporate an H-19B N specific *in vivo* reporter cassette. The background *E.coli* strain selected was K3093 (K37 sup° Tn5::*lacZ lacI* $_{Q1}$; Neely and Friedman, 1998), which we will refer as

RS20 in this thesis. A single copy lysogen (*pLac*-H19B- *nutR* –3T-*lacZYA*) from pK8641 (Neely and Friedman, 1998) was transduced into RS20 using ?RS45. (Refer materials methods for ? transduction). Similarly, a single copy lysogen of (P_{lac}-H19B-*nutR*–ΔT) from pK8628 (Neely and Friedman 1998) was introduced into K3093 (this strain would act as the control in the *in vivo* assays for antitermination as the absence of terminators will result in constitutive expression of the Beta Galactosidase enzyme) and the strain was subsequently designated as RS62. To annul the effect of the chromosomal allele of *rpoB/C* genes, we made this alleles temperature sensitive at a non-permissive temperature (42 °C). RS31 and RS62 strains were made *ts* for *rpoC* allele by moving the *rpoC* 120 (rif^{s}) *btuB*::Tn10 mutation from the strain DJ354 (Din Jin) through P1 transduction (refer materials methods) and the resultant strains were named as RS36 and RS68 respectively. Similarly, strains RS31 and RS62 were made *ts* for the *rpoB* allele by introducing *rpoB ts* 3401 (Ding Jin et.al, 1988) by P1 transduction as discussed earlier and designated as RS80, RS81 respectively. These strains grew in the permissive temperature of 30 °C, while they could not grow at 42 °C. For the supply of H-19B N protein in *trans* in these strains, all these strains were then transformed with the H19BN encoding plasmid pK8601 (Neely and Friedman, 1998). The resultant strains were named RS38, RS69 for *rpoC ts* 3T and ? T background strains respectively; *rpoB ts* 3T and ? T background strains were named RS82 and RS83 for 3T and ? T background strains respectively. RS38 and RS82 served as the background strains for screening of N specific RNAP mutants. Plasmid encoding wild type *rpoC* (pBAD18M; Sen et al., 2002) and *rpoB* (pHYD534) (refer materials methods) was transformed into the respective background strains and were found to fully complement for from the reporter system and was blue in color on X Gal IPTG plates, when grown at 42 °C. Fig 3.1 depicts a schematic representation of the background strain construction.

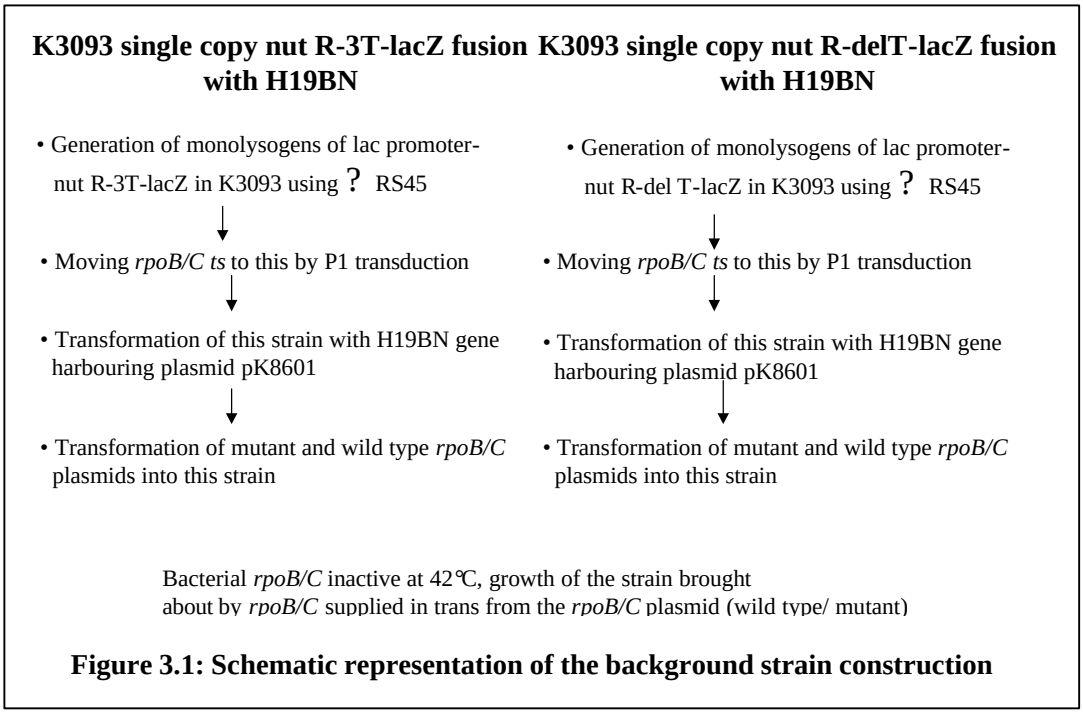


Figure 3.1: Schematic representation of the background strain construction

3.2.2 Preparation of lambda (λ RS45) lysate

The bacterial cells used for infection were grown in LB broth containing 0.4% maltose and 10 mM $MgSO_4$. A λ-sensitive host (MC4100) was grown overnight in 3 ml of LB containing 0.4% maltose and 10 mM $MgSO_4$. To 1 ml of this culture, 0.1 ml of phage lysate (or to a titer value ~10^9 pfu/ml) was added in the presence of 5 mM $MgSO_4$ and after 20 min of infection, the mixture was diluted with 10-20 ml of LB containing 5 mM $MgSO_4$. This was grown with shaking until lysis occurred after ~ 4-6 hours. Once visible cell lysis occurs, this lysate was treated with chloroform (0.3 ml/10 ml), shaken well, transferred into a sterile glass test tube, centrifuged at 4000 rpm. After centrifuging down the debris, the clear supernatant was removed into a sterile test tube, treated again with chloroform (0.3 ml), spun at 4000 rpm and the clear lysate was stored at 4 °C with Chloroform. The lysate thus obtained usually contained 10^9 pfu/ml phage particles. To quantitate the lysate preparation, a titration was done using a sensitive indicator strain such as MG1655. 100 µl each of serial dilutions of the λ RS45 phage made (typically 10^{-5}, 10^{-6}) were mixed with 0.1 ml of the fresh culture. After 15 min adsorption at 37 °C without shaking, each mixture was added in a soft agar overlay to LB agar plates supplemented with 10 mM $MgSO_4$, and incubated overnight at 37 °C. The phage titer was calculated from the number of plaques obtained on the plates.

3.2.2.1 Preparation of recombinant lambda (λ RS45) lysate

λ RS45 is a lambda phage vector used for transferring fusions to single copy by homologous recombination *in vivo* (the DNA that is integrated between the recombination region of *bla'-lacZAY* of the phage vector gets transferred) into the recipient cell's chromosome at the phage λ attachment site (Simons et al., 1987). The λ RS45 that was used as the starting vector has phenotype of blue plaques in XGal IPTG (due to the *lacZYA* gene fused in to the recombinant vector). This recombinant λ RS45 was propagated in *E. coli strain* MC4100, harboring the pK8641 plasmid. This plasmid carries the P_{lac}-*nutR*-$T_{R'}$-T1T2-*lacZYA* fusion. The recombinant λ RS45 vector, harboring the P_{lac}-*nutR*-$T_{R'}$-T1T2-*lacZYA* cassette was screened by the logic that the cassette contains three strong terminators and hence the recombinant λ RS45 harboring this cassette will appear pale blue compared to the deep blue color of the parent vector, on a bacterial lawn of MC4100 in an LB Soft Agar plate supplemented with XGal IPTG. The plates were prepared by pouring a mix of the bacterial overnight (O/N) culture and the recombinant phage suspended in soft agar (the soft agar is maintained at around 37 °C) and a mix of 0.3 ml O/N culture of MC4100 and 10^7 plaque-forming units (pfu) or phage to an MOI of 0.01 of the

recombinant λ RS45 was added; mixed quickly and overlaid on to the LB agar plate supplemented with X Gal IPTG (to a concentration of 40 µg/ml for X Gal and 1 mM for IPTG), 0.4% maltose and 10 mM $MgSO_4$. The pale blue recombinant λ RS45 harboring the sequence of interest was picked up and propagated as follows.

3.2.2.2 Preparation of recombinant l lysate from single plaques

A single plaque of λ contains approximately 10^5-10^6 pfu/ml. The method of propagation of λ from a single plaque was as follows. The contents of a single isolated plaque (the pale blue plaque/s here) were drawn into a sterile 1-ml pipette tip and dispensed into 0.5 ml of LB. After addition of a drop of chloroform, the contents were vortexed for short bursts and centrifuged. The lysate was purified to homogeneity for phages expressing the pale blue phenotype by passaging it once again through fresh MC4100 cells by growing the adsorbed phage-bacterial cells as above in LB XGal plates and picking up pale blue plaques from it and repeating the lysate preparation from the picked up plaque, as above. The clear supernatant obtained was mixed with 0.3 ml of MC4100 cells grown O/N in LB supplemented with 10 mM $MgSO_4$ and incubated for 20 min at room temperature for adsorption. The infection mixture was then added into 10 ml of LB supplemented with 5 mM $MgSO_4$ and incubated at 37 °C with shaking until lysis. The lysate was mixed with chloroform and was purified by centrifugation as given above and clear lysate was stored at 4 °C.

3.2.3 Preparation of λ Lysogens harboring the P_{lac}-*nutR*-$T_{R'}$-T1T2-*lacZYA* system

A bacterial lawn of K3093 (RS20) was prepared by mixing 0.3 ml of O/N culture of K3093 (grown in LB containing 10 mM $MgSO_4$) with 3 ml soft agar (kept at a temperature of around 37 °C- 40 °C) and overlaying it by pouring it into an LB agar plate supplemented with 10 mM $MgSO_4$ and allowing it to solidify for 10 minutes, inside the laminar air flow chamber. Dilutions of (10^{-5} and 10^{-6} in LB) of the recombinant λ RS45 lysate (purified to homogeneity for λ RS45 harboring the P_{lac}-*nutR*-$T_{R'}$-T1T2-*lacZYA* cassette and having the pale blue phenotype on X Gal IPTG plates) was spotted onto this lawn and allowed to adsorb on to the lawn for 30 minutes. The plate was then incubated for 18-20 hours in a 37 °C incubator. Lysogens (RS20 harboring the λ prophage) which were visible on the clearings in the bacterial lawn brought about by phage mediated lysis were picked up using a toothpick and streaked onto a LB agar plate onto and incubated O/N at 37 °C. The lysogens were re-streaked into fresh plates and a colony PCR using the primers specific for the upstream and down stream regions of the P_{lac}-*nutR*-$T_{R'}$-T1T2-*lacZYA* was carried out, to confirm the presence of the cassette in the chromosome of RS20. The lysogens

obtained after transduction were checked to confirm that they were single copy insertions (prophage copy number) at the ? phage insertion site in the bacterial chromosome by PCR screening using *E.coli att*B and ? *att*P specific primers (SBM 15,16,17; also refer Appendix II) by following the methodology described by Powell et al., 1994 (Figure 3.3).

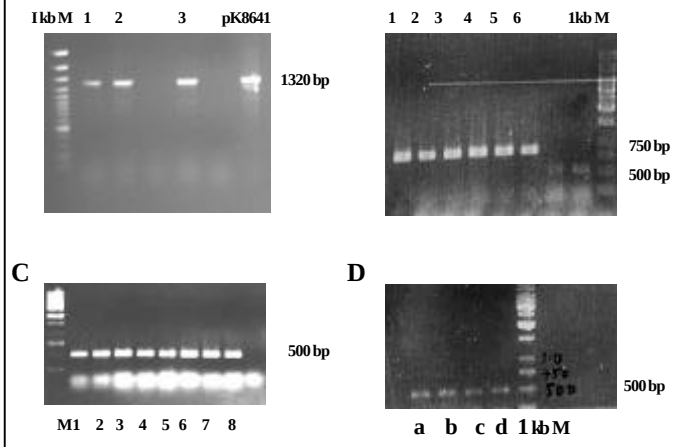


Figure 3.2: Colony PCR for screening the 3T and ? T lysogens
A. **PCR screen for 3T lysogens.** Lanes 1, 2,3. Positive lysogens amplified with RS58/RK1 primers; pK8601 harboring the T7-*nutR*-3T cassette on PCR amplification gives the DNA product as expected with lysogens and hence served as the positive control.1 kb DNA ladder, MBI is shown. **B**. **PCR to screen mono lysogens.** Lanes 1-8.Monolysogen recombinants amplified to the expected size of 500 bp using primers SBM 15, 16, and 17. **C**. **PCR screen for ? T lysogens.** Lanes 1-6.Positive lysogens amplified with RS58/RK1 primers. **D**. **PCR to screen ? T mono lysogens**. Lanes a-d. Monolysogen recombinants amplified to the expected size of 500 bp using primers SBM 15, 16, and 17.

3.2.4 Phage P1 lysate preparation

A single colony of the donor strain (DJ354 for *rpoC ts* and DJ3401 for *rpoB ts*) was inoculated into 3 ml LB supplemented with 1.0 mM $CaCl_2$ and Tetracycline and was grown overnight at 30 °C. 0.3 ml of this culture was taken into a sterile eppendorf tube and was mixed with 10^7 plaque-forming units (pfu) of a stock P1 lysate prepared on MG1655. Adsorption was allowed to occur at 30 °C for 30 minutes.10 ml LB medium supplemented with 5 mM $CaCl_2$ and Tetracycline was added to this and kept on a shaker incubator set at 30 °C for 3-4 hours and monitored for cell lysis. Once visible cell lysis occurs, this lysate was treated with chloroform (0.3 ml/10 ml), shaken well, transferred into a sterile glass test tube, centrifuged at 4000 rpm. After centrifugation, the clear supernatant was removed into a sterile test tube, twice treated with chloroform as earlier and the clear lysate was stored at 4 °C. To quantitate the P1 phage lysate preparations, titration was done using a P1-sensitive indicator strain such as MG1655. 100 µl each of serial dilutions of the recombinant P1 phage made (typically 10^{-5}, 10^{-6}) were mixed with 0.1 ml of the fresh culture. After 15 min adsorption at 37 °C without shaking, each mixture was

added in a soft agar overlay to LB agar plates, and incubated overnight at 37 °C. The phage titer was calculated from the number of plaques obtained on the plates.

3.2.5 P1 transduction

2 ml of overnight culture of the recipient strain (RS31), (grown in LB supplemented with 5mM $CaCl_2$) was spun at 4000 rpm, the supernatant was removed and was re-suspended in 0.5 ml fresh LB supplemented with 2.5 mM $CaCl_2$. This was mixed with 10^7 plaque-forming units (pfu) of a recombinant P1 lysate prepared on strains DJ354 for *rpoC ts* and DJ3401 for *rpoB ts*, as required.

Adsorption was allowed to occur at 30 °C for 30 minutes. This was then spun at 4000 rpm to remove the un-adsorbed P1 phage. The bacterial pellet was resuspended in 4 ml of fresh LB supplemented with 150 µl of 1M Sodium Citrate (to prevent further adsorption P1 phage). This was incubated for 30 min at 30 °C with slow shaking to allow for phenotypic expression of the antibiotic resistance gene. This was then spun at 4000 rpm, the supernatant was taken out and the bacterial pellet was resuspended in 0.3 ml of Citrate buffer. The mixture was then centrifuged at 4000 rpm, and the pellet was resuspended in 0.3 ml of citrate buffer. 100 µl aliquots were plated on appropriate antibiotic containing plates supplemented with 2.5 mM sodium citrate (LB Tet plates here). A control tube without the addition of the P1 lysate was processed in the similar way as described above. The plates were incubated O/N at 30 °C, the transductants that came up in the plate were streaked onto fresh plates with the same medium and grown O/N at 30 °C. The transductants were checked for temperature sensitive phenotype by patching the same colony onto two plates with the same medium, the plates were then incubated separately for O/N at 30 °C and 42 °C; and those clones were selected, which failed to grow at 42 °C but grew well at 30 °C.

3.2.6 Random mutagenesis of RNAP by chemical mutagenesis with MNNG

Mutant plasmid libraries of both *rpoB* and *rpoC* to be used for screening N specific RNAP mutants in the background strains were constructed by random mutagenesis using a chemical mutagen, MNNG (*N*-methyl-*N'*-nitro-*N*-nitrosoguanidine). MNNG methylates numerous positions on bases in DNA. Methylation in the O^6 position of Guanine results in the G: C to A: T transition, which is the principal mutation induced by MNNG in bacteria (Miller, 1992). The *rpoC* and *rpoB* encoding plasmid were transformed into *E.coli* MG1655, separately. Random mutagenesis of the resultant strains (RS44 and RS90 respectively) was carried out by MNNG

mutagenesis as per the protocol of Miller (Miller, J.H 1992) and the efficiency of mutagenesis was assayed by calculating the survival rates of the MNNG treated cells after mutagenesis.

An overnight culture of a MG1655 strain, harboring the plasmid to be mutagenized was made by inoculating a single colony of the strain in 3 ml LB with the appropriate antibiotic. This was then sub cultured to a 1:50 dilution in LB with the appropriate antibiotic and incubated until an A_{600} of 0.6-0.8 was reached. 5 ml of cells were centrifuged, washed twice with equal volumes of 0.1 M citrate buffer (pH 5.5), and resuspended in the same volume of the buffer. MNNG was prepared fresh as a 1 mg/ml stock solution in citrate buffer and added to cells at a final concentration of 50 μg/ml. Incubation was continued for 45 min at 37 °C, followed by washing twice (centrifugation of the cells at 4000 rpm, discarding the supernatant and resuspending in 5 ml 0.1 M Phosphate buffer) with 0.1 M phosphate buffer (pH 7.4). The cells were finally resuspended in 2 ml of LB and grown overnight at 30 °C in 20 ml of LB broth with the appropriate antibiotics. From the overnight grown cells, plasmid was isolated by midi preparation using Qiagen midi prep kit to obtain the mutagenized plasmid libraries. The efficiency of mutagenesis was calculated by plating dilutions of the surviving fraction of the cells after MNNG treatment and the untreated cells (control) in LB agar plates with appropriate antibiotics and growing overnight. The titer of the cells of the MNNG treated and untreated control was taken, and the efficiency of mutagenesis was measured by comparing the values. About 0.45% and 0.35% survival rates were observed for the *rpoC* and *rpoB* encoding MG1655 strains respectively. The glassware and materials used for the treatment of cells with MNNG were deactivated by treatment with 1 N NaOH as described by Miller (Miller, J.H. 1992). The H-19B N gene encoding pK8601 was also mutagenized in a similar fashion and a mutant plasmid library of N gene was obtained. To obtain an enhanced N specific mutant library, the MNNG treated N plasmid was electroporated into the corresponding ? N wild type *rpoC/B* strain, the white transformants were pooled, and the total plasmid was isolated by midi prep.

3.2.7 Screening for N specific RNAP mutants

The mutagenized plasmids were electroporated (see Appendix III for Electroporation protocol) to the corresponding background strain (RS38 for *rpoC* and RS82 for *rpoB*) and the transformants were plated on LB agar supplemented with Ampicillin, Spectinomycin, Tetracycline, XGal and IPTG and grown at 42 °C. Approximately 1.2×10^6 colonies for *rpoB* and 0.9×10^5 colonies for *rpoC* were screened. White mutants were picked, purified thrice by streaking in LB agar plates with the selection medium supplemented with XGal IPTG and the putative mutant *rpoB/C* plasmid was isolated using miniprep kits from Qiagen. These plasmids were then retransformed

into their respective background strain and checked to see if they retain the mutant phenotype or not.

3.2.7.1 Site directed mutagenesis for obtaining single substitutions from double mutants

Site Directed Mutagenesis was carried out to make the P251S *rpoC* mutant from the P251S P254L double mutant. The plasmid with *rpoB/C* expressing the mutant phenotype were sequenced and characterized by *in vivo* assays. Site directed mutagenesis was carried out essentially using the methodology described in Quick Change Site Directed Mutagenesis Kit from Stratagene, with slight modifications. Deep Vent DNA Polymerase enzyme was used in the thermal cycling reactions for the incorporation and extension of the mutagenic oligonucleotide primers (oligos encoding the necessary a. a substitutions, refer Appendix II for oligos) with wild type *rpoC* plasmid as the DNA template, followed by phenol extraction and DpnI digestion. The digested mix was passed through a G25 spin column (Amersham) and was electroporated to DH5a cells; the plasmids obtained from the transformants were screened in the corresponding background strain (RS38) for mutant phenotype at 42 °C and were sequenced to confirm the mutation. Site directed mutagenesis was also carried out to make the single R903H *rpoB* mutant from the R903H G1045D double mutant (refer Results section). The resultant plasmid was sequenced to confirm the unique substitution and was transformed into the *rpoB* background strain and *in vivo* assays showed that the single substitution does not retain the N mediated defective phenotype, as it expressed *in vivo* antitermination rates of around 35%, almost equivalent to the corresponding wild type values.

3.2.8 *In vivo* characterization of RNAP mutants

3.2.8.1 Beta-Galactosidase assay

Beta Galactosidase assay (Beta Gal assay) is commonly used to determine promoter strength, transcriptional activity, termination etc in a construct that have area of interest fused upstream to the Galactosidase gene. Assays for determination of Beta-Galactosidase enzyme activity in cultures were performed as described by Miller (Miller, J.H 1992) after permeabilizing the cells with SDS/chloroform, and the activity values were calculated in Miller units, as defined therein. Beta-Galactosidase enzyme cleaves the Lactose analog ONPG (O-Nitro phenyl-ß-D-Galactosidase), a colorless compound into yellow colored ONP (O- Nitro phenol) and Galactose. After disruption of cells with SDS and Chloroform, ONPG is added and the reaction is allowed to proceed until visible yellow color is developed. Reactions were stopped by adding 0.5 M Na_2CO_3. The Beta gal values are obtained by quantification spectrophotometrically at 420 nm. The white mutants were subjected to Beta Gal assays (Miller 1992) in triplicates over different days for

checking the *in vivo* antitermination rates. The mutant plasmids were transformed into RS69 and RS83 respectively for the ΔTerminator *rpoC* and *rpoB* strain background. The mutant plasmids in the 3T and ΔT strain background were grown in LB media with all the antibiotics and IPTG at 42 °C. % Read through (representing *in vivo* antitermination rates) was calculated as the ratio between the Beta Gal values obtained for the mutant plasmids in the 3T and ΔT strain background. Wild type *rpoB/C* plasmid transformed into the respective background strains served as the controls. A Del N strain (the pK8601 plasmid conferring N protein in *trans* is absent in this strain) for both 3T and ΔT cassettes were constructed by transforming the mutant plasmids into RS36/RS68 and screening for Spectinomycin sensitive colonies, which does not harbor pK8601 plasmid. This served as a control to check for N independent antitermination and the specificity of the mutant phenotype *in vivo* at 42 °C. O/N cultures of the required strains were grown in appropriate media and this was then sub cultured to fresh medium with IPTG; and was grown in a shaker with aeration till O.D. reaches 0.3 to 0.4.The cultures were chilled on ice. The O.D. of the cultures was measured at 600nm. 100 µl of culture was added to 900µl of Z-buffer. Two drops of chloroform (~40 µl) and one drop of 0.1% SDS solution (~20 µl) was added to each assay mixture and the tubes were vortexed for 30 seconds each. The tubes were placed in a water bath set at 28 °C for five minutes. Reactions were started by adding 0.2 ml of ONPG (4mg/ml). After a detectable level of yellow color is developed, the reaction was stopped by adding 500 µl of 1M Na_2CO_3 solution. Each tube was centrifuged at 12000 rpm and the O.D was recorded at 420 nm in a spectrophotometer.ß-Gal activity was calculated by using the following formula.

$$\text{Units of ß–Gal} = \frac{1000 \times O.D_{420}}{t \times V \times O.D_{600}}$$

Where T = time of reaction in minutes and V = volume in ml.

3.2.8.2 Phage spotting as an assay for N mediated antitermination *in vivo*

Bacteriophages can grow on a bacterial lawn and can form plaques, which are formed by the lysis of the host bacterium by a phage in the lytic mode. Plaques can be formed only if the phages can antiterminate and synthesize the late gene products, required in lysis. Here, mutations of the RNA polymerase impair the binding of phage N protein and the formation of a processive elongation complex, which can antiterminate efficiently. Hence this system can be used to study the *in vivo* antitermination rates at 42 °C, in a suitable *rpoB/C ts* strain background, where the entire RNA polymerase ß /ß' subunit function comes from the mutant plasmid. The mutant *rpoB/C* plasmids were transformed into RS158 and RS215 strains (K3093 to which the *rpoB and C ts* mutation respectively was introduced by P1 transduction). These strains were then grown at 42 °C with

appropriate antibiotics. The cells grown O/N were used to make a bacterial lawn on LB Agar with the antibiotics. Dilutions of H-19B Phage and λC1B2 and λC1857 was spotted on this lawn and grown at 42 °C and plaques were counted. Assays with each of the mutant were done four times each, on separate days with different O/N cultures. Strains were grown overnight in LB containing 0.4% maltose and 10 mM $MgSO_4$, sub cultured and grown to early stationary phase in the same medium. 100 µl of the culture (grown at 42 °C) was mixed with 2.5 ml of soft agar and overlaid on LB agar plates supplemented with 10 mM $MgSO_4$. Serial dilutions of Wt H-19 B (obtained from K9774-K37 strain with H19BStx::Cam lysogen, gift from Dr.David Friedman) H-19B *Nin* (this study, see ahead), λ*C*1B2 (from J Gowrishankar) and λY1090 (λ *Nin5*, a gift from Dr. Robert Weisberg) lysates were prepared (in LB) and 10 µl were spotted from each dilution on the soft agar lawn and allowed to dry. Wt *rpoB/C* plasmid transformed into their respective background strain served as controls. The plates were incubated at the 42 °C overnight, and the plating efficiency determined by counting the number of plaques formed. The *in vivo* antitermination rates were obtained by calculating the Efficiency of Plating (EOP) of phage growth between the mutants and the corresponding wild type allele for *rpoB/C*. Efficiency of plaque formation was measured relative to WT, where efficiency for WT was fixed at 1. H-19B *Nin* phage was made by screening for and purifying plaques that grew in the G1045D mutant background and was checked for N independence by its ability to grow and form plaques in a NusA1 strain, where Wt H-19B failed to form plaques. We also confirmed the deletion in the *Nin* region by PCR amplification using *Nin* specific primers (RS172, RS173, RS175, RS200; refer Appendix II). N independent anti termination defect was assessed by spotting the H-19B *Nin* and the λ *Nin*5 phages along with the respective Wt phages as above.

3.2.9 Identification of the substitutions in *rpoB/C*

The *rpoC/B* mutants with significantly reduced antitermination activities with respect to the Wt were sequenced to identify the mutations, with a series of *rpoC/B* gene specific internal primers (see Appendix II for oligo list). The entire *rpoC/B* genes were sequenced to rule out the chance of secondary mutation being present (if any) and the point mutations were identified by using pair wise BLAST (BLAST2SEQ, NCBI) against the *rpoC/B* wild type sequence. The substitutions were confirmed by re-sequencing the region of interest.

3.2.9.1 Localization of the mutants on TEC

The localization of the mutants in a model of Ternary elongation complex (TEC) (provided by Dr.Seth Darst) was carried out by aligning the corresponding *E. coli* residues to the *T. acquaticus*

and the TEC residues using CLUSTALX program followed by mapping the substitutions on the TEC. The alignments were manually corrected to represent the gaps present in the model structure such that the aligned positions represent the structurally equivalent positions in all the three structures. The location of the *rpoB/C* mutations in the TEC model was pinpointed this way. The positions of the mutations, in the three dimensional structure of *E.coli* RNA polymerase were found with the help of the model of the TEC (provided by Dr. Seth Darst) of *Thermus aquaticus* in complex with DNA-RNA hybrid between *E.coli* RNA polymerase and *T.acquaticus* RNAP subunit sequences were obtained from the HOMSTRAD Database. The structure based sequence alignments database (http://wwwcryst.bioc.cam.ac.uk/~homstrad/). To these profiles, the corresponding subunit sequence of the EC model was aligned using CLUSTALX (Thompson et al., 1997). The alignments were manually corrected to properly represent the gaps occurring in the model structure. *E. coli* a. a substitutions was then mapped onto the structurally equivalent positions in the EC model. The distances of these mutation sites to the structures associated with the active site region was obtained by computing the distance (using a shell script) between the mutation sites and each of the residues in structures comprising the active site region. For this the distances between the C-a atom at the mutation site and the phosphorous atoms in each of the nucleotides of the DNA/RNA were considered.

3.2.10 *In vitro* characterization of RNAP mutants

3.2.10.1 Establishment of a functional antitermination system

We established an *in vitro* transcription assay to check the efficiency and specificity of H-19B N mediated antitermination at the triple terminator ($T_{R'}$-T1-T2) cassette, the similar one that has been used in the *in vivo* experiments. We used a linear DNA template, bearing a *nutR* site followed by the triple terminator cassette. Transcription was initiated from the strong T7A1 promoter. Here, as RNA is being synthesized *in vitro*, H-19B N protein can bind to the *nut* site (coded by the *nutR* in the template DNA) and along with the host elongation factors supplied can interact with the RNA polymerase, forming a functional antitermination complex, *in vitro*. In the absence of N, the transcription terminates at one of the three terminators present in the template. But, when N is supplied and when it is able to form a functional antitermination system, RNAP can read through the terminator and 'Run Off' RNA (full length RNA) is made. A Δ *nutR* variant of the template was also made, which was used as a control for checking the N mediated specificity of the antitermination system (Figure 3.2). The RNA products can be analyzed on a gel, if the RNA synthesized is radiolabelled. The antitermination rates can be calculated by

comparing the run off RNA band intensity between the –N and +N RNA transcripts. We cloned the required DNA templates; and also cloned, expressed and purified H-19B N and the other host *nus* factors required, for setting up such an *in vitro* antitermination system.

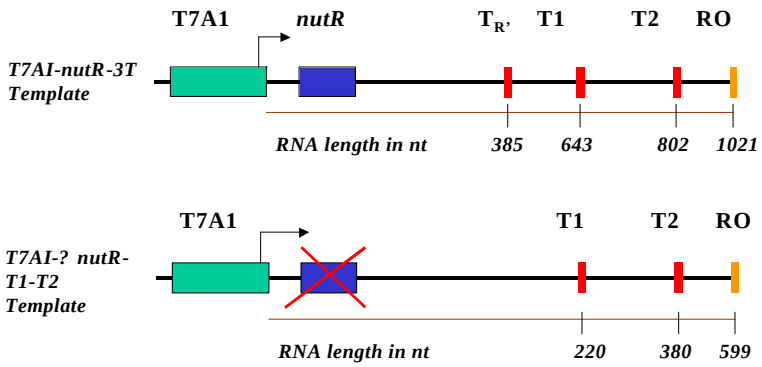


Figure 3.3: Cartoon representing the two different templates used for *in vitro* antitermination assays.

The T7A1-*nutR*-T_R'T1-T2 and the T7AI-? *nutR*-T1-T2 templates are shown.

3.2.10.2 Construction of templates for *in vitro* transcription

The plasmid pK8641 (Neely and Friedman, 1998b) having the sequencP_{lac}-H19B- *nutR* –3T-*lacZ* was used for making the DNA templates for the *in vitro* transcription system. As the transcription signal from Lac promoter was very weak under single round *in vitro* transcription conditions, we replaced the lac promoter with T7A1 promoter, by cloning a PCR amplified (primers RS67 and RS68) T7A1 region into EcoRI – HindIII site of pK8641 (pRS18); the resultant plasmid having a DNA cassette of T7A1-*cro*-Δ*nut* – *lacZ*. H19B *nutR*-T_R' (PCR amplified from pK8641 using primers RS58/RK1) was introduced to the HindIII site of pRS18 forming pRS22, having the pT7A1-*nutR*-3T-*lacZ* cassette. A T7A1-*nutR* (? *cII*)-3T-*lacZ* (this was designed so that the total transcript length will be lesser, and for this the *cII* region in the earlier cassette was deleted) was amplified from pRS22 using RS58/RS76 primers and cloned at its EcoRI/SmaI site (pRS25).

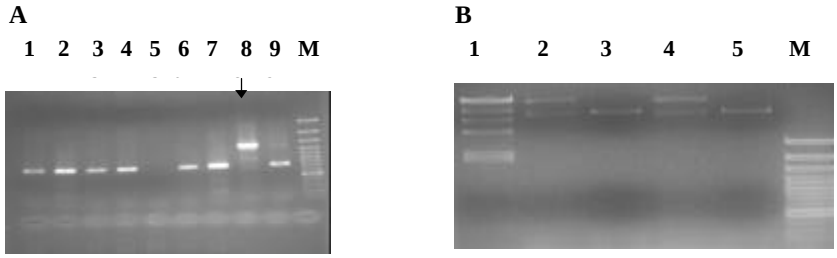


Figure 3.4: Cloning of pRS22 (pT7A1-*nutR*-3T-*lacZ*)
A. Colony PCR with RS58/RK1 primers to screen for *nutR*-3T insertion. Lanes 1-7 and 9 are PCR products, which didn't give the expected size for the 3T insertion. Lane 8 (marked with and arrow) is the PCR product giving the correct size, amplified from clone #39 colony; M denotes DNA marker (100 bp DNA ladder; NEB) **B.** Restriction analyses (double digestion with EcoRI/HindIII) of the positive #39 plasmid along with control pK8641 plasmid.

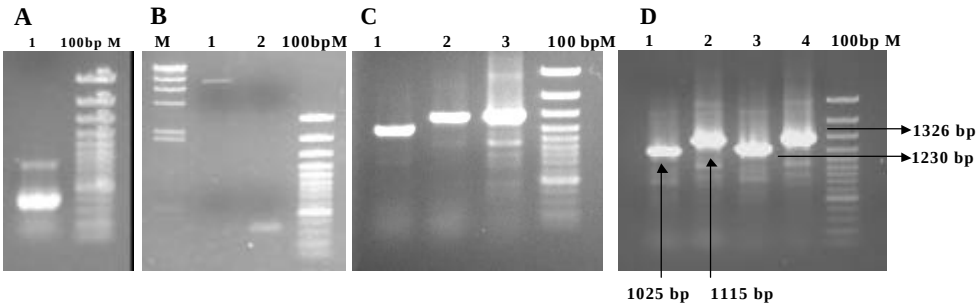


Figure 3.5: Cloning of pRS25 (pT7A1-*DcII nutR*-3T-*lacZ*)

A. PCR amplified T7A1-*nutR* DNA using RS58/RS76 primers used for cloning. Lanes: 1. PCR amplified DNA; 2. 100 bp DNA Marker (NEB). **B**. **EcoRI/SmaI double digested pK8641 plasmid (vector) and T7A1-*nutR* DNA (insert) used for cloning**. **C**.**PCR screen of colonies with RS58/RK1 primers Lanes**: 1.Positive clone showing the correct insert size (1.15 kb); 2. Clone with a product higher in size (~ 1.32 kb); 3.Control PCR product with pK8641 DNA as template (1.32 kb).**D**. **PCR confirmation of the +ve clone (pRS25).** Lanes: 1. Positive clone DNA amplified with RS58/RK1 primers; 2.pK8641 DNA amplified with RS58/RK1 primers; 3. Positive clone DNA amplified with RS67/RK1 primers; lane 4- pK8641 DNA amplified with RS67/RK1 primers.

This plasmid was used as the template DNA for PCR amplification of the shortened pT7A1-*nutR*-3T DNA template used in the *in vitro* antitermination transcription assays.

During the course of design of various *in vitro* experiments with this system, different template DNA was amplified using appropriate primer, but essentially using the plasmid pRS25 as the template DNA in the PCR reactions. The construction of such transcription templates will be discussed with their relevant experiments.

3.2.10.3 Cloning, expression and purification of proteins

3.2.10.3.1 Cloning of Wt and C terminal His tagged H-19B N

The Wt H-19B N gene was PCR amplified using gene specific primers (RS65; RS66a) from plasmid pK8601 (Neely and Friedman 2000) and was cloned into XhoI-NdeI sites of the pET vector pIA200 (pET21b), transformed into ultracompetant *E coli* DH5a cells. The recombinants having the insert were screened using colony PCR, confirmed by double digestion with XhoI-NdeI enzymes for the insert release, sequenced and the plasmid was named as pRS8.

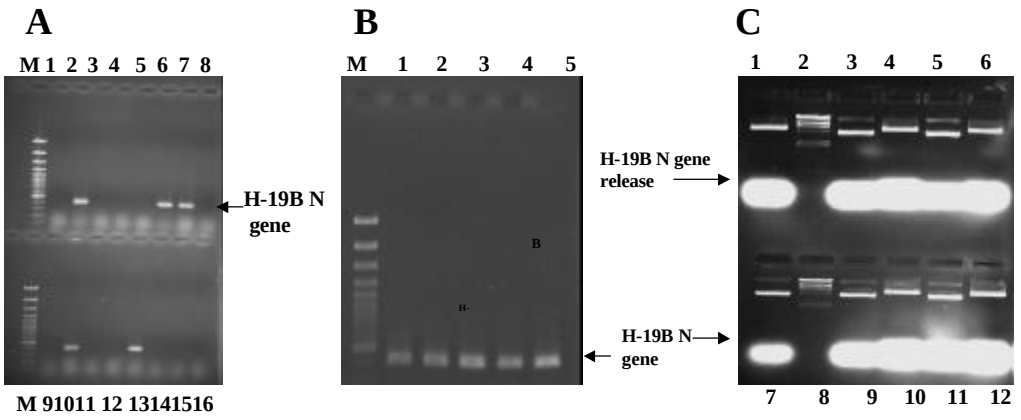


Figure 3.6: Cloning of Wt H-19B N gene into pET28B

A. **Colony PCR for screening H-19B N gene**. Lanes: 1,3,4,5,8,9,11,12,14,15,16 show clones where there were no proper integration of the N gene; lanes 2,6,7,10,13 show positive clones.

B. **NdeI/XhoI digestion of recombinant pET21b with H-19B N cloned in**. Lanes: 1,3,5,7,9 and 11 are uncut plasmid DNA samples; lanes 2,4,6,10,12 are NdeI-XhoI double digested plasmids. The released N gene is shown with arrows down, just above the RNA.

C.**PCR screen to confirm H-19B N gene amplified from the plasmid using RS66a and RS65 primers.** Lanes 1-5 shows the N gene amplified with the gene specific primers (RS66a-RS65) from the recombinant plasmids. M stands for 100 bp DNA Marker (NEB).

We also made a His-tagged (C terminal 6-Histag) of Wt H-19B N, following the same procedure as above, with the exception that the downstream primer RS64, did not have a stop codon and this allowed the addition of the 6-His tag from the vector to the C terminus of N transcript. The vector provided translation stop codon. The positive recombinants were sequenced and the plasmid was named as pRS35.

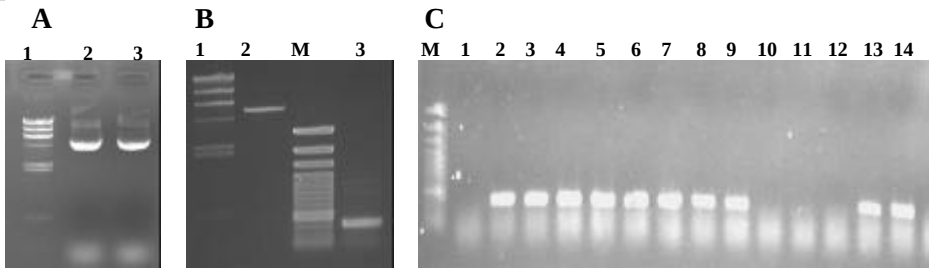


Figure 3.7: Cloning of His tagged H-19B N

A. *Lanes.1.? HindIII digest DNA marker; 2,3 pET21b plasmid purified.*

B. **Double digested vector and insert**. Lanes: 1.? HindIII digest DNA marker; 2. NdeI-XhoI double digested pET28b plasmid (vector); 3.NdeI-XhoI double-digested H-19B N DNA (amplified using RS66a-RS64 primers; insert); M-100 bp DNA ladder, NEB.

C.**PCR screen to confirm H-19B N gene amplified from the plasmid using RS66a and RS64 primers**. Lanes: 2-9 and 13-14 shows the His tagged H-19 B N gene amplified with the gene specific primers (RS66a-RS64) from the recombinant plasmids. M stands for 100 bp DNA Marker

3.2.10.3.2 Purification of Wt H-19 B N Protein

H-19B N protein was purified according to the published protocol for ? N protein (Das et al., 1996). pRS8 (refer plasmid list, Appendix I) was transformed into BL21 (DE 3) over expressing strain of *E. coli*. Over expression of Wild type N protein from the pET21b with H-19B N gene cloned under the control of phage T7 promoter regulated through a lac operator and resident *lacI* gene, was achieved in the *E.coli* over expression strain BL21 (DE3) that contains T7 RNA Polymerase gene under *lac OP* control. The N protein in the induced culture was largely in an insoluble form. It was recovered in a low speed pellet of the crude cell lysate. The N protein was then solubilized with 8 M Guanidium hydrochloride and a pure preparation of N protein was obtained after two affinity chromatographic steps. Single colonies were inoculated into LB with Ampicillin and grown O/N at 32 °C. This was then sub cultured (1000 fold dilution) into 500 ml Terrific Broth (TB) with Ampicillin and grown in a shaker incubator till an OD_{600} of 0.5-0.6 .The culture was then induced with IPTG to a final concentration of 0.5 mM. The growth was continued for another 4 hours and was stopped by chilling the culture on ice. The cells were harvested by centrifugation at 8000 rpm for 5 minutes at 4 °C. The pellet was then resuspended in 20 ml of Buffer A (20 mM Tris-HCL, pH 7.5; 1mM DTT; 0.5 mM PMSF and 10% (v/v) Glycerol) and centrifuged again as above and the pellet was chilled on ice. 20 ml of Buffer B (20 mM Tris-HCL pH 7.5; 100 mM NaCl; 10 mM EDTA; 10% (v/v) Glycerol; 10 mM 2-mercaptoethanol and 0.5 mM PMSF) was added to the pellet, resuspended well by repeated pipetting and was then sonicated to 50 cycles for about 10 minutes in a sonicator for disrupting the cells. The lysate was centrifuged at 13,000 rpm for 30 minutes at 4°C and the supernatant was removed. The pellet was resuspended in 20 ml of Buffer B and centrifuged as before and the supernatant was removed. The pellet was kept on ice and was then solubilized using 15 ml of Buffer A containing 8 M Guanidium hydrochloride and was incubated in ice for 1 hour and was then centrifuged at 13,000 rpm for 30 minutes at 4 °C. The clear supernatant was collected and was dialyzed against 750 ml of buffer C (Buffer B containing 2 M Urea and 1 mM DTT) overnight in a cold room with two changes of buffer. The supernatant containing the protein was centrifuged at 12000 rpm for 15 minutes at 4 °C the next day to remove any protein aggregates. The supernatant was transferred to fresh pre-cooled Falcon tubes and was stored at 4 °C or was loaded on to a pre-packed CM Sepharose column (Amersham). The column was earlier equilibrated with buffer B. The protein was loaded into the column manually using a syringe, was washed with buffer C and a stepwise gradient of Buffer C with NaCl in concentrations ranging from 0.15 M to 1 M (2 ml each) was applied to the column. Aliquots were collected for each fraction. The aliquots were then checked for the presence of protein by running in a 10% SDS

PAGE. The fractions with H-19BN protein in them was pooled and was dialyzed O/N against buffer C without NaCl. The protein was centrifuged at 12000 rpm for 15 minutes at 4 °C the next day to remove any protein aggregates. The supernatant was then diluted with buffer C without NaCl to double the volume and was loaded on to a Heparin column (Amersham), which was earlier equilibrated with buffer C without NaCl. N protein was eluted with a 2 ml linear gradient of NaCl (0.2-1.0 M). The pure N protein elutes out in the range of 0.4- 0.6 M NaCl. The fractions were checked for the presence of protein by running in a 12% SDS PAGE. The pure fractions were pooled and were dialyzed against buffer B and were concentrated in a 3 Kilo Dalton (KD) concentrator (Amicon) against 2X storage buffer (20 mM Tris-HCL pH 7.5; 100 mM NaCl; 10 mM EDTA; 10% (v/v) Glycerol; 10 mM DTT and 0.5 mM PMSF). Glycerol to the final concentration of 50% v/v was added to the concentrated pure protein, aliquoted into fresh sterile eppendorf tubes and stored in $^{-}$70 °C deep freezer.

3.2.10.3.3 Purification of His tagged N Protein

pRS35 (C-terminal Histag H-19B N in pET21b) was transformed into *E.coli* BL21 DE 3 cells. The N protein in the induced culture was largely in the insoluble form. It was recovered in a low speed pellet of the crude cell lysate. The N protein was then solubilized with 8M Urea and a pure preparation of N protein was obtained after two affinity chromatographic steps. Single colonies were inoculated into LB with Kanamycin and grown O/N at 32 °C. This was then subcultured (1000 fold dilution) into 500 ml TB with Kanamycin and grown in a shaker incubator till an OD_{600} of 0.5-0.6 and was induced with IPTG to a final concentration of 0.5 mM. The growth was continued for another 4 hours and was stopped by chilling the culture on ice. The cells were harvested by centrifugation at 8000 rpm for 5 minutes at 4 °C. The pellet was then resuspended well by repeated pipetting, in 25 ml of Binding buffer (50 mM $NaH_2PO_4.H_2O$, 300 mM NaCl, 15 mM Imidazole, pH adjusted to 8.0) with 8 M Urea. It was incubated for 1 hour at 4 °C, for the solubilization of the protein, with constant stirring. The lysate was centrifuged at 13,000 rpm for 30 minutes at 4 °C and the supernatant was carefully transferred to a fresh pre-cooled falcon tube and centrifuged briefly at 5000 rpm at 4 °C for 5 minutes to remove any aggregates. The lysate was loaded on to a pre-equilibrated His Trap HP column (Amersham; pre-equilibrated with Binding buffer with 8 M Urea, as per manufacturer's instructions). The flow through was collected and the column was washed with 15 ml wash buffer (50 mM $NaH_2PO_4.H_2O$, 300 mM NaCl, 20 mM Imidazole) and the wash was collected. His tagged H-19B N protein was eluted with 10 ml of Elution buffer (20 mM Phosphate, 0.5 M NaCl and 500 mM Imidazole). The load, wash and elute samples were loaded onto a 10% SDS PAGE for analyses. The pure H-19B N

fractions in the elute were then pooled, and was dialyzed O/N against buffer C with 2 M Urea (20 mM Tris-HCL, pH 7.5; 100 mM NaCl; 10 mM EDTA; 10% (v/v) Glycerol; 10 mM 2-mercaptoethanol and 0.5 mM PMSF and 2 M Urea) with two buffer changes. Next day, the protein was centrifuged at 12000 rpm for 15 minutes at 4 °C to remove any protein aggregates. The supernatant was then diluted with buffer C without NaCl to double the volume and was loaded on to a Heparin column (Amersham), which was earlier equilibrated with buffer C without NaCl. N protein was eluted with a 2 ml linear gradient of NaCl (0.2-1.0 M). The pure N protein elutes out in the range of 0.4-0.6 M NaCl. The fractions checked for the presence of protein by running in a 12% SDS PAGE. The pure fractions were pooled and were dialyzed against buffer B and were concentrated in a 3 KD concentrator (Amicon) against 2x storage buffer (20 mM Tris-HCL pH 7.5; 100 mM NaCl; 10 mM EDTA; 10 % (v/v) Glycerol; 10 mM DTT and 0.5 mM PMSF). Glycerol to the final concentration of 50% v/v was added to the concentrated pure protein, aliquoted into fresh sterile eppendorf tubes and stored in ⁻70 °C deep freezer. We found that the C terminal His tag of H-19B N made it less active, with respect to its *in vitro* antitermination efficiency- the purified protein showed reduced levels of transcription antitermination *in vitro*, compared to its wild type non-His tagged counterpart. Hence in this entire study, we have used non-His tagged Wt H-19B N protein unless otherwise mentioned.

3.2.10.4 Sub cloning of the suppressor N gene

L108F was sub cloned into the expression vector pET21b. The suppressor N gene was PCR amplified using primers RS65and RS66a from pRS243 (pGB with L108F H-19BN; double digested with XhoI/NdeI and was cloned into the XhoI/NdeI sites of pET21b). The positive transformants were confirmed by restriction digestion and sequenced. ß' P254L suppressor N was purified according to the published protocol for ? N protein (Das et al., 1996) and as given earlier for wild type H-19B N protein purification.

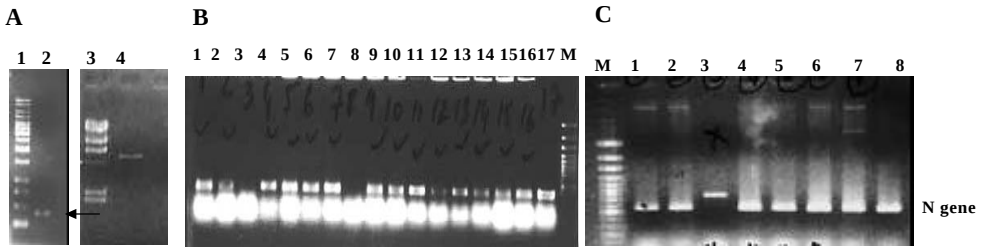


Figure 3.8: Cloning of L108F suppressor H-19B N

A. **Double digested vector and insert.** Lanes: 1.100 bp DNA marker, NEB; 2. NdeI-XhoI digested, gel purified H-19BN gene PCR amplified from pRS243 (insert); 3. ? HindIII digest DNA marker; 4.NdeI-XhoI double digested pET28b plasmid (vector). **B**. **Colony PCR to screen the recombinants with RS65/66a primers**. Lanes 1-2, 4-7, 9-17 is positive recombinants showing N gene amplification. **C. PCR to screen the recombinant plasmids with RS78/79 (T7 promoter and terminator specific) primers**. Lanes: 1-2 and 4-8 are the positive recombinants, giving the H-19B N gene amplification with T7 primers.

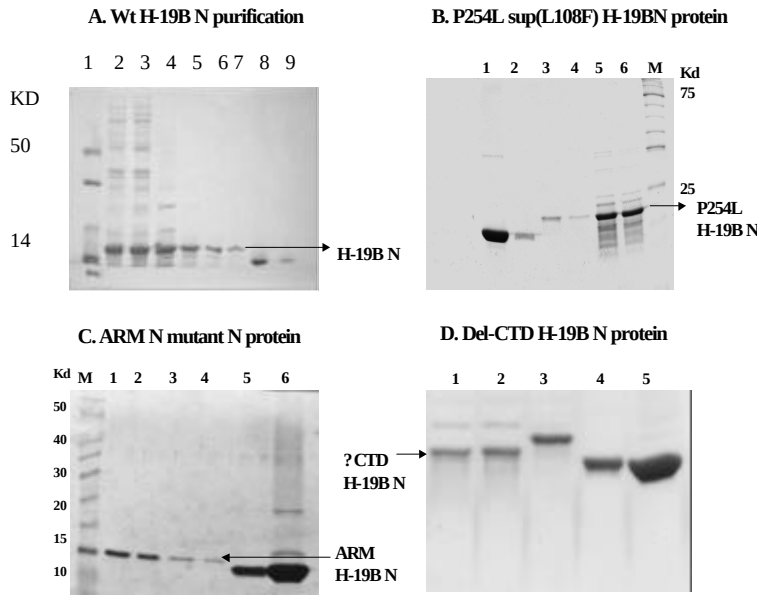


Figure 3.9: Purification of H-19B N proteins
A. **Purification of wild type H-19B N protein**. Lanes.1.MW marker (NEB); 2, 3.Cell pellet; 4. N protein after passage through MonoQ column; 6,7. Purified protein after passage through Heparin column.8, 9.Lysozyme. **B**. **Purified L108F H-19B N protein.** Lanes. 1,2. Lysozyme.3,4.Wt H-19B N 5,6.Purified L108F N protein. M- Protein molecular weight marker (MBI). **C. Purified ARM H-19B N protein.** Lanes.. Lanes 1, 2. Purified ARM mutant H-19B N. 3, 4. Wt H-19B N protein. 5, 6. Lysozyme. **D**. **Purified CTD (?100-127 a. a) H-19B N.**

3.2.10.5 Cloning, Expression and purification of *E.Coli* NusA and NusG proteins

NusA cloned in the PET (pET28b) vector (a gift from Dr. Irina Artsimovitch), was transformed into *E.coli* BL21 DE 3 strain. The majority of NusA remains in the soluble form and since here, NusA has a 6X Histag at the N terminus, it was purified to homogeneity using Ni-NTA column, followed by Heparin column. *NusG* gene was amplified from *E.coli* genomic DNA and the PCR product was cloned in the NdeI/XhoI site of pET21a to introduce a His-tag at C-terminal. The resultant plasmid was called pRS119; the sequenced plasmid was transformed into BL21 cells for over expression and NusG protein was purified using Ni-NTA columns (Amersham).

3.2.10.5.1 Protocol for purification of NusA

Single colonies were inoculated into LB with Kanamycin and grown O/N at 32 °C. This was then sub cultured (1000 fold dilution) into 500 ml TB with Kanamycin and grown in a shaker incubator until an OD_{600} of 0.4-0.45 and was induced with IPTG to a final concentration of 1.0 mM. The growth was continued for another 3 hours and was stopped by chilling the culture on ice. The cells were harvested by centrifugation at 8000 rpm for 5 minutes at 4 °C. The pellet was then resuspended in 10 ml of Lysis Buffer (50 mM $NaH_2PO_4.H_2O$, 300 mM NaCl, 20 mM Imidazole, pH adjusted to 8.0) and 1mg/ml Lysozyme, kept for 30 minutes on ice, resuspended

well by repeated pipetting and was then sonicated to 20 cycles for about 5 minutes in a sonicator for disrupting the cells. The lysate was centrifuged at 13,000 rpm for 30 minutes at 4 °C and the supernatant was carefully transferred to a fresh pre-cooled falcon tube and centrifuged briefly at 5000 rpm at 4 °C for 5 minutes to remove any aggregates. The sample was then loaded onto a pre-equilibrated His Trap HP column (Amersham, the column was pre-equilibrated as per manufacturer's protocol with 10 ml of binding buffer (20 mM Phosphate, 0.5 M NaCl and 20 mM Imidazole, pH adjusted to 7.4). The flow through was collected. The column was washed with 15 ml binding buffer and the wash was collected. NusA protein was eluted with 10 ml of Elution buffer (20 mM Phosphate, 0.5 M NaCl and 500 mM Imidazole). The load, wash and elute samples were loaded onto a 10% SDS PAGE for analyses. The pure NusA fractions in the elute were then pooled, equilibrated against buffer B (20 mM Tris-HCL pH 7.5; 100 mM NaCl; 10 mM EDTA; 10% (v/v) Glycerol; 10 mM 2-mercaptoethanol and 0.5 mM PMSF) and was then concentrated in a concentrator (Amicon-10 KD cut off) and concentrated against 2X Storage buffer (20 mM Tris-HCL pH 7.5; 100 mM NaCl; 10 mM EDTA; 10% (v/v) Glycerol; 0.2 mM DTT). Glycerol to the final concentration of 50% v/v was added to the concentrated pure protein, aliquoted into fresh sterile eppendorf tubes and stored in the $^{-}$70 °C deep freezer.

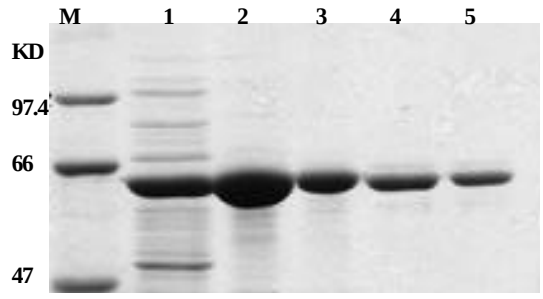


Figure 3.10: Purification of NusA protein

Different stages of purification of NusA protein Lanes: 1-2: Purification through Ni-NTA column: 1.Flow through; 2,3 show the Ni NTA fractions; lanes 4,5 show the pure NusA after passage through Heparin column; M stands for protein molecular weight marker, NEB.

3.2.10.6 Purification of mutant RNA Polymerases

Purification of mutant RNA Polymerases of either *rpoC* (for *rpoC* mutants) or *rpoB* (for *rpoB* mutant) with His tag at C-terminal were purified from the background strains carrying a plasmid containing the mutant *rpoB* and *rpoC* genes and a chromosomal copy of temperature sensitive *rpoB* or *rpoC* genes. The cells were grown at 42 ^{o}C to get majority of the holoenzyme assembled with mutant *rpoB* or *rpoC* subunits and also the presence of his-tag helped us to purify only mutant holoenzyme from the mixed pool, if any, in the cytoplasm. In all the cases RNAPs were purified from 20 to 30gms of cells essentially following the method of Kashlev (Kashlev et al., 1996) and using Ni–NTA agarose and Heparin-agarose affinity columns (as follows). As the

RNAP preparations did not have stoichiometric amount of σ^{70}, all *in vitro* transcription reactions were supplemented with purified σ^{70}. All the plasmids carrying *rpoC* mutants have His-tag at C-terminal. Plasmid pBAD18M (Sen et al., 2002) encoding the *rpoC* mutants in the *rpoC* background strain (refer Table 1) were used. For the purification of the G1045D *rpoB* mutant, a C terminal Histag was introduced to the pHYD534 by PCR amplification. For the purification of the G1045D *rpoB* mutant a C terminal Histag was introduced. A DNA fragment of the *rpoB* consisting of the G1045D mutation was PCR amplified using oligos SeqB4, an *rpoB* specific upstream primer and RS129, specific to *rpoB* C-terminal and having a 6 Histidine-tag with a BamH1 site. The amplified *rpoB* gene fragment encompassing the G1045D was then double digested with HpaI (site upstream to the mutation) and BamHI. pHYD534 was double digested with HpaI/BamHI and was ligated with the double digested C terminal His tagged G1045D *rpoB* gene. The ligation mix was electroporated into DH5α cells. The positive transformants were screened; the plasmid was sequenced and both the presence of G1045D mutation as well as C terminal Histag was confirmed. It then was transformed into RS82 (*rpoB* background strain) to check for complementation and for antitermination rates. The *in vivo* antitermination rate was found to be same as G1045D mutant and subsequently, *rpoB* G1045D RNAP was purified from it.

3.2.10.6.1 Protocol for RNAP purification

RNAPs were purified following the method of Kashlev (Kashlev et al., 1996) using Ni–NTA agarose and Heparin-agarose affinity columns. The cells were grown at 42 °C to get majority of the holoenzyme assembled with mutant *rpoB* or *rpoC* subunits and also the presence of His-tag helped us to purify only mutant holoenzyme from the mixed pool, if any, in the cytoplasm. Single colonies of the *rpoB/C* background strains with the required *rpoB/C* mutation harboring *rpoB/C* plasmid was inoculated in LB supplemented with Ampicillin, Tetracycline and Spectinomycin and grown O/N at 42 °C. Next day, it was subcultured to 1:1000 dilutions in 5 L of the same media, grown till $O.D_{600}$ reached around 1.5. The cells were harvested by centrifugation at 12000 rpm and the cell pellet was frozen. In all the cases RNAPs were purified from 20 to 30 gm of cells, grown at 42 °C. To the pellet, 50 ml of buffer A (50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 5% (v/v) Glycerol, 1 mM DTT, 300 mM NaCl). It was mixed well by stirring, on ice. 0.3 mg/ml Lysozyme was added to this, and was mixed well by pipetting and by using a blender. This was then incubated on ice for 10 min. 50 ml of Buffer A and Lysozyme (0.3 mg/ml) was added, mixed well as above and incubated for 10 min, on ice. Sodium deoxycholate (4%) was added drop by drop with stirring, to a final concentration of 0.2% and was incubated at 4 °C in a cold

room, with stirring. The cells were aliquoted to 25 ml each in fresh pre-cooled Falcon tubes and sonicated for 10 minutes for complete lysis. The lysate was then centrifuged at 12000 rpm for 45 min at 4 °C. The supernatant was transferred to a clean beaker (pre-cleaned with DEPC water and chilled) and 10% Polyethylamine (Polymin P, pH 7.9) was slowly added with constant stirring, to a final concentration of 0.8%. The stirring was continued for another 20 min (the solution turns turbid) at 4 °C. This was then centrifuged at 12000 rpm for 45 min at 4 °C. The pellet obtained was thoroughly resuspended in 100 ml TGED (10 mM Tris-HCl, pH 8.0, 0.5 mM EDTA, 5 % (v/v) Glycerol, 0.1 mM DTT) plus 0.5 M NaCl by mixing thoroughly in a blender in the cold room. The suspension is centrifuged at 12000 rpm for 15 min at 4 °C, the supernatant is stored for analysis, and the pellet is resuspended completely in TGED plus 0.5 M NaCl as before and the washing is repeated twice more. To elute RNAP from Polymin P, the pellet was resuspended in 100 ml of TGED plus 1M NaCl. The mixture was centrifuged at 12000 rpm for 30 min at 4 °C. The clear supernatant was measured, transferred into a clean and pre-cooled glass beaker and finely ground Ammonium sulphate was added slowly to the supernatant with constant stirring, to an amount of 35 g to 100 ml solution. pH was adjusted to the range of 7-7.5 with 2 N NaOH. This was left with stirring for 30 minutes. The suspension was then centrifuged at 12000 rpm for 45 min at 4 °C. The pellet obtained was resuspended in 10 ml of buffer B (20 mM Tris-HCl, pH 7.9; 10 mM $MgCl_2$; 300 mM KCl; 1 mM 2-mercaptoethanol). The protein was then dialyzed against 1 L of Buffer B O/N with two buffer changes. Next day, the protein was centrifuged at 12000 rpm for 15 minutes at 4 °C to remove any protein aggregates. The protein was then mixed with Ni-NTA beads (washed with DEPC water and then equilibrated with Buffer B) and left for binding in a rotating orbiter in the cold room for 15 minutes. It was then transferred into a Ni-NTA super flow column, by slowly pouring the slurry into the empty column, to bed volume of 5-10 ml and the resin was allowed to settle. The flow was collected and was then washed with 20 ml of Wash buffer (50 mM Phosphate, 0.3 M NaCl and 20 mM Imidazole, pH adjusted to 8.0) and the wash was collected. RNAP was then eluted by adding 15 ml of Elution buffer (50 mM Phosphate, 0.3 M NaCl and 250 mM Imidazole, pH adjusted to 8.0). The samples were analyzed on an 8% SDS PAGE. The eluted protein is then dialyzed against 1L of TGED plus 0.1 M NaCl, with two buffer changes O/N. Next day, the protein was centrifuged at 12000 rpm for 15 minutes at 4 °C to remove any protein aggregates. The supernatant was then diluted with TGED with 0.1 M NaCl and was loaded on to a Heparin column (Amersham), which was earlier equilibrated with TGED with 0.1M NaCl. RNAP was eluted with a 2 ml linear gradient of NaCl (0.2-1.0 M). The pure RNAP fraction elutes out in the range of 0.4-0.65 M NaCl. The fractions were checked for the presence of protein by running in an 8% SDS PAGE. The pure fractions were pooled and

were concentrated in a 100 KD concentrator (Amicon) against 2x storage buffer (80 mM Tris-HCl; 0.4 M KCl; 2 mM EDTA, 2 mM DTT). Glycerol to the final concentration of 50% v/v was added to the concentrated pure protein, aliquoted into fresh sterile eppendorf tubes and stored in $^{-}$ 70 °C deep freezer.

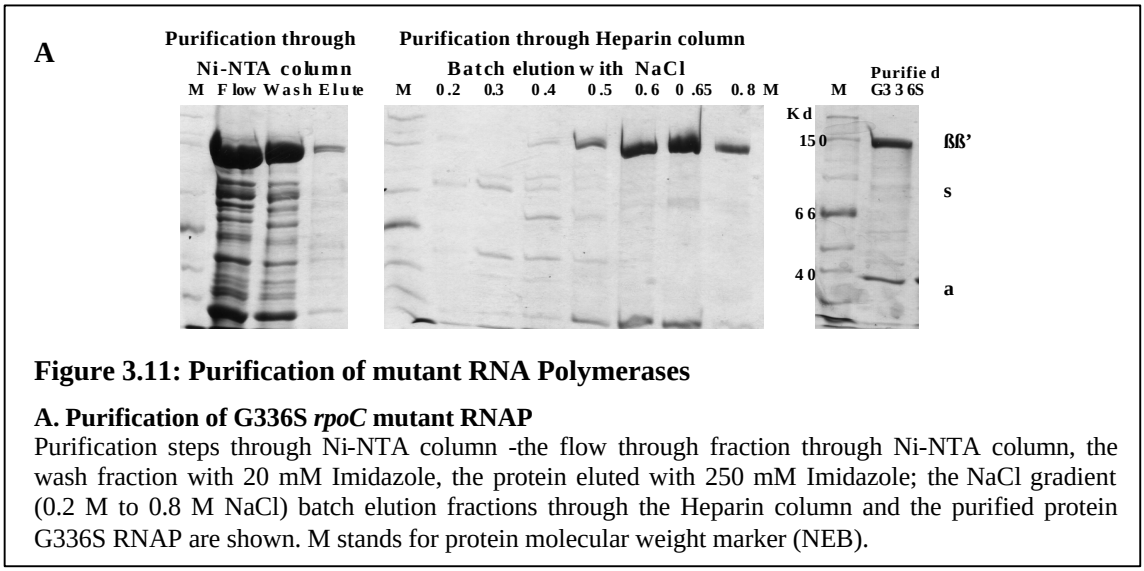
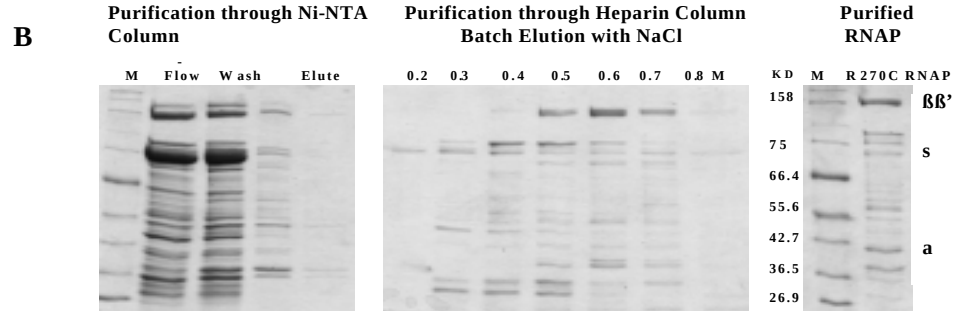
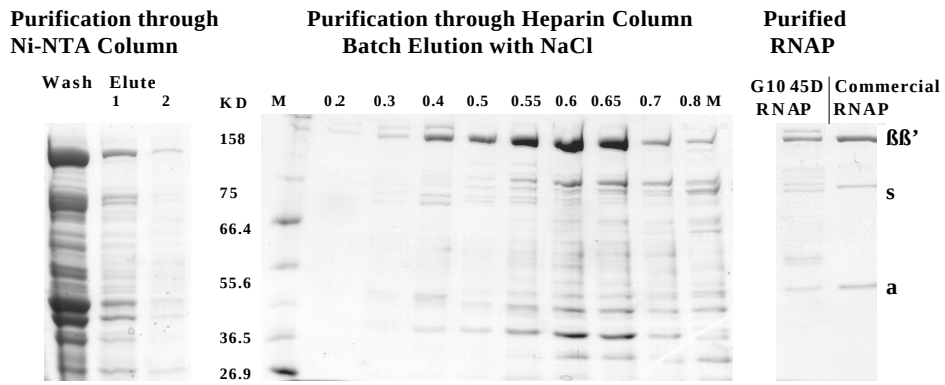


Figure 3.11: Purification of mutant RNA Polymerases

A. Purification of G336S *rpoC* mutant RNAP
Purification steps through Ni-NTA column -the flow through fraction through Ni-NTA column, the wash fraction with 20 mM Imidazole, the protein eluted with 250 mM Imidazole; the NaCl gradient (0.2 M to 0.8 M NaCl) batch elution fractions through the Heparin column and the purified protein G336S RNAP are shown. M stands for protein molecular weight marker (NEB).



Purification of R270C *rpoC* mutant RNAP
Purification steps through Ni-NTA column -the flow through fraction through Ni-NTA column, the wash fraction with 20 mM Imidazole, the protein eluted with 250 mM Imidazole; the NaCl gradient (0.2 M to 1.0 M NaCl) batch elution fractions through the Heparin column; and the purified R270C RNAP are shown. M stands for protein molecular weight marker (NEB).



C. Purification of G1045D *rpoB* mutant RNAP

Purification steps through Ni-NTA column - the wash fraction with 20 mM Imidazole, two fractions of protein eluted with 250 mM Imidazole; the NaCl gradient batch elution fractions through the Heparin column and the purified G1045D RNAP are shown. Commercial RNAP stands for wild type RNAP purchased from Epicenter. M stands for protein molecular weight marker (NEB).

3.2.11 *In vitro* transcription

Single round *in vitro* transcription assays like the one we employed helps to examine the kinetics of transcription, subassemblies of various transcription complexes, termination and antitermination. The run off transcription assay monitors directly the synthesis of the terminated and runoff transcripts. Typically, the reactions were carried out in two steps; in the first step, transcription complexes poised at a specific template location were generated by incubating RNA polymerase and the template DNA with small amounts of rNTPs and a dinucleotide primer specific for initiation at the desired promoter. Here a ^{32}P-CTP is one of the rNTPs for the labeling of short transcript synthesized. In the second step, further elongation is carried out with the addition of an equal volume of pre-warmed transcription buffer that contains necessary elongation factors (N can be added just prior to this step for antitermination assays) as well as higher amounts of all four rNTPs. The dilution of the label with almost four fold excess of unlabelled NTP yields 5'-endlabelled transcripts from the first round of synthesis and diminishes the labeling of transcripts in further rounds. Rifampicin was also added to block new rounds, when required. The reactions were terminated by the addition of phenol or Formamide loading

dye. Transcripts were resolved by running on 6%, 8% or 10% Acrylamide-6 M Urea sequencing gels.

3.2.11.1 *In vitro* transcription protocol for antitermination assay

Transcription reactions were done at 32 °C in transcription buffer (T-Glu) containing 20 mM Tris-Glutamate pH 8.0, 10 mM Mg-Glutamate and 50 mM K-Glutamate. The initial mix contained RNA Polymerase (20 nM final concentration), 5-10 nM DNA template (PCR amplified using the specific plasmid construct as template DNA, phenol extracted and concentrated through a 100/50 KD cut off concentrator Millipore), along with 1mM DTT and 0.1 µg/ml BSA (both of them from 10x stocks made using RNAase free water, Ambion). Reactions were initiated with by adding the solution containing 175 µM ApU, 5 µM each of GTP and ATP and 2.5 µM of CTP and a ^{32}P-CTP to the RNAP-DNA mix, to make a 22-mer, labeled EC. This was incubated for 7 minutes at 32 °C. Then as per requirement of the reaction, this mix was aliquoted into different eppendorf tubes. N protein (Wt or ARM/CTD) was added to the 'with N' tubes, equal amounts of transcription buffer was added to the 'without N' tube. It was then chased by the addition of a mix containing 400 µM each of all the rNTPs, NusA, NusG along with BSA and DTT and 1x Tris-Glu buffer. Under these conditions, reactions were essentially single round as the labeled CTP was diluted more than 1000-fold by the cold CTP. After chasing (depending on the template size, from 2-3 minutes to 15 minutes), the reactions were stopped by addition of phenol, followed by ethanol precipitation in presence of glycogen and Ammonium Acetate (both from Ambion). Incubated at ⁻20 °C for half an hour, spin at 13000 rpm at 4 °C and were washed with 70% ethanol; the pellets were air dried at RT/37 °C. The dried down RNA pellets were then re-suspended in Formamide loading dye (Ambion), heated at 95 °C for 4 minutes and load onto a 6%/ 8%/ 10% Urea- sequencing gel and run with 1x TBE buffer at 1500 volts for 8% and 1800 volts for 6%. (The gels were pre-run for 20-30 minutes before loading).

3.3 Results

3.3.1 Genetic screen for isolation of N specific substitutions in ß and ß' subunits of RNAP

We designed a genetic screen, for isolating N specific mutants of RNA polymerase. The first objective in the design of the strain for screening N specific mutants was to incorporate an H-19B N specific *in vivo* reporter system. The background *E.coli* strain selected was K3093 (K37 sup° Tn5::*lacZ lacI* $_{Q1}$ (Neely and Friedman, 1998b, see Appendix 1). A single copy lysogen (P_{lac}-H19B- *nutR* –3T-*lacZYA*) from pK8641 (Neely and Friedman, 1998b) was transduced into RS20 using ? RS45 phage vector (see Materials methods). Here, the DNA cassette is designed such that, it will serve as an *in vivo* antitermination probe. The transcripts starting from the *pLac* promoter will be terminated at the three terminators (T_{R1}, T1, T2), unless H-19B N protein binds to the *nutR* site and brings about transcription antitermination. In such a case, the RNAP can read through these three terminators and express the *lacZYA* genes located further downstream in the cassette. Under such conditions, the otherwise *lac*$^-$ strain will be able to express the Beta Galactosidase enzyme and appear blue in a plate with medium supplemented with X Gal IPTG. To annul the effect of the chromosomal allele of *rpoB/C* genes, we made this alleles temperature sensitive at a non-permissive temperature. The strains were made temperature sensitive (*ts*) for the *rpoC* allele by moving *rpoC* 120 (rif^s) *btuB::Tn10* mutation from the strain DJ354 (Din Jin) and *ts* for the *rpoB* allele by introducing *rpoB ts* 3401 (Jin et al., 1988a) by P1 transduction (refer materials methods). These strains grew in the permissive temperature of 30 °C, while they could not grow at 42 °C. For the supply of H-19B N protein in *trans* in these strains, all these strains were then transformed with the H19BN encoding plasmid pK8601 (Neely and Friedman, 1998b). Plasmid encoding wild type *rpoC* (pBAD18M) and *rpoB* (pHYD534) (refer materials methods) was transformed into the respective background strain and was found to fully complement for the *rpoB/C* function at 42 °C. In the presence of the Wild type (Wt) H-19B N expressing plasmid, the wild type RNAP overcomes the cluster of the three intrinsic terminators (T_R-T1-T2), expressed Beta Galactosidase from the reporter system; and gave rise to blue colonies on X Gal IPTG plates, when grown at 42 °C. These strains with the *pLac-nutR*-3T –*lacZYA* cassette, the *rpoC/rpoB ts* mutation (separately in different strains) and harboring the plasmid pK8601 which provides the H-19BN gene product in *trans*, served as the background strains for N specific mutant RNAP screening (Figure 3.1). The approach for isolating N specific mutants was that, on transformation into such a background strain (*rpoB* encoding plasmid to *rpoB ts* strain and *rpoC*

encoding plasmid to *rpoC ts* strain) of an *rpoB/C* plasmid that has a mutation, which is specific to defective N interaction, will make the RNAP defective for transcription antitermination through the three terminators present. Such a strain would appear white on an XGal IPTG plate with all the necessary antibiotics required, as there is no read through to the *lacZYA* genes present downstream to the terminators. We expected that this screen would isolate mutations in *rpoB* and *rpoC* genes, which do not interfere with the viability of the bacterium, but which were specifically defective for N mediated transcription antitermination. Figure 3.12 gives a schematic representation for the strategy employed for isolation of N specific mutations.

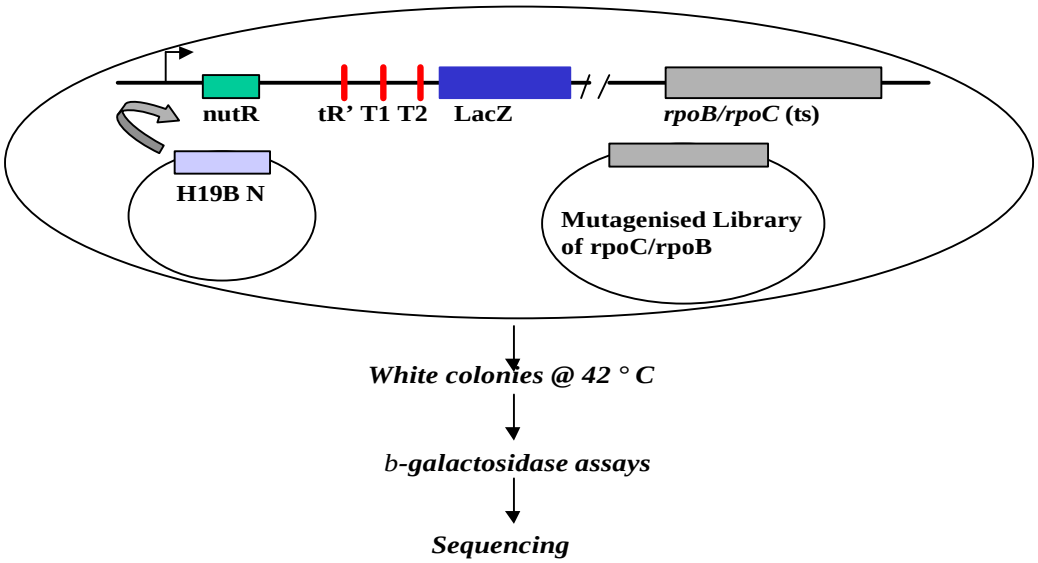


Figure 3.12: Schematic representation of the genetic screen employed for identifying N specific RNAP mutants

3.3.2 Screening and isolation of mutants

The *rpoB/C* plasmid library was transformed by electroporation (refer Appendix III) into the background strains having the *rpoB ts or rpoC ts* allele, which will be compatible with the mutagenized plasmid library. Around 1.2×10^6 colonies for *rpoB* and 0.95×10^5 colonies for *rpoC* were screened at 42 °C. White mutants were picked, purified thrice by streaking in the same medium and the *rpoB/C* plasmid was isolated using plasmid miniprep kits from Qiagen. These plasmids were then retransformed into their respective background strain to see if they retain the mutant phenotype or not. The confirmed white mutants were then subjected to Beta Gal assays

(refer Material methods). The mutant plasmids were transformed into strains RS69 and RS83 respectively for the P_{Lac}-*nutR*- ΔTerminator (ΔT) fusion in the *rpoC* and *rpoB ts* strain background. The mutant plasmids in the 3T and ΔT strain background were grown in LB media with all the antibiotics and IPTG at 42 °C. % Read through (which denoted the *in vivo* antitermination efficiency) was calculated as the ratio between the Beta Gal values obtained for the mutant plasmids in the 3T and ΔT strain background. Wild type (Wt) *rpoC/B* plasmid transformed into the respective background strains served as the controls.

3.3.3 *In vivo* antitermination assay through Beta Gal assays

We measured the amount of Beta-Galactosidase made from the single copy reporter fusion both in the presence or absence of triple terminator cassette present upstream of the *lac ZYA* genes at 42°C. At this temperature, the chromosomal copies of *rpoB* or *rpoC* will be inactive and the active holoenzyme will be made only from the Wt or mutant copy of the *rpoB* and *rpoC* genes supplied from the plasmid. H-19B N was supplied from the plasmid pK8601. Table 1 describes all the *in vivo* measurements. Read through efficiency (%RT) at the end of the triple terminator cassette (Table 1) is the measure of antitermination. In our assays, we obtained about 40% read through for WT RNA Polymerase, which is less than the previously reported activity of the plasmid borne multicopy reporter (Neely and Friedman, 2000). The discrepancy could perhaps arise from the difference in plasmid copy numbers between the cells having fusions with and without terminators on the plasmid. The *rpoB* mutants have about 10-fold less read through efficiency compared to WT while the mutations in *rpoC* reduce read through efficiency by 4 to 10-fold. The double mutant P251S P254L exhibited the least *in vivo* antitermination rates among the *rpoC* mutants. In the absence of N, the read through efficiency ranged from 1 to 2% and was similar in Wt and all the mutants, which suggest that these mutations are impaired for N-dependent antitermination. We also moved one of these mutations G1045D into the chromosome of RS117.We grew the mutant RS117 with plasmid encoded ß G1045D mutation at 37 °C and in the absence of Ampicillin in the media, required for the maintenance of the plasmid. After prolonged growth for several generations under these conditions, colonies expressing white color on XGal IPTG plates were screened for. Two out of a thousand colonies screened has this phenotype. The *rpoB* gene was amplified from the genomic DNA of these two strains and was sequenced with *rpoB* specific primers. It was found that G1045D mutation had recombined into the chromosome and the *ts* defect in the chromosome reverted back to wild type. This chromosomal *rpoB* G1045D mutant (named as RS223) retained all the N specific antitermination

defective phenotype of the plasmid borne *rpoB* G1045D mutant. We also moved in the *rpoB* G1045D mutation to the chromosome of RS20 by P1 transduction. Here a recombinant P1 lysate was made by growing P1 on strain RS117, which has plasmid encoded ß G1045D mutation. This recombinant P1 lysate was used for transduction into RS20 and the transductants were screened for white color after the transformation of the plasmid harboring the N gene (pK8601) into the transductants and growing on XGal IPTG plates. The white colonies were purified and the *rpoB* gene was amplified from the genomic DNA of this strain and was sequenced with *rpoB* specific primers and the presence of G1045D mutation was confirmed and strain was named as RS232.

Table 1: H-19B N specific mutations of *rpoB* and *rpoC* and *in vivo* antitermination rates

Mutations in rpoB and rpoC	***+H-19B N b-galactosidase (M.U.)***			***-H-19B N b-galactosidase (M.U.)***		
	+3T	**3T**	**%RT**	**+3T**	**-3T**	**%RT**
rpoB						
WT	299 ± 8	856.5 ± 15	**35.0**	8 ± 0 .5	568 ± 20	**1.4**
G1045D	22 ± 3	873 ± 26	**2.5**	8 ± 0.1	540 ± 47	**1.5**
G1045D, G544S	26 ± 2	879 ± 24	**3.0**	5 ± 1.2	572 ± 33	**0.9**
G1045D, D185N	30 ± 4	868 ± 20	**3.5**	10 ± 1.2	578 ± 5	**1.7**
G1045D, V296M	37 ± 5	913 ± 23	**4.0**	10 ± 0.6	579 ± 15	**1.7**
G1045D, G162S	30 ± 5	929 ± 22	**3.2**	11 ± 0.7	551 ± 32	**2.0**
G1045D, R903H	22 ± 5	712 ± 35	**3.1**	5 ± 0.8	565 ± 66	**0.9**
rpoC						
WT	331 ± 28	854.5 ± 22	**39.0**	8 ± 0.9	606 ± 16	**1.4**
P254L	58 ± 2	900 ± 28	**6.4**	10 ± 1.1	519 ± 44	**1.9**
P251S, P254L	23 ± 5	929.5 ± 29	**2.5**	4 ± 0.3	533.4 ± 41	**0.7**
P251S[#]	78 ± 2	861 ± 25	**9.0**	6 ± 1.0	545 ± 23	**1.1**
R270C	50 ± 9	788 ± 30	**6.3**	4 ± 0.3	341 ±26	**1.1**
G336S	96 ± 4	888 ± 26	**10.8**	9 ± 0.8	521 ± 28	**1.1**

Beta Galactosidase values were average of 4 to 9 independent measurements. +3T and –3T denote Beta Galactosidase values in the presence and absence of terminators. %RT is the read through efficiency at the end of the triple terminators and was calculated by taking the ratios of +3T and –3T values. [#] This mutant was obtained by site directed mutagenesis.

3.3.4 *In vivo* antitermination assay through phage spotting

We tested the efficiency of the mutants to support the growth of the H-19B and λ phages.

It was observed that all the mutants were severely defective in supporting the growth of H-19B phage. On the other hand, growth of λ phage was normal in the *rpoB* mutants and only partially affected in *rpoC* mutants (Table 2). The *Nin* (N-independent; deletion of *Nin* region) derivatives of both the phages grew equally well on the strains with WT and mutant RNAPs. These data strongly suggest that the mutants are specifically defective for H-19B N-mediated antitermination and that they are unaffected for its Q-mediated antitermination. The *Nin* deletion removes the terminators present at that region and allows a TEC with no N modified antitermination property to reach the site in DNA encoding the Q site. If the RNAP mutants were defective for Q function also, one would expect that the phage would not be able to grow well, as it cannot antiterminate past the terminators into the lytic genes via a Q modification of RNAP. Since no such defect was observed for the *Nin* deletion phages, we could conclude that the mutants were not defective for Q function. The partial defects of *rpoC* mutants in supporting the growth of λ phage suggests that the target site of λ N could also be close to this region.

Table 2: *In vivo* antitermination assay of the mutants by phage spotting

Mutations in rpoB and rpoC	***H-19B growth (E.O.P)***		***l growth (E.O.P)***	
	WT	***Nin***	**WT**	***Nin***
rpoB				
WT	(1)	(1)	(1)	(1)
G1045D	~ 10^{-5}	1	1	1
G1045D, G544S	~ 10^{-5}	1	1	1
G1045D, D185N	~ 10^{-5}	1	1	1
G1045D, V296M	~ 10^{-5}	1	1	1
G1045D, G162S	~ 10^{-5}	1	1	1
G1045D, R903H	~10^{-5}	1	1	1
rpoC				
WT	(1)	(1)	(1)	(1)
P254L	~10^{-4}	1	0.1	1
P251S, P254L	~10^{-5}	1	0.03	1
P251S[#]	~10^{-5}	1	0.1	1
R270C	~10^{-4}	1	0.15	1
G336S	~10^{-3}	1	0.1	1

Relative efficiency of platting of H-19B and λC1857 phages is shown in column 1 and 2, respectively. Number of plaques obtained with WT type enzyme was set as 1.

[#] This mutant was obtained by site directed mutagenesis.

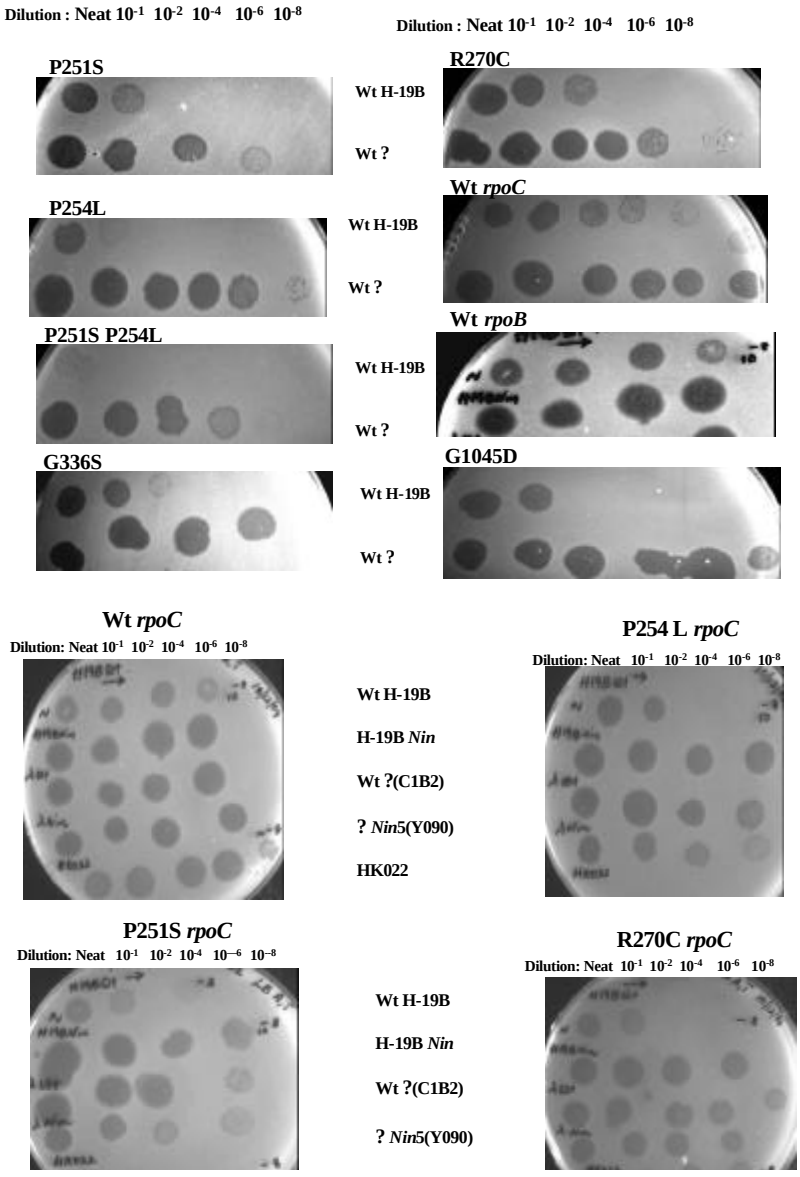


Figure 3.13: *In vivo* antiterminati on assay by phage spotting
A. Representation of phage spotting assay, with H-19 B phage on the mutants; dilutions of the phages used for spotting is given on top. **B**. Phage spotting assay, with H-19 B, H-19B *Nin*, Wt ? (?C1B2) and ? *Nin5* phages on the *rpoB/C* mutants; dilutions of the phages used for spotting is given on top.

3.3.5 *In vitro* characterization of the RNAP mutants

We established an *in vitro* transcription assay to check the efficiency and specificity of H-19B N mediated antitermination at the triple terminator ($T_{R'}$-T1-T2) cassette (Figure 3.3), the similar one that has been used in the *in vivo* experiments. We used a linear DNA template, bearing a *nutR* site followed by the triple terminator cassette. Transcription was initiated from the strong T7A1 promoter. Here, as RNA is being synthesized *in vitro*, H-19B N protein can bind to the *nut* site (coded by the *nutR* in the template DNA) and along with the host elongation factors supplied can interact with the RNA polymerase, forming a functional antitermination complex, *in vitro*. In the absence of N, the transcription terminates at one of the three terminators present in the template. But, when N is supplied and when it is able to form a functional antitermination system, RNAP can read through the terminator and 'Run Off' RNA (full length RNA) is made. The RNA was radiolabelled and products were analyzed on a 6% Urea sequencing gel. The antitermination rates were calculated by comparing the run off RNA band intensity between the –N and +N RNA transcripts.

3.3.5.1 Processive antitermination with H-19BN and *in vitro* antitermination assays with the mutant RNAPs

We observed that *in vitro*, H-19B N requires both NusA and NusG for efficient and processive antitermination (where as NusA alone is reported to be sufficient for processive antitermination) and that the read through efficiency was similar to that observed *in vivo*. Though it has not been shown earlier, our results show that NusG increases processivity of the TEC. Recently it has been demonstrated that NusG is essential for λ N mediated antitermination *in vivo* (Zhou et al., 2002). Though it is not tested for H-19B phage, it is likely that NusG will also be required *in vivo*. Though it is not tested here, it is very likely that NusG is required *in vivo* for H-19B phage also, based on our observation that *in vitro*, H-19B N requires both NusA and NusG for efficient and processive antitermination. We could obtain around 39% antitermination through the triple terminator cassette in the presence of NusG. The read through efficiency obtained was similar to that observed *in vivo* (Figure 3.13, part A). The observed antitermination was N specific, since mutations in the RNA binding motif (ARM) of N as well as the deletion of *nutR* site prevented antitermination, even when N protein was supplied at high concentrations (Figure 3.14, part B).

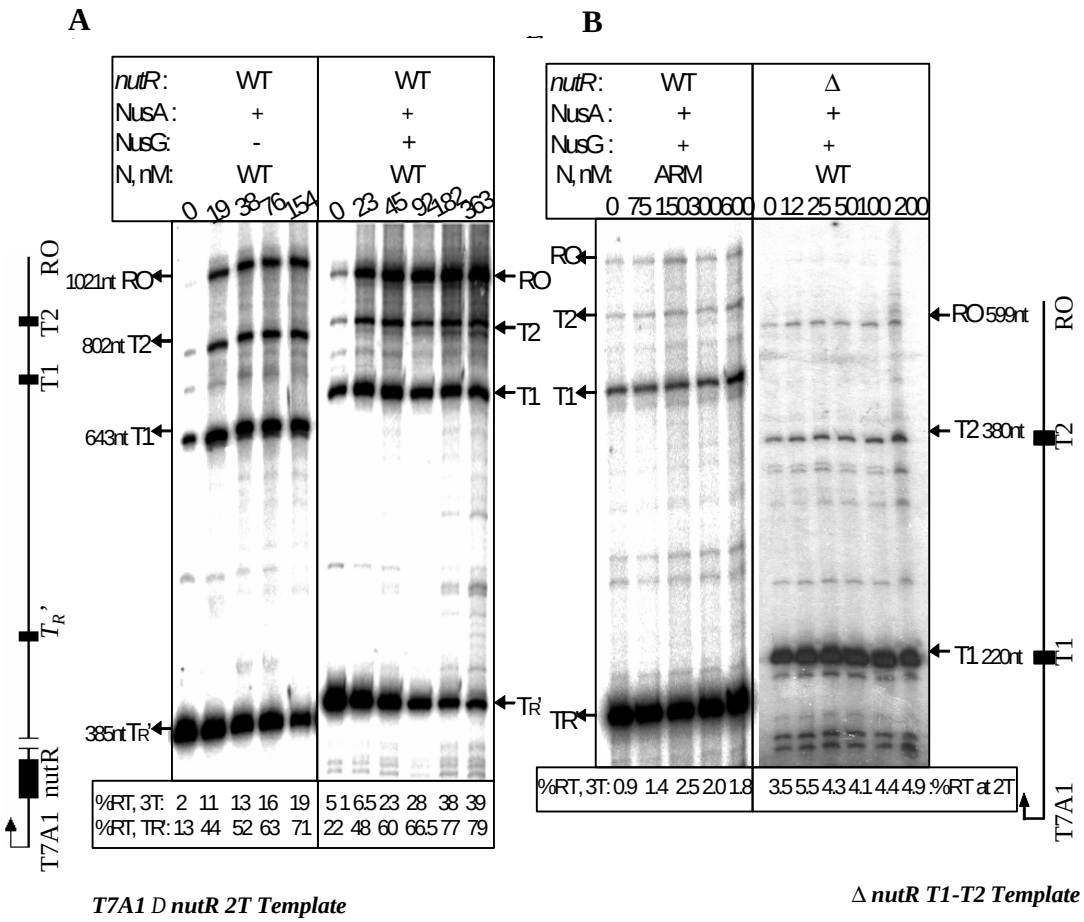


Figure 3.14: A functional *in vitro* antitermination system with H-19B N

Autoradiograms showing the *in vitro* single round transcription assays under different conditions. Linear DNA templates with and without (Δ*nutR*) *nutR* sites are shown on the both sides of the autoradiograms with the length of the transcripts terminated at different terminators (T1 and T2) and at the end of the template (RO). Transcription was initiated from T7A1 promoter. Concentrations of DNA template, RNA polymerase, NusA and NusG were 5nM, 20nM, 300nM and 200nM, respectively. Concentrations of N are indicated. Mutations in N-terminal Ariginine rich motif of N are denoted as ARM.

Antitermination efficiency (%RT) of the three terminators and of the first terminator was calculated from the amount of transcript (band intensity) at different terminators as follow. %RT, at 3T = [RO]/[+ T1 + T2 + RO], %RT, at $T_{R'}$ = [T1 + T2 + RO]/[+ T1 + T2 + RO] and for Δ *nutR* template, % RT, at 2T = [RO]/[T1 +T2 +RO].

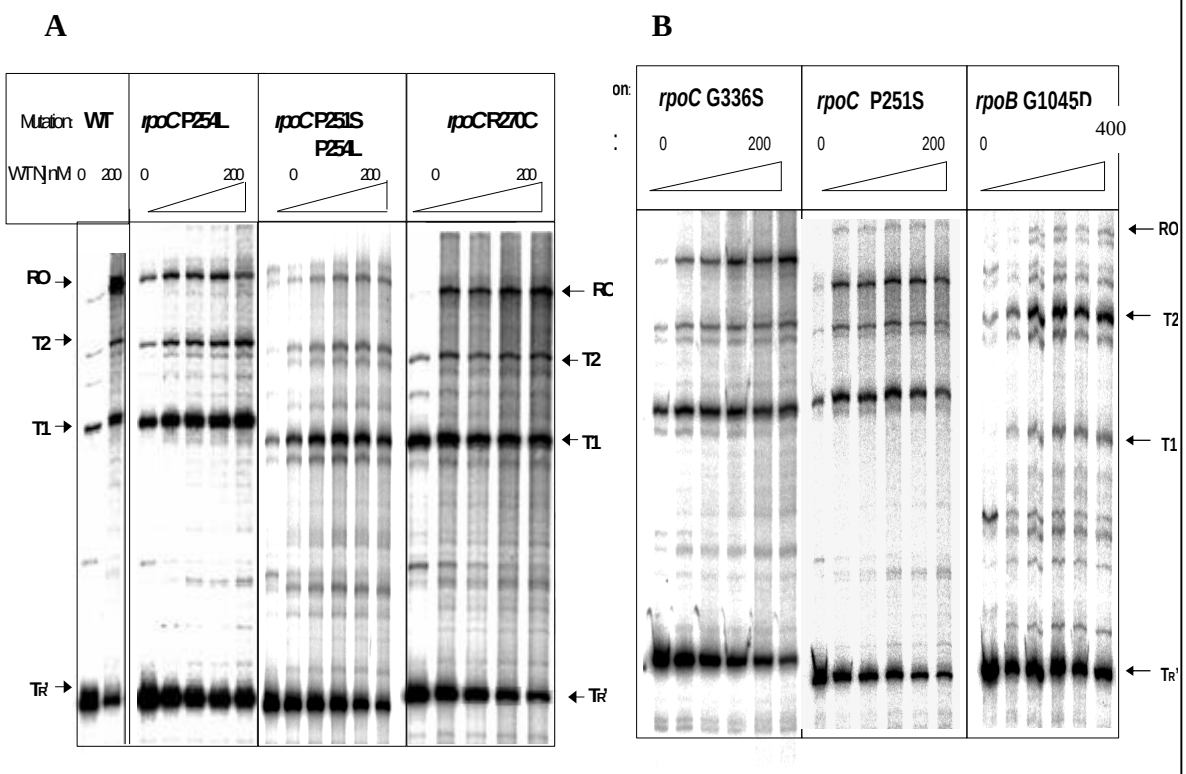


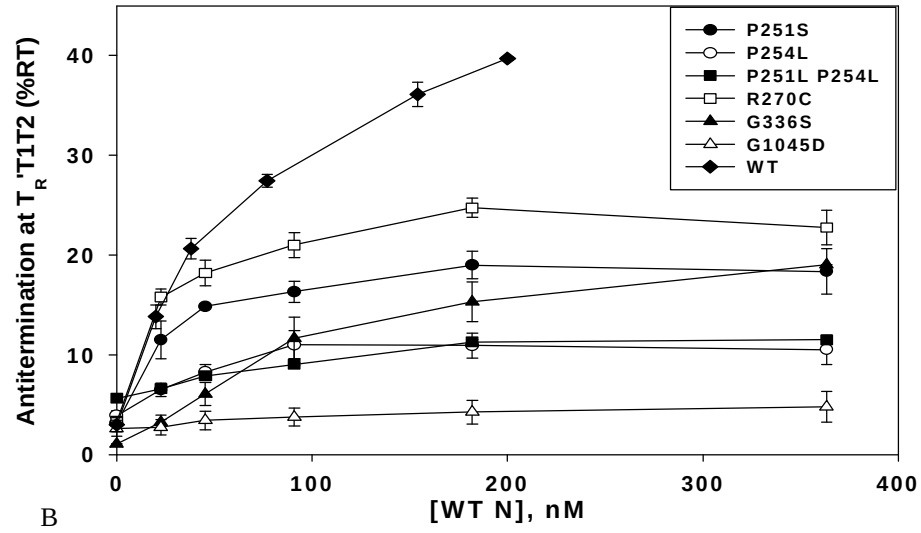
Figure 3.15: *In vitro* antitermination assays with the mutant RNA polymerases

In vitro transcription assays with the mutant RNAPs were done using the same amounts of the triple terminator template, NusA and NusG and under the same conditions as described in Figure 3.14. Concentrations of mutant RNAPs were ~100nM.

A, **B** depicts autoradiograms of *in vitro* transcription with the mutants as indicated.

Concentrations of H-19B N protein were 0, 25, 50, 100 and 200 nM for *rpoC* mutants and 0, 25, 50, 100, 200 and 400 nM for *rpoB*-G1045D mutant

A



B

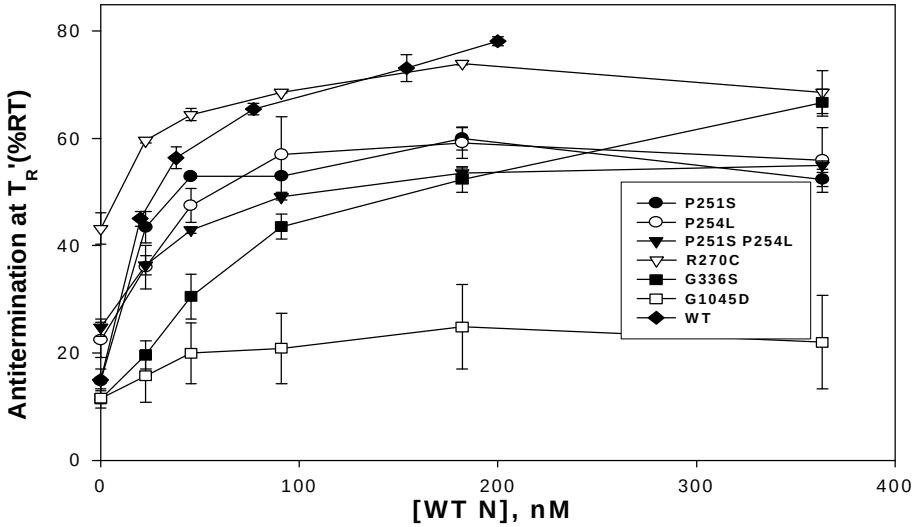


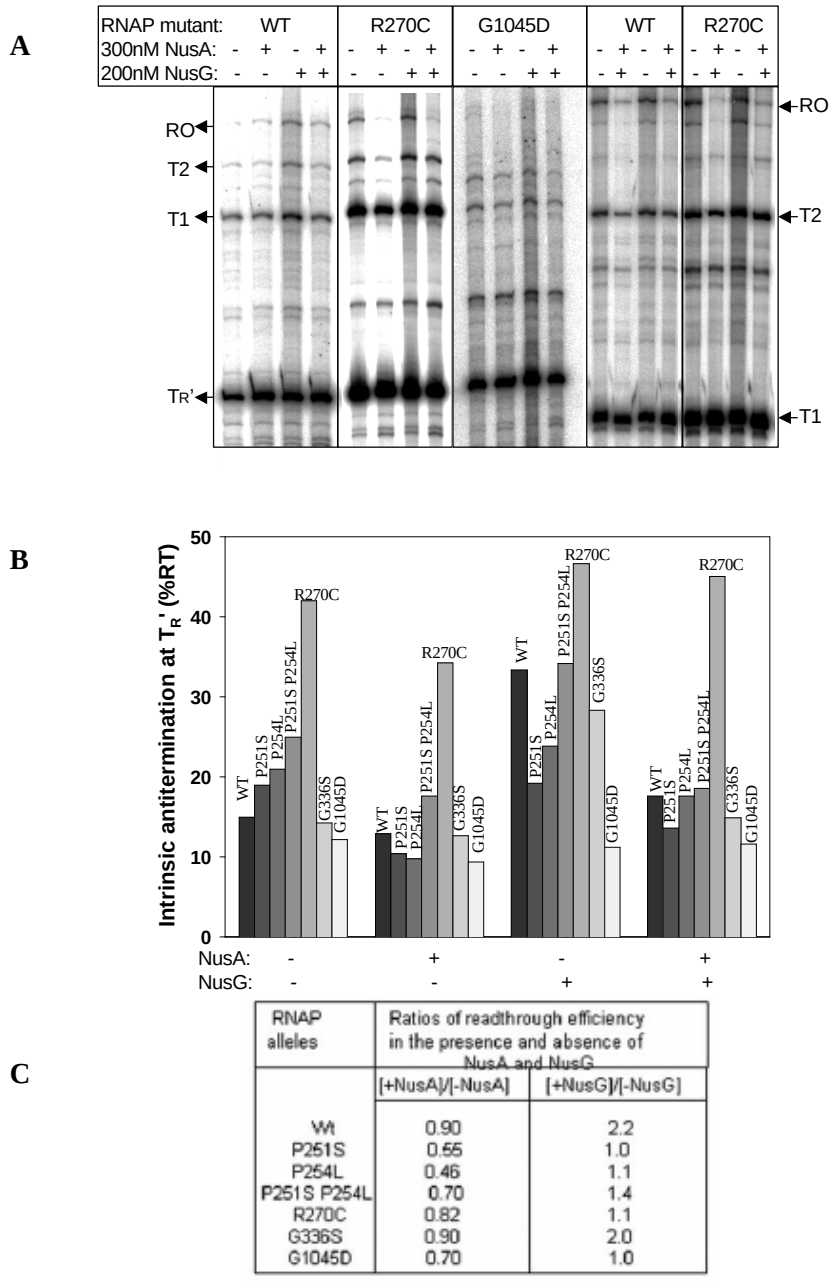
Figure 3.16: Plots representing the antitermination efficiency (%RT) of the mutants

A. Antitermination rates through the three terminators ($T_{R'}$- T1-T2) for WT RNAP and all the *rpoB/C* RNAP mutants, against increasing concentration of H-19B N.
B. Antitermination rates through the first terminator ($T_{R'}$) for WT and all the mutants, against increasing concentration of H-19B N. Symbols of each plot are described in the inset.
Antitermination efficiency (%RT) of the three terminators and of the first terminator was calculated from the amount of transcript (band intensity) at different terminators as follow. %RT, at 3T = [RO]/[+ T1 + T2 + RO], %RT, at $T_{R'}$ = [T1 + T2 + RO]/[+ T1 + T2 + RO] and for Δ *nutR* template, %RT, at 2T = [RO]/[T1 +T2 +RO].

As could be inferred from comparing the results depicted in Figures 3.14 and 3.15, antitermination efficiency (%RT) at the end of triple terminators was reduced between 2 to 8 fold with different mutants (Figure 3.16). The ß -G1045D RNAP was found to be most defective, whereas ß'-R270C exhibited about 20% antitermination at high concentration of N (compared to 40% read through with WT RNAP). Except for ß -G1045D, higher concentrations of N protein partially suppressed the antitermination defect for most of the mutants. Interestingly, except for ß -G1045D; at higher concentrations of N, all the mutants could antiterminate the first terminator $T_{R'}$, close to the WT efficiency (Figure 3.16 B). This could mean that N dissociates more rapidly from the mutant ECs as the length of the RNA-tether between *nut* site and the EC increases, which in turn decreases the local concentration of N. All this points to the weaker binding affinity of the ß' mutants for N. Higher concentrations of N could not overcome the defect of ß -G1045D, which indicates that this position of ß subunit may be involved in a very crucial interaction. It is also possible that G1045D mutation impairs some other steps in antitermination.

3.3.5.2 *In vitro* N-independent antitermination of the RNAP mutants

We went on to probe the intrinsic termination defects of the mutants. The termination properties of each of these mutants at the triple terminator cassette were assayed using the same template as above, either in presence or absence of NusA and NusG, but in absence of H-19B N. Figure 3.17 A shows the three representative autoradiograms for WT, ß'-R270C and ß -G1045D holoenzymes. Figure 3.17 B summarizes the termination defects, which is expressed as antitermination efficiency (%RT; inverse of termination efficiency) at $T_{R'}$ terminator and Figure 3.17 C quantitatively describe the effect of NusA and NusG on termination for all the mutants.



RNAP alleles	Ratios of readthrough efficiency in the presence and absence of NusA and NusG	
	[+NusA]/[-NusA]	[+NusG]/[-NusG]
Wt	0.90	2.2
P251S	0.55	1.0
P254L	0.46	1.1
P251S P254L	0.70	1.4
R270C	0.82	1.1
G336S	0.90	2.0
G1045D	0.70	1.0

Figure 3.17 N independent antitermination of RNAP mutants

In vitro transcription reactions were done under similar conditions as described earlier, either in presence or absence of NusA and NusG; but in the absence of H-19B N.

A. Autoradiogram showing the transcription assays with WT, β'-R270C and β-G1045D RNAPs **B.** Antitermination efficiencies describe at the first terminator ($T_{R'}$) for WT and for all the mutant RNAPs in presence and absence of NusA and NusG.

C. Ratio of %RT in the presence and absence of NusA or NusG. %RT was calculated according to that described in Figure 3.16

From these experiments, it was observed that in the absence of any host factors, all the mutants, except ß'-R270C, exhibited antitermination efficiencies within the range of 15 % to 25% which is similar to that of WT. ß'-R270C showed a specific defect in termination at $T_{R'}$ as the read-through transcripts past this terminator was efficiently terminated at the second terminator, T1 (Figure 3.16 A). We also confirmed that this is a specific defect for $T_{R'}$, by performing separate assays with the template containing only the T1 and T2 terminators (? T1-T2) template, where this mutant did not show any anomaly. The reason for this behavior of ß'-R270C is not clear and we have not explored this phenomenon any further. In accordance with previous studies (Rees et al., 1997); for the WT RNAP, about 15% antitermination which was observed at $T_{R'}$ terminator was further decreased to about 10% in presence of NusA and increased to about 35% in presence of NusG. Termination defects of ß'-P251S, ß'-P254L, ß'-P251S-P254L, and ß G1045D mutants were more efficiently corrected by NusA. Other mutants responded to NusA in a similar way as the WT RNAP holoenzyme (Figure 3.16 C). On the other hand, NusG did not stimulate the read-through activity significantly for any of the mutant RNAPs, except for ß' G336S (Figure 3.16 C). As all the mutants were viable, this *in vitro* defect could be specific for $T_{R'}$ terminator and may not reflect a general binding defect for NusG. Though the mutants exhibited a varied range of response to NusA and NusG in N-independent terminations assays, the antitermination defect at the terminator was negated at higher concentrations of N for all the mutant RNAPs, except that for G1045D. Hence it is likely that the antitermination defect observed is more due to weak N-binding and than due to impaired binding of the Nus factors. In case of the ß -G1045D RNAP, however, we cannot rule out the possibility that defective antitermination exhibited by it could also involve the defective binding of NusA and or NusG to it in the antitermination complex.

3.3.6 Isolation of suppressors from H-19B for the N specific RNAP mutants

We employed a similar genetic screen as described for the isolation of the RNAP mutants for screening suppressors in H-19B N for the RNAP mutations. We transformed a random mutagenized library of the plasmid containing H-19B N gene into the background strains, each of which carried an antitermination-defective RNAP mutation and lacking the Wt N gene product in *trans* (these strains lack pK8601 plasmid) and then screened for *lacZ* expression. In this work, we had concentrated only on the unique *rpoB/C* mutants- G1045D *rpoB* (both Beta Galactosidase assays and phage growth tests showed that all the double mutants of *rpoB* have the phenotype of the single G1045D mutant and for the subsequent studies we chose only G1045D mutation and ignored other second-site mutations of *rpoB,* other than R903H, which later on we obtained as a single substitution with site directed mutagenesis) and P251S, P254L, R270C, G336S *rpoC*

mutants. The logic of the screen was that from the mutagenized library of N, if there were mutants of H-19B N that will specifically suppress the RNAP substitution and its resultant antitermination defect, such transformants would appear blue in color on X Gal IPTG plates. Here, the mutant RNAP expressing strains having the Wt N gene will be white (due to the N specific defect of RNAP and hence the failure to antiterminate), so will be the transformants with non-specific mutations in N gene. Only a transformant, with an H-19B N gene having a mutation, which can specifically suppress the N specific defect, can form a functional antitermination complex *in vivo* and antiterminate through the terminators, expressing Beta Galactosidase and forming blue colonies on X Gal IPTG plates. Such blue colonies were picked, and were purified by retransformation to the same background strain with the mutant RNAP to confirm the mutant phenotype and were sequenced with a series of *rpoB/C* gene specific primers.

3.3.6.1 Screening for the *rpoB/C* substitution specific suppressor in H-19B N

The mutagenized gene library was transformed into each of the *rpoC* mutant background strains (P251S, P254L, R270C and G336S) as well as to the G1045D *rpoB* mutant. Despite exhaustive screening against each of these mutants (for each mutant, above 100,000 colonies were screened), we could obtain a suppressor in N for one *rpoC* mutant. The mutant library transformed into P254L yielded a blue colored colony, which retained its phenotype after the isolation of the N plasmid and its subsequent re-transformation back into the Del N ß' P254L background strain. Beta Galactosidase assays was carried out to see if the suppressor could bring the antitermination read through to near about wild type levels in the P254L strain. This plasmid was sequenced and a single mutation, L108F was found in the C terminal region of H-19 B N gene. The C-terminal domain of N is required for binding to RNAP (Mogridge et al., 1998). The *in vivo* antitermination assays revealed that in the presence of L108F N, efficiency of antitermination of the ß'-P254L mutant RNAP increased to the level of antitermination efficiency of Wt RNAP obtained in the presence of WT N (see Table 3 for comparison). To check the allele specificity of this suppressor N, we also transformed it into all the mutant RNAP background strains and Beta Gal assays were carried out as before. The specificity of the suppressor N was checked by transforming it into the background strain lacking N plasmid (Del N) strains; harboring the ß' P251S, ß' P251SP254L double mutant, ß' G336S, ß' R270C, ß G1045D, Wt *rpoB* and Wt *rpoC* plasmid separately. Wt pK8601 transformed into the each of the respective background strain served as the control. Beta Gal assays was carried out to obtain the *in vivo* antitermination rates. L108F H-19B N partially suppressed the antitermination defect of ß' P251S and ß' G336S but had no effect on ß' R270C and ß G1045D. The antitermination efficiency of Wt RNAP was

about 2-fold reduced in presence of this mutant N. These observations could be related to the spatial location of the different RNAP mutations (discussed in later section). ß' P254L is close to ß' P251S and ß' G336S in the structural model of EC, where as ß G1045D and ß' R270C form another cluster.

Table 3: Suppressor mutations in H-19B N

Mutations in H-19B N	***Mutations in rpoB and rpoC***		***b-galactosidase (M.U.) +3T***	***-3T***	***Antitermination (%RT)***
L108F	***rpoB***	WT	161 ± 2	905 ± 24	17.7
		G1045D	31 ± 5	839 ± 23	3.7
	rpoC	WT	197.5 ± 3.5	917.3 ± 10	21.5
		P251S	173.9 ± 10.5	858.5 ± 28.5	20.2
		P254L	314.4 ± 7	862.4 ± 8	**36.5**
		P251S, P254L	153.5 ± 7.9	887 ± 29.4	17.3
		R270S	93.0 ± 3.2	863 ± 31.5	10.8
		G336S	191 ± 5.0	902 ± 21	21.3

Beta Galactosidase values were average of 4 to 9 independent measurements. +3T and –3T denote Beta-Galactosidase values in the presence and absence of terminators. %RT is the read through efficiency at the end of the triple terminators and was calculated by taking the ratios of +3T and –3T values.

We observed from the *in vivo* assays that L108F N also partially suppressed the antitermination defect of β'-P251S and β'-G336S but had no effect on β'-R270C and β-G1045D. The antitermination efficiency of WT RNAP was about 2-fold reduced in presence of this mutant N. These observations could be related to the spatial location of the different RNAP mutations (as discussed later, Figure 3.18 B). β'-P254L is close to β'-P251S and β'-G336S in the structural model of EC (Figure 3.18), where as β-G1045D and β'-R270C form another cluster.

We purified the L108F N protein and used it to study *in vitro* antitermination with the WT and mutant RNAPs. Figure 3.18 shows the autoradiogram of *in vitro* antitermination assays for some representative RNAP mutants and Figure 3.19 A and B summarize the read through efficiencies (%RT) for all the mutants at the end of triple terminators and at the first terminator, respectively.

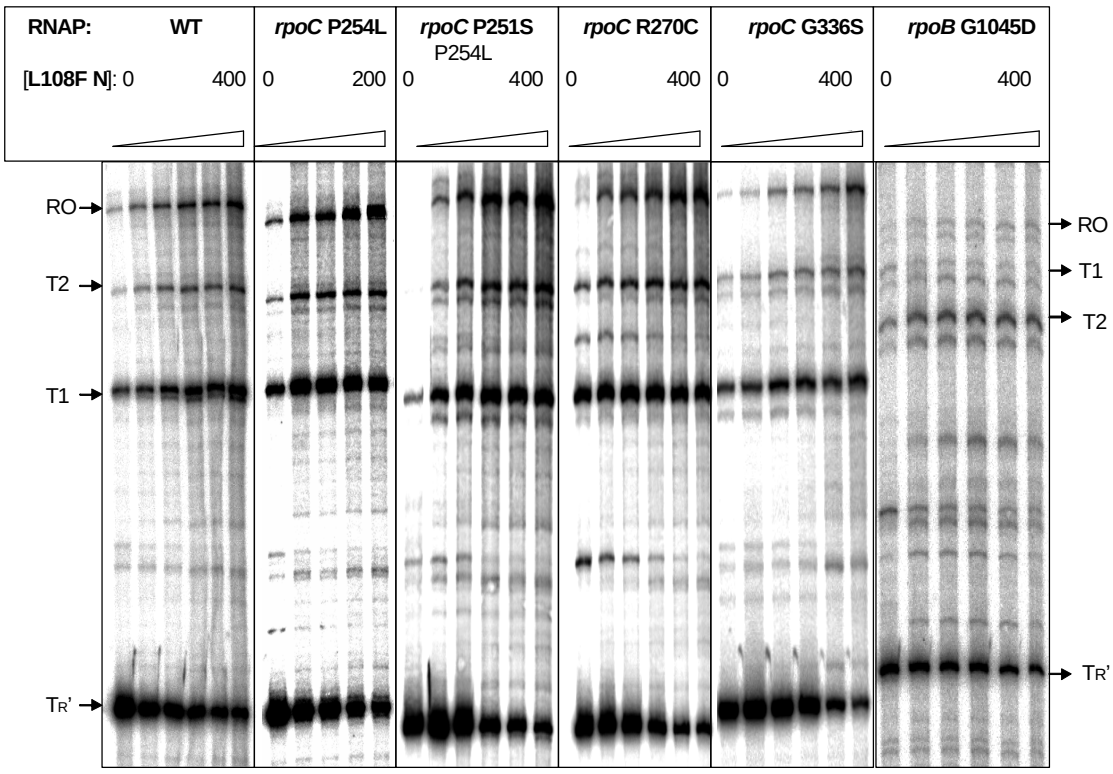


Figure 3.18: *In vitro* antitermination assays with L108F mutant of H-19B N

In vitro transcriptions were done under the same condition as described in Figure 3.14 and 3.15.

Autoradiogram showing the *in vitro* transcription of WT, β'-P254L, β'-P251S P254L double and β-G1045D mutants of RNAP in presence of increasing concentration of L108F mutant of H-19B N. Concentrations of N in different lanes were 0, 25, 50, 100 and 200 nM for P254L and 0, 25, 50, 100, 200 and 400 nM for other mutants.

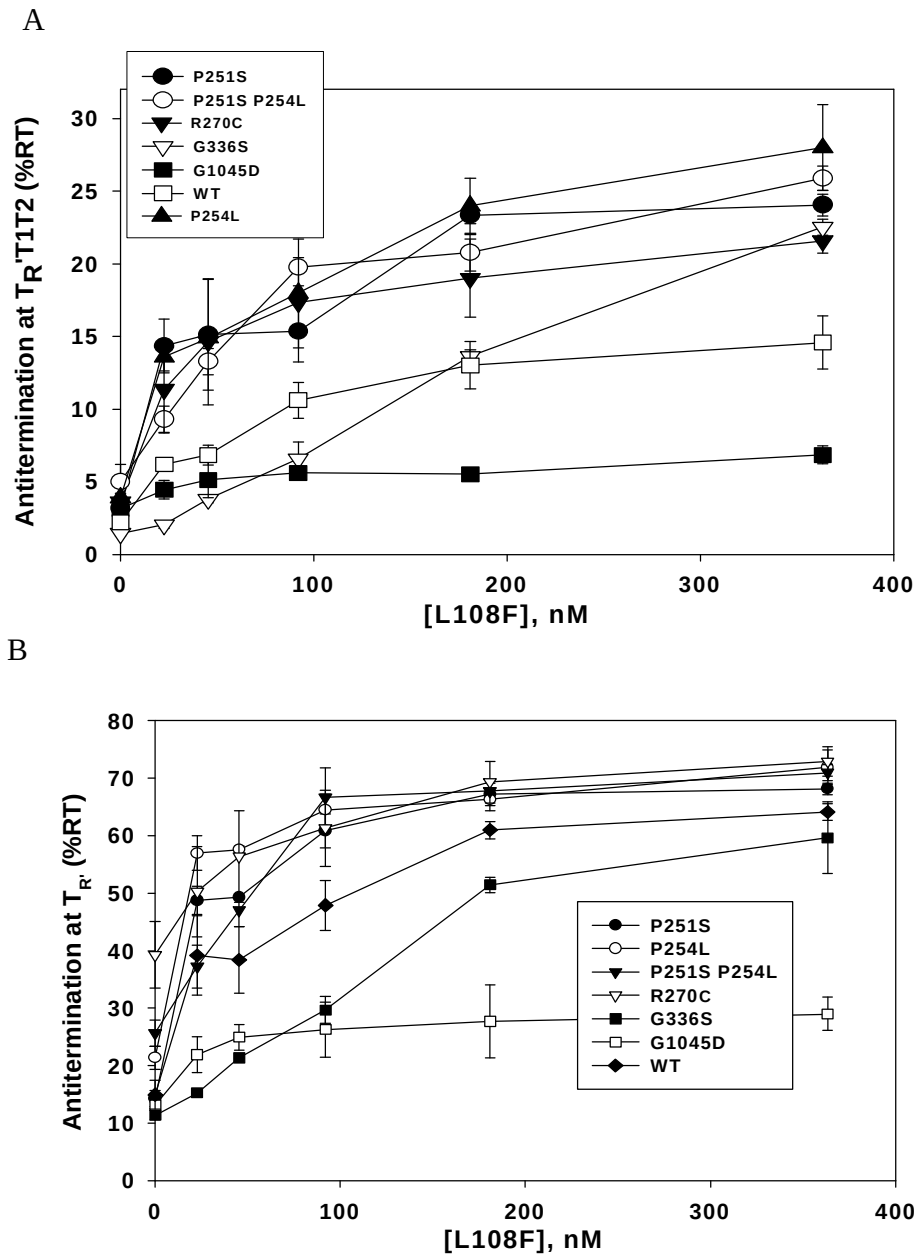


Figure 3.19: Plots representing the antitermination efficiency (%RT) of L108F N

A. Antitermination rates through the three terminators ($T_{R'}$- T1-T2) are plots of the antitermination efficiencies (%RT) of the three terminators for WT and all the mutants, against the increasing concentration of L108F N.

B. Antitermination rates through the first terminator ($T_{R'}$) for WT and all the mutants, against the increasing concentration of L108F N. Symbols of each plot are described in the inset. Antitermination efficiency was calculated as described in Figure 3.16.

The *in vitro* antitermination rates that we obtained complimented the *in vivo* suppression that we observed for L108F H-19B N. The antitermination defects of the *rpoC* P254L mutant was suppressed by L108 F N with antitermination read through rates nearing wild type levels (around 30%) being obtained in the *in vitro* triple terminator assay. The suppressor protein could also partially suppress the antitermination defects of the *rpoC* P251S P254L double mutant and also the *rpoC* P251S mutant *in vitro*. The L108F N protein could bring about only around 14% antitermination with wild type RNA polymerase; the CTD substitution in N clearly affecting its interactions with wild type RNA polymerase. The N suppressor protein had no discernible effect on the *rpoB* G1045D mutation. At higher concentrations, the L108F mutant N improves the read through efficiency of ß' mutants, P254L, P251S, P251S P254L double, and G336S, by at least two to three-fold at the end of triple terminators (compare with Figure 3.15). ß'-P254L, the RNAP mutation on which N-L108F was isolated, showed the highest efficiency. Read through efficiency for WT RNAP went down by three-fold (from 38% to 12%, compare with Figure 3.14), whereas it was unchanged for ß'-R270C and ß -G1045D. These results follow the same trends as observed *in vivo*. When antitermination is calculated at the end of the first terminator, the response to L108F for all the *rpoC* mutants, as well that as for Wt N was similar (Figure 3.19 B). L108F mutation in N did not make a more active H19-B N, but instead makes an N protein that works better with the mutant RNAPs than with WT. It suppressed the defect of the three *rpoC* mutations, which are located close to each other. This may indicate possible interactions between a region of ß' subunit (comprising of the residues ß' P251, ß'P254 and ß'G336), and of this part of N. However, as this suppression is not allele specific, we cannot rule out the possibility that effect of this mutation in N could be indirect.

3.3.7 Mapping the mutations on the model structure of EC

We localized the five RNAP mutations on the available model of EC, which has been developed by Seth Darst and co-workers, on the basis of crystal structures of *Thermus aquaticus* RNA polymerase (Zhang et al., 1999), elongation complex of yeast RNA polymerase II (Gnatt et al., 2001) and cross-linking data on protein-nucleic acids interactions from the EC of *E.coli* RNAP (Korzheva et al., 2000). We first determined the equivalent amino acids of these mutations in the model structure by structural alignment using CLUSTALX program (Thompson et al., 1997; refer Materials methods). All these changed amino acids, except ß'-P254, are highly conserved among the bacterial RNA Polymerases. Once the equivalent amino acids of these *E.coli* RNAP mutations were found in the structural model of TEC that we had (Figure 3.20), we localized these

substituted residues onto the TEC model. Figure 3.21 shows the location of the mutations on TEC.

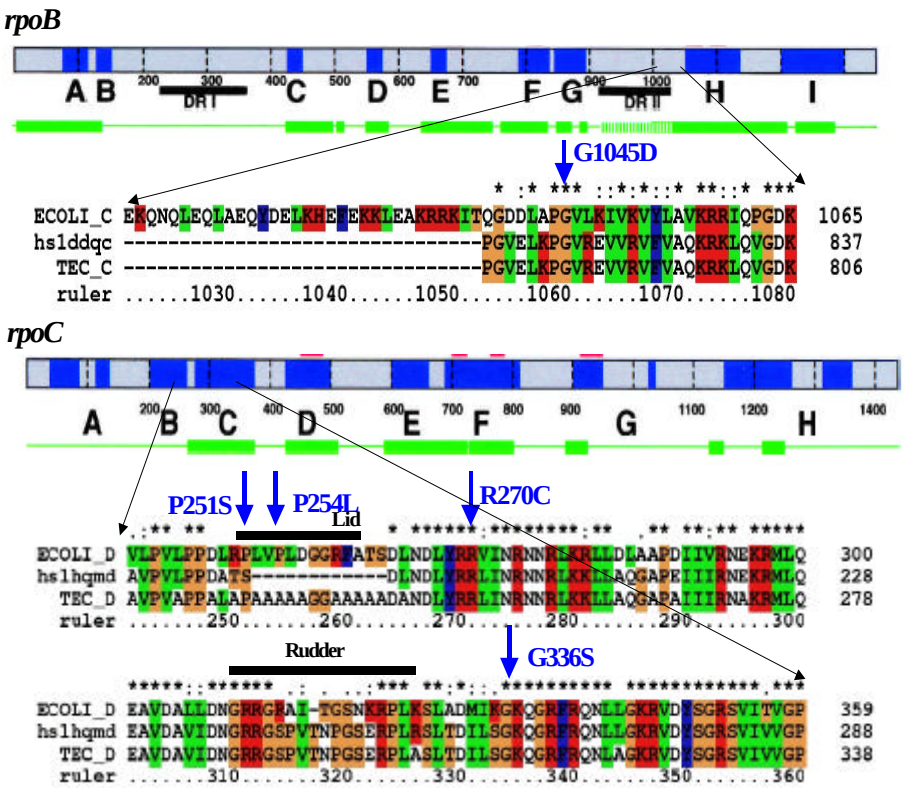


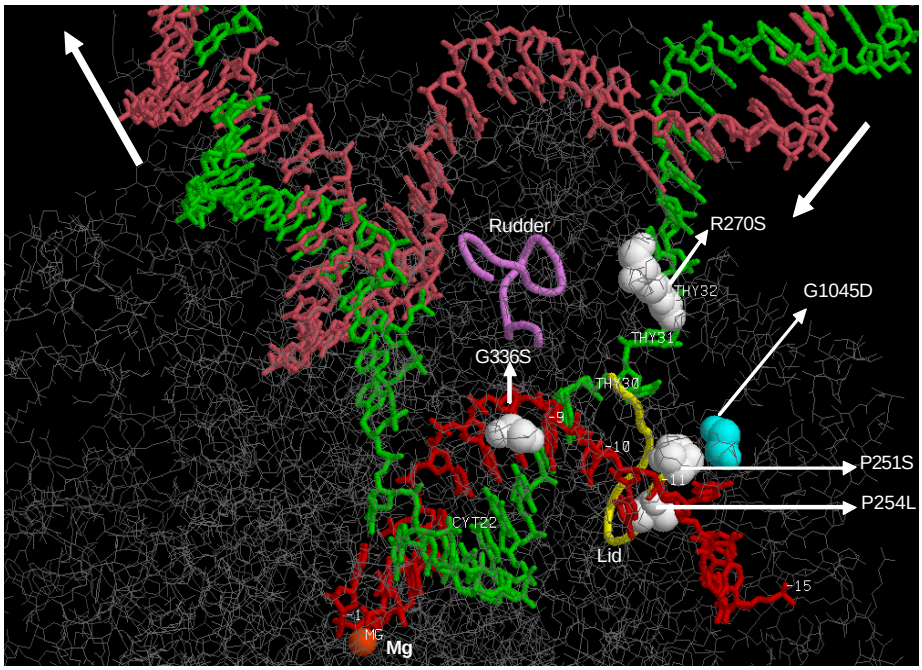
Figure 3.20: Localization of the substituted residues on to the *rpoB/C* genes

Structural alignment of the relevant amino acid regions of ß and ß' subunits of *E.coli* RNAP with the equivalent regions of *Thermus aquaticus* RNAP using CLUSTAL W is shown.
The bars represent maps of *rpoB* and *rpoC* with conserved segments in blue. The regions displaying essentially identical tertiary folds between *T. aquaticus* RNAP and yeast Pol II are shown beneath in green (Ebright 2000). Structural elements like ß' lid and ß' rudder are shown with underlining. DRI and DRII represent the two dispensable regions in *E.coli rpoB*. Color scheme is according to CLUSTALX color code. hs1ddqc and hs1hqmd stand for pdb ids of *Thermus acquaticus* RNAP C *(rpoB*) and D chains (*rpoC*), respectively, obtained from Homestrad database. TEC_C and TEC_D stand for ternary elongation complex C and D chains, respectively.

3.3.7.1 Localization of the *rpoB/C* substitutions in the model structure of TEC

We observed that in the model structure of EC, these mutations map very close to RNA:DNA hybrid, DNA template strand, RNA-exit channel and important structural elements like rudder and lid. We took a closer look at the location of these substituted residues with respect to their location in the EC and calculated the distance of these residues from structural elements like ß'

rudder and also their distance from RNA and DNA. Figure 3.21 gives an expanded view of the nucleic acid framework of the model structure of TEC using RASMOL and the molecular distances.



B

Mutation	Distance from RNA (Å)	Distance from DNA (Å)	Distance from rudder (Å)
rpoC			
P251S	6.2 (-10)	15.3 (31T)	19.2 (325K)
P254L	11.9(-10)	14.7 (30T)	23.8 (322R)
R270S	19.3 (-9)	15.0 (32T)	13.1 (314R)
G336S	9.3 (-8)	12.2 (22C)	10.3 (325K)
rpoB			
G1045D	20.4 (-11)	14.8 (31T)	24.3 (316I)

Figure 3.21: Localization of the substitutions and their spatial locations

A. An expanded view of the nucleic acid framework in the of the model structure of EC using RASMOL. Locations of different mutations to the equivalent positions of this model are shown as space-filled in the wire-framed gray background of other parts of RNAP. The color-coding of template DNA, non-template DNA and RNA are green, pink and red, respectively. Active site Mg^{2+} is shown in orange. Flexible loop structures of "rudder" and "lid" are also indicated.
B. A cartoon describing the numbering scheme of the RNA:DNA hybrid used in the model. RNA positions -1 to -8 are in RNA:DNA hybrid and positions -9 to -15 are in the RNA exit channel. Numbers of DNA residues are also mentioned and template and non-template DNA are labeled as "T" and "NT". The distances of each mutation from the nearest residues in the DNA, RNA or Rudder element were calculated (see experimental procedures) and are as given in the table. Nearest residue positions are shown in parenthesis. For an example, P251S is 6.2Å away from -10 position of nascent RNA, 15.3Å away from 31T of template DNA and 19.2Å away from the 325K amino acid in the rudder element.

The rudder and lid are the two loop-structures in the ß' subunit, and their role was originally proposed to guide nascent RNA into the RNA-exit channel (Gnatt et al., 2001). Recent deletion analysis of the ß' rudder and the ß' lid regions was shown to destabilize EC (Toulokhonov and Landick, 2006).After the determination of the spatial distances of the substituted residues from Nucleic acids and structural elements of the EC, it was found that ß'-P251S and ß'-P254L is located in the ß' lid structure and close to the -10 position of RNA in the RNA exit channel. ß'-G336S comes within 10Å of the upstream edge of RNA:DNA hybrid and the rudder element. ß'-R270C is also located close to rudder and template DNA upstream of the RNA:DNA hybrid, whereas ß -G1045D is situated a small distance away from these nucleic acid framework but comes within 12Å of P254L.

3.3.7.2 Surface accessibilities of the substituted *rpoB/C* residues

We then mapped these substituted residues on the space-filling model of the EC to know the surface accessibility of the mutant amino acids (Figure 3.22). ß -G1045D and ß'-R270C was found to be on the surface, ß'-P251S was partially exposed, while ß'-P254L and ß'-G336S were buried inside the structure. If these *amino acids* residues were to take part in direct interactions with N, a major conformational change in the EC will be required to get access to the ß'-P254L and ß'-G336S amino acids.

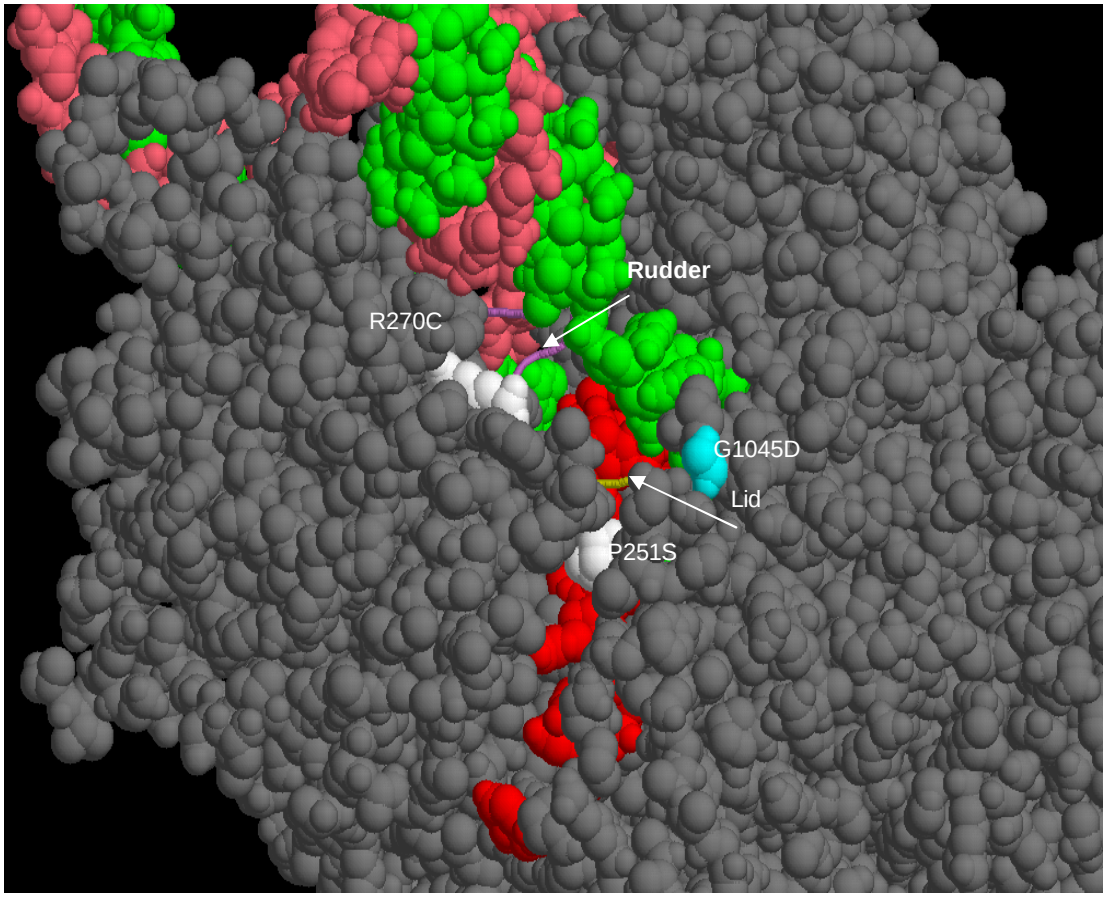


Figure 3.22: Space filled model of the EC

Rasmol space filled image, showing the EC with the nucleic acid framework and the substituted amino acid residues showing their surface accessibility. ß G1045D, ß' R270C and ß' P251S are visible on the surface while other mutants are buried inside the structure. Color-coding is same as in Figure 3.21.

Table 4: **Relative solvent accessibilities of the mutation site residues**

Mutation site	Total side chain	Relative total side chain accessible
ß G1045D	12.57	115.9
ß' P251S	17.65	44.9
ß' P254L	11.11	48.0
ß' R270C	28.70	45.8
ß' G336S	4.25	39.2

The relative solvent accessibilities of the mutation site residues were found using the PSA (JOY) program, which calculates the solvent accessibility of residues as the percentage of its accessibility in Ala-X-Ala dipeptide in extended conformation. By the definition of JOY software (Mizuguchi et al., 1998), a relative total side chain accessibility of less than 7.0 % is considered inaccessible. As could be inferred from Table 4, all the substituted residues that we obtained are accessible to solvent, among them *rpoB* G1045D showing the maximum accessibility and *rpoC* G336S being the least accessible. We mapped these mutations on the space-filling model of the EC to know the surface accessibility of the substituted amino acids (Figure 3.22). ß -G1045D and ß'-R270C were found to be on the surface, ß'-P251S is partially exposed, while ß'-P254L and ß'-G336S are buried inside the structure. If these residues were to take part in direct interactions with N, a major conformational change in the EC will be required to get access to the ß'-P254L and ß'-G336S amino acids. In this space filled image of the TEC, the ß G1045D mutant (cyan colored residue) as well as ß 'R270C and ß 'P251S (white residues) can be seen. The red colored residue is RNA in the exit channel; green and pink colored residues are Template and Non-template DNA respectively.

3.4 Discussion

RNA Polymerase mutations in the ß subunit (Georgopoulos, 1971; Ghysen and Pironio, 1972; Jin et al., 1988a; Sternberg, 1976; Jin et al., 1988; Jin et al., 1988b) and the a subunits of RNA Polymerase (Obuchowski et al., 1997; Schauer et al., 1996; Szalewska-Palasz et al., 2003; Weilbaecher et al., 1994), defective in supporting the growth of ? phage and as well as impaired for N mediated transcription antitermination has been reported earlier but none of these mutants were shown to be specifically impaired for binding of N. This is the first report describing the isolation of RNAP mutants specifically defective for transcription antitermination and for binding to the N antiterminator protein.

We provide the following lines of evidence which strongly suggest that these amino acids of ß and ß' subunits of RNAP are involved in direct binding of H-19B N.

1) Mutants exhibited significantly reduced *in vivo* antitermination efficiency and did not support H-19B phage growth (Tables 1, 2).

2) *In vitro* transcription studies showed increased concentration dependence of the RNAP mutants for N and exhibited lack of processive antitermination (Figure 3.15 and 3.16).

3) A mutation, L108F in the RNAP-binding domain of N (Mogridge et al., 1998), efficiently suppressed the defect of β'-P254L mutation (Table 3, Figure 3.19).

4) The mutants (as demonstrated in the case of β-G1045D) exhibited a failure to prevent RNA release from, and to stabilize the stalled elongation complex at a terminator (discussed later in Chapter V). These are strong indications for the binding defect of each of the RNAP mutants for H-19B N.

5) Close proximity of the substituted residues among each other and their location surrounding the RNA:DNA hybrid in the model structure of EC (Figure 3.21), clearly defines a binding surface for H-19BN.

Our results did not show unequivocally that these positions of amino acids in RNAP are involved in direct interaction with H-19B N and it is possible that effect of these mutations are indirect. N may bind to a different part of RNAP and the effect might be exerted allosterically through this region. Due to the spatial location of the mutations close to the RNA:DNA hybrid and upstream part of the RNA exit channel, these altered amino acid residues can affect the interactions of RNAP with RNA:DNA hybrid and with RNA in the exit channel, which in turn can impair the process of antitermination. Alternatively, it is also possible that these changes affect NusA and NusG functions in the antitermination complex. These mutant RNAPs were severely defective for phage H-19B growth but were either partially (for ß'-mutants) or not (for ß G1045D) defective for phage λ growth. Amino acid sequence of H-19B N is closely related to HK97 N rather than that to λ N (Weisberg and Gottesman 1999). Amino acid sequence of H-19B N is closely related to HK97 N and other *Stx* encoding phages than that to λ N (Figure 1.30). In general, RNAP binding domain of different N proteins are not well conserved (Weisberg and Gottesman, 1999) which suggests that their site of action may not be same but it is also unlikely that different N will have different mechanisms of antitermination. Partial defect of ß'-mutants for the growth of λ phage may suggest that site of action of λ N is also near this region. However, until we obtain mutations in this same region specific for λ N, or show through biochemical means that the site of actions of λ N and H-19B N are close by, we will restrict ourselves to the conclusions that region defined by these mutants is the specific site of action of H-19B N.

CHAPTER IV

Biochemical Probing of Binding Surface of H-19B N on the Elongation complex

4.1 Introduction

The isolation of N specific mutations in RNAP close to the RNA exit channel and the RNA:DNA hybrid (as described in Chapter III) encouraged us to attempt the identification of H-19B N interaction site on RNAP, through biochemical means. The RNAP mutations obtained were located very close to important structural elements like the RNA exit channel, the lid and the rudder elements of the EC (Cheeran et al., 2005). These elements are important for guiding the nascent RNA out of the RNA:DNA hybrid present at the active center (Gnatt *et al*., 2001; Naryshkina *et al*., 2006; Toulokhonov and Landick, 2006). Location of these mutations suggests that H-19B N exerts its action by modulating the interactions at the active center of the EC. The genetic screen that we employed had a draw back that only viable mutants could be isolated. Since the N interaction sites could be around the structurally important elements in and around the active site, and such mutants being mostly non-viable, we might have missed them in our genetic screen. To address this issue, we developed a highly specific *in vitro* set up for the footprinting of an N modified TEC using an array of biochemical tools.

We designed a stalled EC modified with NusA, NusG and H-19B N (both WT and mutant), using Lac repressor as a roadblock (as depicted in figure 4.1). The *lac* operator sequence is placed about 80 nt downstream of the *nutR* site to facilitate the *nutR*-hairpin dependent binding of H-19B N to the stalled EC. This stalled EC remains is found to stay transcriptionally active over the period of *in vitro* assays (King *et al*. 2003, Cheeran *et al.* 2005). In all the protein footprinting assays (other than the cases otherwise mentioned), this stalled EC was formed using radiolabeled RNAP. The radiolabeling of the protein was achieved through the HMK-tag sequences at the N-terminal of ß and C-terminal of ß' subunit of RNAP (refer Materials methods) for the detection of cleavage products generated from the Fe-BABE, Fe-DTT and Trypsin. Detailed analysis of the DNA-RNAP-*nut* RNA-N-NusA-NusG complex (to be referred as the N modified EC from now on in this chapter), were carried out by hydroxyl radical footprinting using Fe-EDTA (Price and Tullius, 1992; Heyduk et al., 1996; Tullius and Greenbaum, 2005; Wang et al., 1997), localized hydroxyl radical cleavages by an Fe^{2+} ion placed in the catalytic center of RNAP (Mustaev et al., 1997; Zaychikov et al., 1996) and footprinting by limited proteolysis with Trypsin (Borukhov et al., 1991). We also carried out footprinting of the EC using a chemical nuclease called Fe-BABE (Datwyler and Meares, 2000; Meares et al., 2003), linked to the C terminal domain of H-19B N. Specific cleavage patterns of the EC were generated using this techniques. We used hydroxyl radical footprinting to identify areas on the solvent-accessible surface of RNAP that interact with N protein, the second method (localized footprinting with Fe-DTT) to detect local N-TEC

interactions within 5-20 Å from the Mg^{2+} binding site and the third method to look for conformational changes brought in by N modification which renders residues on the RNAP more accessible for/sensitive towards Trypsin digestion. Footprinting with H-19B N conjugated with Fe-BABE was used to directly map the regions which comes within ~12 Å radius of the Fe-BABE conjugated at the C terminus of N (refer Materials methods). With all these methods, N-TEC interactions were inferred by changes in the protection/enhancement patterns of the various cleavages, which occur as either a direct or indirect consequence of H-19B N protein binding.

4.2 Materials and methods

4.2.1 Cloning and purification of the HMK tagged proteins

An HMK (Heart Muscle Kinase) tag sequence was cloned to the C terminus or N terminus of various proteins, where by these proteins can be radiolabelled and subsequently the labeled proteins or the clevage products from these proteins can be visualized and analyzed in a phosphor imager screen, after electrophoresis. The labeling method utilizes the catalytic subunit of PKA to catalyze the transfer of the gamma phosphate of **?**-P^{32} ATP to the Serine residue of the PKA recognition site (Arg Arg Ala **Ser** Val) present on fusion proteins

4.2.1.1 Cloning, expression and purification of C terminal HMK-*rpoC* protein

pBAD18M (Sen et al., 2002) encoding Wt *rpoC* was digested with restriction enzymes HindIII and SacI. RpoC gene was amplified using RS63 (upstream of *rpoC* gene) and RS228 (with C-terminal HMK tag, 6-His tag and SacI of *rpoC*) as the downstream primer using a proof reading DNA polymerase enzyme, Deepvent (NEB). The PCR product was double digested with HindIII/SacI and was cloned into pBAD18 at its HindIII/SacI site. The plasmid (pRS513d) was sequenced and was transformed into RS38e (*rpoC ts* background strain), grown at 42 ℃, checked on LB A, S, T, XGal, IPTG plates for complementation and *in vivo* antitermination. This strain (RS514d) was used for large-scale cell pellet preparation at 42 °C in TB supplemented with Amp, Spec, and Tet media. C-terminal HMK tagged RNAP was purified according to the published protocol (Kashlev et al 1996) and the same way as given earlier in Chapter III using Ni–NTA beads (Quiagen) and Heparin-agarose affinity column (Amersham).

4.2.1.2 Expression and purification of N terminal-HMK *rpoB* protein

pTRC99A encoding ß S531Y *rpoB* was a gift from Sergei Borukhov (Laptenko et al., 2003).This plasmid was transformed into MG1655 cells (RS487) grown large scale in TB and ß S531Y RNA Polymerase was purified according to the protocol (Kashlev et al., 1996) and the same way as given earlier in Chapter III using Ni–NTA beads (Quiagen) and Heparin-agarose affinity column (Amersham)

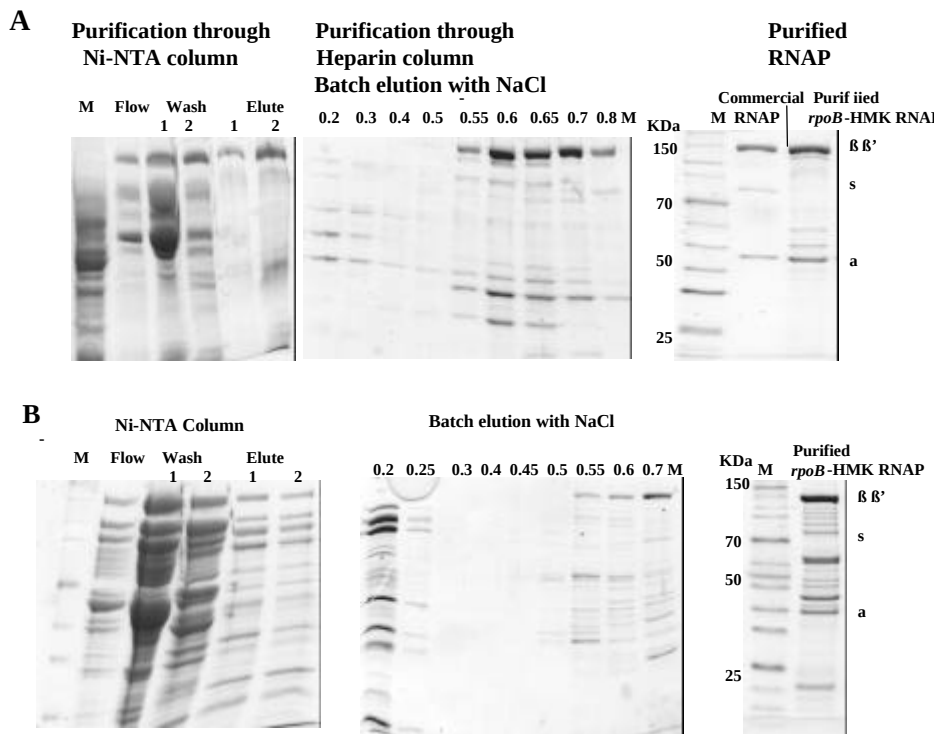


Figure 4.1: Purification of HMK tagged RNA polymerase proteins

A.**HMK tagged *rpoB* RNAP purification.** Purification steps through Ni-NTA column (the flow through fraction through Ni-NTA column, two wash fractions with 20 mM Imidazole, two fractions of RNAP eluted with 250 mM Imidazole); the NaCl gradient batch elution fractions through the Heparin column and the purified protein are shown. M stands for protein molecular weight marker (NEB). Commercial RNAP (Holoenzyme, Epicenter) is also shown, for comparison.

B. HMK tagged *rpoC* RNAP purification. Purification steps through Ni-NTA column (the flow through fraction through Ni-NTA column, two wash fractions with 20 mM Imidazole, two fractions of RNAP eluted with 250 mM Imidazole); the NaCl gradient batch elution fractions through the Heparin column and the purified protein are shown. M stands for protein molecular weight marker (NEB). M stands for protein molecular weight marker (broad range protein marker; MBI).

4.2.1.3 Cloning, expression and purification of N terminal HMK tagged H-19B N

H-19B N gene was amplified from pK8601 (Neely and Friedman, 1998a) using RS66a (upstream primer with NdeI site) and RS65 (Downstream primer of H19B N gene with XhoI site and with stop codon) primers using Dnazyme (Finnzyme); double digested with XhoI/NdeI restriction enzymes and was cloned into the XhoI/NdeI site of pET33b (+) (Merck). The sequenced plasmid (pRS464) was transformed into *E. coli* BL21 strain; the protein was over expressed and purified according to the published procedure for non-Histagged ?N purification (Das et al., 1996) and as given in the materials methods section of Chapter III for non Histagged H-19B N protein purification.

4.2.1.4 Cloning, expression and purification of N terminal HMK NusA

NusA gene was obtained by double digestion of the pET28B-NusA (gift from Irina Artsimovitch) with XhoI/NdeI enzymes and was subsequently cloned into the XhoI/NdeI sites of pET33b (+) vector and transformed into DH5a cells. The transformants were screened for the insertion of the gene, plasmids were isolated from the positive clones and further checked with restriction digestion for insert release. The resulting positive plasmid was sequenced, named (pRS523) and was transformed into BL21 cells. It was then over expressed and NusA protein was purified using Ni-NTA column (Amersham), as described earlier in Chapter III for wild type NusA purification.

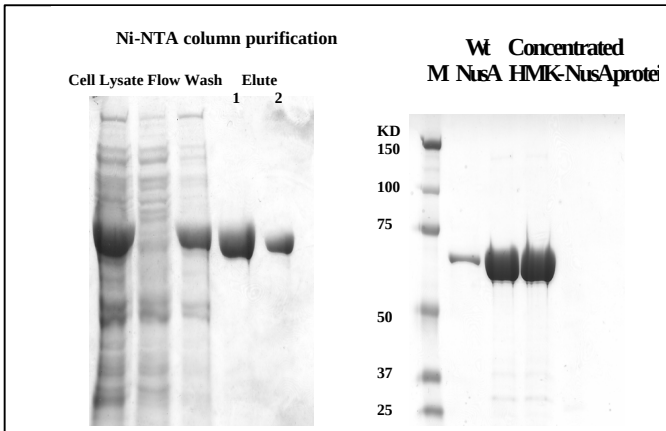


Figure 4.2: HMK tagged NusA purification

Purification steps, including passage through Ni-NTA column (the lysed cells before passage through column, the column flow, wash and elute fraction); and the purified protein is shown. Wt NusA protein stands for purified NusA protein, shown for comparison. M stands for protein molecular weight marker (broad range protein marker; MBI).

4.2.2 Cloning and purification of C107A A116C H-19B N

The mutant H-19B N gene was amplified by overlap PCR from pK8601, using primers RSS66a (upstream primer with NdeI site) and RS282, RS284 (downstream primers with the C107A, A116C changes incorporated, RS284 also had the stop codon and NdeI site). The amplified gene was double digested with NdeI/XhoI enzymes and cloned into the NdeI/XhoI site of pET21b and was transformed into DH5a cells, the transformants were screened for the insertion of the gene

and plasmids were isolated from the positive clones and further checked with restriction digestion for insert release. The resulting positive plasmid was sequenced, named pRS 598 and was transformed into BL21 cells. It was then over expressed and purified under the same conditions as given for that of the purification of non-His tagged H-19B N purification in the materials methods section of Chapter III.

4.2.3 Cloning, expression and purification of ? CTD H-19B N

C-terminal deletion mutant of H-19B N was made by amplifying Wt H-19B N gene from pK8601 (Neely and Friedman, 1998) using RS66a (upstream primer with NdeI site) and RS 202 (downstream primer for making del-CTD from 100th amino acids of H-19B N, having XhoI site and stop codon) using Deepvent enzyme (NEB), the PCR product was double digested with XhoI/NdeI and cloned into the XhoI/NdeI sites of pET21a. The resulting plasmid (pRS422a) was sequenced, transformed into BL21 strain and the CTD H-19B N protein was over expressed and purified according to the published procedure for ? N and as given for the purification of non Histag H-19B N purification in chapter III

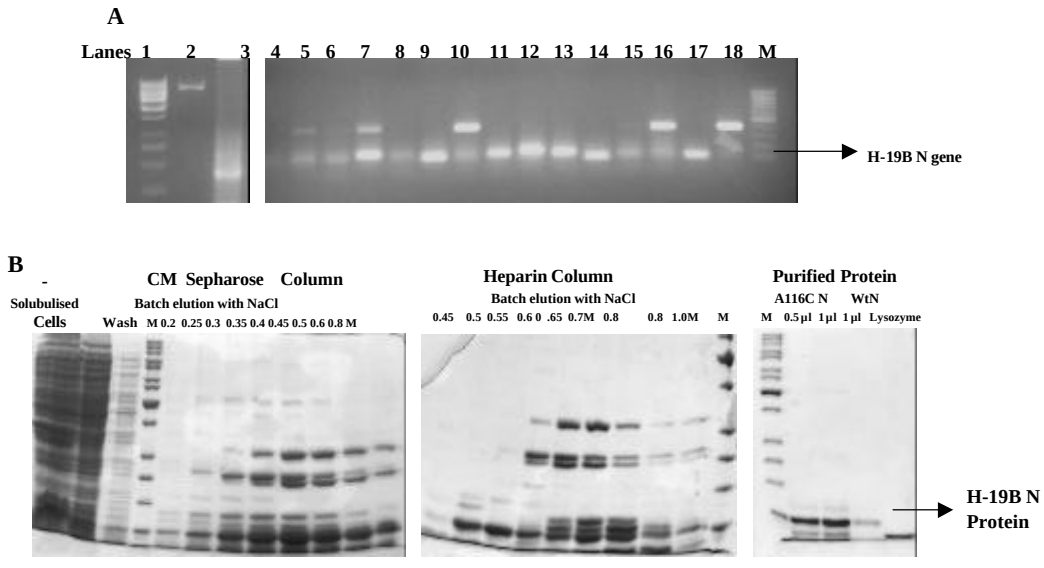


Figure 4.3: Cloning and purification of C107A A116C H-19B N protein
A. Lanes 2 and 3.Double digested pET21b and C107A A116C H-19B N gene. Lanes 7,10,16, 18.Clones that gave the correct product size in the colony PCR for screening the recombinants. **B. Purification of C107A A116C H-19B N protein.** Different purification steps including passage through CM Sepharose and batch elution of H-19B N protein with NaCl; batch elution through Heparin column with NaCl; and the purified protein, is shown. M stands for protein molecular weight marker (broad range protein marker; MBI).

4.2.4 ?-P^{32}ATP labeling of HMK tagged proteins

HMK tagged RNAPs, H-19B N and NusA proteins were radiolabelled with **?**-P^{32}ATP (Amersham) using Protein Kinase A enzyme (Sigma) resuspended in MES buffer (I M MES dissolved in DEPC water, pH adjusted to 6.5 and filter sterilized. The labeling reaction was done essentially as given in the protocol for **?**-P^{32} ATP labeling of HMK tagged proteins in the PKAce kit from Novagen, which helped in the efficient, high specific activity isotopic labeling of proteins bearing the five amino acid target sequence for cAMP-dependent protein kinase (Protein Kinase A; PKA). The labeling method utilizes the catalytic subunit of PKA to catalyze the transfer of the gamma phosphate of **?**-P^{32} ATP to the serine residue of the PKA recognition site (Arg Arg Ala **Ser** Val) present on fusion proteins.

4.2.5 DNA Templates for transcription reactions and footprinting

In order to immobilize the template DNA on the Streptavidin coated magnetic beads (Promega), a biotinylated primer (RS83) was used as the upstream primer and the *lac* operator sequence was incorporated in the downstream primer (RS99). The template DNA was amplified with this primers using pRS25 (Cheeran *et al*., 2005) as DNA template. The PCR amplified DNA was phenol extracted and was further purified by concentrating in a Cenricon-100 (Millipore).

4.2.5.1 Cloning of the Lac Operator downstream to H-19B *nut* site

RS70 oligo (*lac* operator palindrome flanked by XbaI site) was double digested with XbaI restriction enzyme and was cloned into the XbaI cut, CIP treated pRS25 (Cheeran *et al*., 2005), and the positive recombinant was sequenced and named pRS385.

4.2.6 Footprinting of the TEC

4.2.6.1 Generation of radiolabelled RNAP

The radiolabeling was achieved through the HMK-tag sequences at the N-terminal of ß and C-terminal of ß' subunit of RNAP for the detection of cleavage products generated from the Fe-BABE, Fe-DTT and Trypsin mediated cleavage reactions. The labeling reaction was done as earlier described (50 U of PKA enzyme and 5 µM of RNAP were used per reaction), the ß S531Y RNAP getting labeled at the N terminus and the ß'-C terminal HMK tagged RNAP being labeled at the C terminus.

4.2.6.2 Hydroxyl radical footprinting

Fe-EDTA Hydroxyl radical footprinting was carried out on RB TEC in the modified with either Wt H-19B N or a CTD (? 100-127 a. a of H-19B N) H-19B N. Hydroxyl radical footprinting was developed in mid-1980s and has proven to be of great utility in studies involving DNA-Protein or Protein-Protein interactions (Tullius and Greenbaum, 2005). The hydroxyl radical is a small, diffusible, highly reactive, neutral species. The hydroxyl radical is developed *in situ*, through a variant of the Fenton's reaction. The [Fe (II)(EDTA)] $^{2-}$ molecule is allowed to react with Hydrogen Peroxide, producing the hydroxyl radical, hydroxide ion and [Fe (III)(EDTA) $^{1-}$.

$[Fe(EDTA)]^{2-} + H_2O_2 \longrightarrow + OH^- + OH^\cdot$ Ascorbate is also added to the reaction mixture to reduce the resultant Fe (III) to Fe (II), making the system catalytic in iron and for decreasing the amount of [Fe(EDTA)] ^{2} needed.

4.2.6.2.1 Fe-EDTA footprinting protocol

The initial mix containing 1 µM ?-P^{32} ATP labeled RNAP (either *rpoC* or *rpoB* labeled), 100 nM template DNA, 1x TCl buffer (transcription buffer with 50 mM TrisCl pH 8.0, 100 mM KCl, 10 mM $MgCl_2$, 1 mM DTT and 1 mg/l BSA) was mixed with pre-equilibrated streptavidin magnetic beads (Promega) and incubated for 10 minutes at 30 °C. It was washed 1x to remove the excess RNAP and DNA and was resuspended in 1x TCl (Transcription buffer). Lac Repressor protein was added to a concentration of about 100 nM to this and was incubated for 10 minutes. This mix was aliquoted to fresh tubes and the different chase solutions having 500 µM each of all the NTPs with Transcription buffer, 0.1 mg/ml BSA, 1 mM DTT and proteins (600 nM NusA, 400 nM NusG, 1 µM H-19B N, ARM H-19B N, CTD H-19B N) as per requirement of each reaction was added (for forming EC with no elongation factors, EC with NusA and NusG, EC with Wt H-19BN, EC with ARM H-19BN and EC with CTD H-19BN) and transcription was carried out at 30 °C for 3 minutes. Each tube was washed 3x with 100 µl of 1x Transcription buffer, resuspended in 7.1 µl of the same buffer, without BSA added. (An Fe-EDTA mix was made earlier by adding 250 mM Ferrous Ammonium Sulphate and 500 mM EDTA and incubating at room temperature for 20 minutes). To the walls of the tubes, 0.3 µl 600 mM H_2O_2 (final concentration of 20 mM) and 1 µl of 200 mM Ascorbate (final concentration of 20 mM) was added. 1.6 µl of the earlier made 125 mM Fe-EDTA mix was added to the tubes (to a final concentration of 20 mM) and the other two drops of H_2O_2 and Ascorbate were pushed in and mixed well. This was incubated at room temperature and the reaction was stopped at required time points (30 s, 3 minute) by taking out 5µl each and mixing with 5x SDS load dye; and the tubes were stored on ice. Samples were heated at 95 °C for 4 minutes and were loaded on to an

8% SDS-PAGE. The gel was exposed overnight to a Phosphor imager film (Amersham) and the bands were analyzed using Phosphor imager (Typhoon 9200) and Imagequant software (Amersham.).

4.2.6.3 Footprinting of TEC by limited proteolysis with Trypsin

We probed the N-induced conformational changes of the elongation complex by following the partial Trypsin digestion pattern. RB TECs without N, with wild type N and with CTD N were made as above and were subjected to partial trypsin digestion (Borukhov et al., 1991). Limited digestion with Trypsin cuts the protein at Lysine and Arginine a. a residues.

4.2.6.3.1 Reaction protocol

The initial mix containing 1 µM ?-P^{32}ATP labeled RNAP (either *rpoC* or *rpoB* labeled), 100 nM template DNA, 1x TCl buffer (transcription buffer with 50 mM Tris-Cl pH 8.0, 100 mM KCl, 10 mM $MgCl_2$, 1 mM DTT and 1 mg/l BSA) was mixed with pre-equilibrated Streptavidin magnetic beads (Promega) and incubated for 10 minutes at 30 °C. It was washed 1x to remove the excess RNAP and DNA and was resuspended in 1xTCl (Transcription buffer). Lac Repressor protein was added to a concentration of about 100 nM to this and was incubated for 10 minutes. This mix was aliquoted to fresh tubes and the different chase solutions having 500 µM each of all the NTPs with Transcription buffer, 0.1mg/ml BSA, 1mM DTT and proteins (600 nM NusA, 400 nM NusG, 1 µM H-19B N, ARM H-19B N, CTD H-19B N) as per requirement of each reaction (for forming EC with no elongation factors, EC with NusA and NusG, EC with Wt H-19BN, EC with ARM H-19BN and EC with CTD H-19BN) were added; transcription was carried out at 30 °C for 3 minutes. Each tube was washed 3x with 100 µl of Transcription buffer and finally resuspended in 8 µl of the same buffer. To each of these tubes 2 µl of 0.3 U (1 µg/ml) of Trypsin was added, incubated at 30 °C and the reactions were stopped at 30 seconds and at 2 minutes respectively by taking out 5 µl each and adding equal amounts of 5x SDS loading dye, the tubes were then stored on ice. Samples were heated at 95 °C for 4 minutes and were loaded on to an 8% SDS-PAGE. The gel was exposed overnight to a Phosphor imager film (Amersham) and the bands were analyzed using Phosphor imager (Typhoon 9200) and Imagequant software (Amersham).

4.2.6.4 Fe-DTT footprinting

Fe^{2+} mediated localized cleavage of beta and beta' subunits of elongation complex was carried out to probe the conformational changes (if any) and the active site dynamics that could be

brought about by the binding of N to TEC. Active site Mg (II) can be substituted with Fe (II) and *in situ* hydroxyl radicals can be generated in presence of DTT. These activated oxygen species are extremely reactive and can cause degradation of biopolymers within ~1 nm range (Mustaev et.al., 1997). Using this principal, we observed the cleavage patterns of the regions of ß and ß' subunit that are surrounding the active center. The changes observed in the cleavage pattern (between N modified and unmodified EC) should be an indication of the conformational changes in the active center, brought about by N binding. We have monitored the cleavage pattern of the elongation complex in the presence and absence of WT or mutant N. All the other conditions were same as before in the case of Fe-BABE mediated footprinting, except that after the RB complex was formed, active site Mg (II) was replaced by Fe (II) by repeated washing of the complex with Magnesium less transcription buffer and by subsequently adding Fe (II). The Fe-DTT clevage of the EC was carried out as follows, largely following the published protocols (Laptenko et al., 2003; Mustaev et al., 1997).

4.2.6.4 1 Fe-DTT cleavage protocol

The initial mix containing 1 µM ?-P^{32}ATP labeled RNAP (either *rpoC* or *rpoB* labeled), 100 nM template DNA, 1x TCl buffer (transcription buffer with 50 mM Tris-Cl pH 8.0, 100 mM KCl, 10 mM $MgCl_2$, 1 mM DTT and 1 mg/l BSA) was mixed with pre-equilibrated Streptavidin magnetic beads (Promega) and incubated for 10 minutes at 30 °C. It was washed 1x to remove the excess RNAP and DNA and was resuspended in 1x TCl (Transcription buffer). Lac Repressor protein was added to a concentration of about 100 nM to this and was incubated for 10 minutes. This mix was aliquoted to fresh tubes and the different chase solutions having 500 µM each of all the NTPs (Amersham) with Transcription buffer, 0.1 mg/ml BSA, 1 mM DTT and proteins (600 nM NusA, 400 nM NusG, 1 µM H-19B N, ARM H-19B N, CTD H-19B N) as per requirement of each reaction (for forming EC with no elongation factors, EC with NusA and NusG, EC with Wt H-19BN, EC with ARM H-19BN and EC with CTD H-19BN) was added and transcription was carried out at 30 °C for 3 minutes. Each tube was washed 2x with 200 µl of Transcription buffer without Mg and was resuspended in 8.5 µl of the same buffer. Into this, 0.5 µl 10 mM DTT (final concentration of 5 mM) and 1 µl 0.4 mM of freshly made Ferrous Ammonium Sulphate (Sigma) (final concentration of 40 µM) was added, kept for the required time period at room temperature (2 minutes and 10 minutes) and the reaction was stopped by the addition of 5x SDS load dye and the tubes were stored on ice. Samples were heated at 95 °C for 4 minutes and were loaded on to an 8% SDS-PAGE. The gel was exposed overnight to a Phosphor imager film (Amersham) and

the bands were analyzed using Phosphor imager (Typhoon 9200) and Imagequant software (Amersham).

4.2.6.5 Footprinting of the EC with Fe-BABE

Fe-BABE (p-Bromoacetamidobenzyl-EDTA), is a unique tool for determining the three dimensional structures of proteins as well as the binding surfaces of proteins on protein-DNA and protein-protein complexes. Fe-BABE is a Cysteine-tethered chemical nuclease, developed in Claude Meares laboratory (Meares et al., 2003). Fe-BABE can be specifically conjugated to Cysteine a. a residues of a protein. Fe-BABE (Fe Bromoacetamidobenzyl-EDTA) is a protein modification reagent (Meare's reagent) and can be attached to a protein through sulfhydryl groups. Under reducing conditions (the reaction is initiated by addition of Ascorbic acid), whereby an Iron (II) chelate is generated by the reduction of an Iron (III) chelate with Ascorbic acid), and upon addition of Hydrogen peroxide, this reagent generates hydroxyl radicals, which cleaves nearby (~12 Å) peptide and nucleic acid backbones. Here, it was expected that the Hydroxyl radical generated *in situ* will cleave the peptide backbones of the different domains of RNAP situated =12 Å radius of the C-terminal Cysteine residue of the H-19B N bound to the EC and thereby will define the regions of the EC coming close to the C-terminal of N. H-19B N has a single Cysteine residue near the C-terminal of the protein and Fe-BABE was conjugated to it. The sequence of H-19B N protein is as follows; the unique Cysteine residue is underlined

MTRRTQFKGNSRSRRRERLKAKALANGVLAREEAISSEVLHRPTLSRAQIQAKGTHETP
DRIEDAKPIKFMAQDVIWQQEEYRRNLERAAIVYANEFGHKQPETGV**C**LPNVALYAAG
YRKSKQLTAR. The presence of a unique Cysteine at the C terminus allowed us to use it for linking with Fe-BABE. The single Cysteine (a. a 107) at the C-terminal of H-19B N was conjugated to Fe-BABE.

4.2.6.5.1 Conjugation of H-19B N/ARM H-19BN protein with Fe-BABE reagent

The single Cysteine (a. a 107) at the C-terminal of H-19B N was conjugated to Fe-BABE. For effective conjugation, it is required to remove reducing agents such as DTT, if present, from the protein. In order to remove DTT from H-19B N stock solution (20 µM), 100 µl H-19B N protein was first equilibrated in exchange buffer (20 mM Tris-HCl, pH 8.0, 0.5M NaCl and 1 mM EDTA) and concentrated by using Amicon-10 membranes (Millipore). 0.5 mM of Fe-BABE (Dojindo, Japan) was incubated with 10 µM H-19B N in exchange buffer for 1 hr at 37 ^{o}C. Un-conjugated Fe-BABE was removed by passing the reaction mixture through Amicon 10-membranes, equilibrated in storage buffer (20 mM Tris-HCl (pH 8.0), 100 mM NaCl, 10%

glycerol and 1 mM EDTA). Cysteine modification was monitored by the sensitivity for DTNB (Riddles et al., 1979). The free –SH groups in Fe-BABE conjugated H-19B N was determined spectrophotometrically at 412 nm using DTNB. Fe-BABE conjugated H-19B N (1.6 µM) was incubated at 25 ^{o}C with 0.8 mM DTNB in 50 mM Phosphate buffer, pH 7.5. The stoichiometry of the reaction was calculated by using extinction coefficient of 14150 M^{-1} cm $^{-1}$. Fe-BABE conjugation of H-19B N was found to be 76.2%.

4.2.6.5.2 Fe-BABE conjugated H-19B N probing of EC with labeled RNAP

The initial mix containing 1 µM ?-P^{32}ATP labeled RNAP (either *rpoC* or *rpoB* labeled), 100 nM template DNA, 1x TCl buffer (transcription buffer with 50 mM Tris-Cl pH 8.0, 100 mM KCl, 10 mM $MgCl_2$, 1 mM DTT and 1 mg/ml BSA) was mixed with pre-equilibrated Streptavidin magnetic beads (Promega) and incubated for 10 minutes at 30 °C. It was washed 1x to remove the excess RNAP and DNA and was resuspended in 1x TCl (Transcription buffer). It was washed 1x to remove the excess RNAP and DNA and was resuspended in 1x Transcription buffer. Lac Repressor protein was added to a concentration of about 100 nM to this and was incubated for 10 minutes. This mix was aliquoted to fresh tubes and the different chase solutions having 500 µM each of all the NTPs with 1x Transcription buffer, 0.1 mg/ml BSA and proteins (600 nM NusA, 400 nM NusG, 1 µM Fe-BABE conjugated H-19B N, Fe-BABE conjugated ARM H-19B N) as per requirement of each reaction were added and transcription was carried out at 30 °C for 3 minutes. Three separate sets of EC s having un-conjugated Wt H-19BN, Fe-BABE conjugated Wt H-19BN and Fe-BABE conjugated ARM N was made. Each tube was washed 3 x with 100 µl of 1x Transcription buffer and finally resuspended in 8 µl of the same buffer. Cleavage reaction was initiated by adding 25 mM H_2O_2 and 100 mM ascorbic acid, and allowed to continue for different times at 32 ^{o}C. The reactions were stopped by adding 6 µl 5x SDS loading dye supplemented with 0.1 mM EDTA and were stored on ice. Samples were heated to 95 ^{o}C for 4 min prior to loading onto an 8% SDS-PAGE. Gels were exposed overnight to a phosphor-imager screen and the bands were analyzed using Phosphor imager (Typhoon 9200) and Imagequant software (Amersham).

4.2.6.5.3 Fe-BABE probing of RNA on EC

The initial mix containing 1 µM Wt RNAP, 100 nM template DNA, 1x TCl buffer (transcription buffer with 50 mM Tris-Cl pH 8.0, 100 mM KCl, 10 mM $MgCl_2$, 1 mM DTT and 1 mg/l BSA) was mixed with pre-equilibrated Streptavidin magnetic beads (Promega) and incubated for 10 minutes at 30 °C. It was washed 1x to remove the excess RNAP and DNA and was resuspended

in 1x TCl (Transcription buffer). Lac Repressor protein was added to a concentration of about 100 nM to this and was incubated for 10 minutes. Transcription reaction was initiated with 175 µM ApU, 5 µM GTP, 5 µM ATP, and 2.5 mM CTP to make a 22-mer EC. This EC was labeled with a-P^{32}-CTP. This mix was aliquoted to fresh tubes and the different chase solutions having 500 µM each of all the NTPs with 1x TCl, 0.1 mg/ml BSA and proteins (600 nM NusA, 400 nM NusG, 1 µM Fe-BABE conjugated H-19B N, Fe-BABE conjugated ARM H-19B N) as per requirement of each reaction was added and transcription was carried out at 30 °C for 3 minutes. Under these conditions, the reactions were essentially single-round, as the labeled CTP was diluted more than 1000-fold by the cold CTP. Three separate sets of EC having un-conjugated Wt H-19B N, Fe-BABE conjugated Wt H-19B N and Fe-BABE conjugated ARM N was made. Each tube was washed 3x with 100 µl of Transcription buffer and finally resuspended in 8 µl of the same buffer. Fe-BABE cleavage reaction was initiated by adding 25 mM H_2O_2 and 100 mM ascorbic acid to this tube, mixed well and the reaction was incubated for 1 minute, 2 minute and 5 minutes respectively. Reactions were stopped by extracting with phenol, followed by precipitation by ethanol in the presence of Glycogen and Ammonium Acetate, mixed with loading dye and run on 8% sequencing gel and exposed overnight to a phosphor imager screen. Transcript intensities were quantified using Typhoon 9200 (Amersham).

4.2.6.5.4 Fe-BABE probing of DNA on EC

The template DNA was labeled with **?**-P^{32}ATP by end labeling the reverse primer (RS99) with Poly Nucleotide Kinase enzyme (NEB, USA) and then using this primer along with a Biotinylated upstream primer (RS83) and pRS25 plasmid as the template a PCR reaction was carried out to amplify the template DNA. The PCR amplified product was phenol extracted and was purified by concentrating on a Centricon-100 (Millipore). The non-template strand of DNA was labeled by end labeling forward primer RS58 and PCR amplification (using RS2 as downstream primer) and purification as above. Clevage reaction was carried out as follows: 1 µM RNAP, 100 nM template DNA, 1x TCl (transcription buffer with 50 mM Tris-Cl, pH 8.0, 100 mM KCl, 10 mM $MgCl_2$), 1x DTT and 1x BSA was mixed with pre-equilibrated Streptavidin magnetic beads and incubated for 10 minutes at 30 °C. It was washed with transcription buffer to remove the excess RNAP and DNA and was resuspended in same buffer. Lac Repressor protein was added to a concentration of about 100 nM to this and was incubated for 10 minutes. This mix was aliquoted to fresh tubes and equal amounts of different chase solutions having 500 µM each of all the NTPs with Transcription buffer, 0.1mg/ml BSA and proteins (600 nM NusA, 400 nM NusG, 1 µM Fe-BABE conjugated H-19B N, Fe-BABE conjugated ARM H-19B N) as per

requirement of each reaction was added and transcription was carried out at 30 °C for 3 minutes. Three separate sets of EC having un-conjugated Wt H-19BN, Fe-BABE conjugated Wt H-19BN and Fe-BABE conjugated ARM N was made. Each tube was washed 3x with 100 µl of Transcription buffer and finally resuspended in 8 µl 1x TCL. As non-Biotinylated primers were used for Non-template strand DNA, ECs were purified by passing through Microcon-100 (Millipore). The Fe-BABE cleavage reaction was initiated by adding 25 mM H_2O_2 and 100 mM ascorbic acid to this tube, mixed well and the reaction was incubated for 1 minute, 2 minute and 5 minutes respectively. The reactions were stopped by extracting with phenol, followed by precipitation by ethanol in the presence of Glycogen and Ammonium Acetate, mixed with loading dye and run on 10% sequencing gel, exposed overnight to a phosphor imager screen. The DNA retained on the EC was quantified using Typhoon 9200 (Amersham).

4.2.6.6 Generation of radiolabelled RNAP markers

4.2.6.6.1 CNBr digested RNAP marker

Radiolabelled RNA polymerase (either at the ß or ß') was subjected to chemical cleavage at its Methionine residues using CNBr. The labeled protein (1µM) was incubated with 0.4% SDS in TCl transcription buffer in a final volume of 20 µl for 30 minutes at 30 °C. 50 mM HCl (1 µl from a 1M stock made in DEPC water) and 25 mM CnBr (1 µl from a stock of 0.5 M CNBr in Acetonitrile) was added to this, incubate for 5 minutes at 30 °C (Heyduk et al., 1996). The reaction was stopped by the addition of equal amount of 5x SDS load dye and was chilled on ice.

4.2.6.6.2 Generating LysC RNAP markers

Reaction mixture for Lysine specific cleavage contained 1 µM labeled RNAP, 8 M Urea (in 10 mM Tris HCl, pH 8.0) in excess (10 µl in a 20 µl reaction mix) and 30 ng LysC protein (Sigma) (0.3 µl from a 0.1 µg/µl LysC dissolved in 50 mM Tris-HCl stock). This was incubated at 25 °C for 10 minutes (enzyme concentration and time of incubation modified from Heyduk et. al, 1996). The reaction was stopped by the addition of equal amount of 5x SDS load dye and was chilled on ice.

4.2.6.6.3 Generating NTCB RNAP markers

Radiolabel led RNA Polymerase was dissolved in 8 M Urea with 10 mM ß-Mercaptoethanol (Sigma, pH 8.5) and kept for 2 hours at 37^{0} C. Five fold excess NTCB (Sigma) was added and this was incubated for 15 minutes at room temperature. The pH was then adjusted to 9.0 with 1M NaOH for the clevage reaction; and it was incubated for a further 30 minutes at 37 °C. Reaction was stopped by the addition of equal amount of 5X SDS load and was chilled on ice.

4.3 Results

4.3.1 *Design* of an N-modified stalled EC for biochemical probing

We have designed a stalled EC modified with NusA, NusG and H-19B N (both WT and mutant), using Lac repressor as a roadblock (cartoon shown in Figure 4.4A).

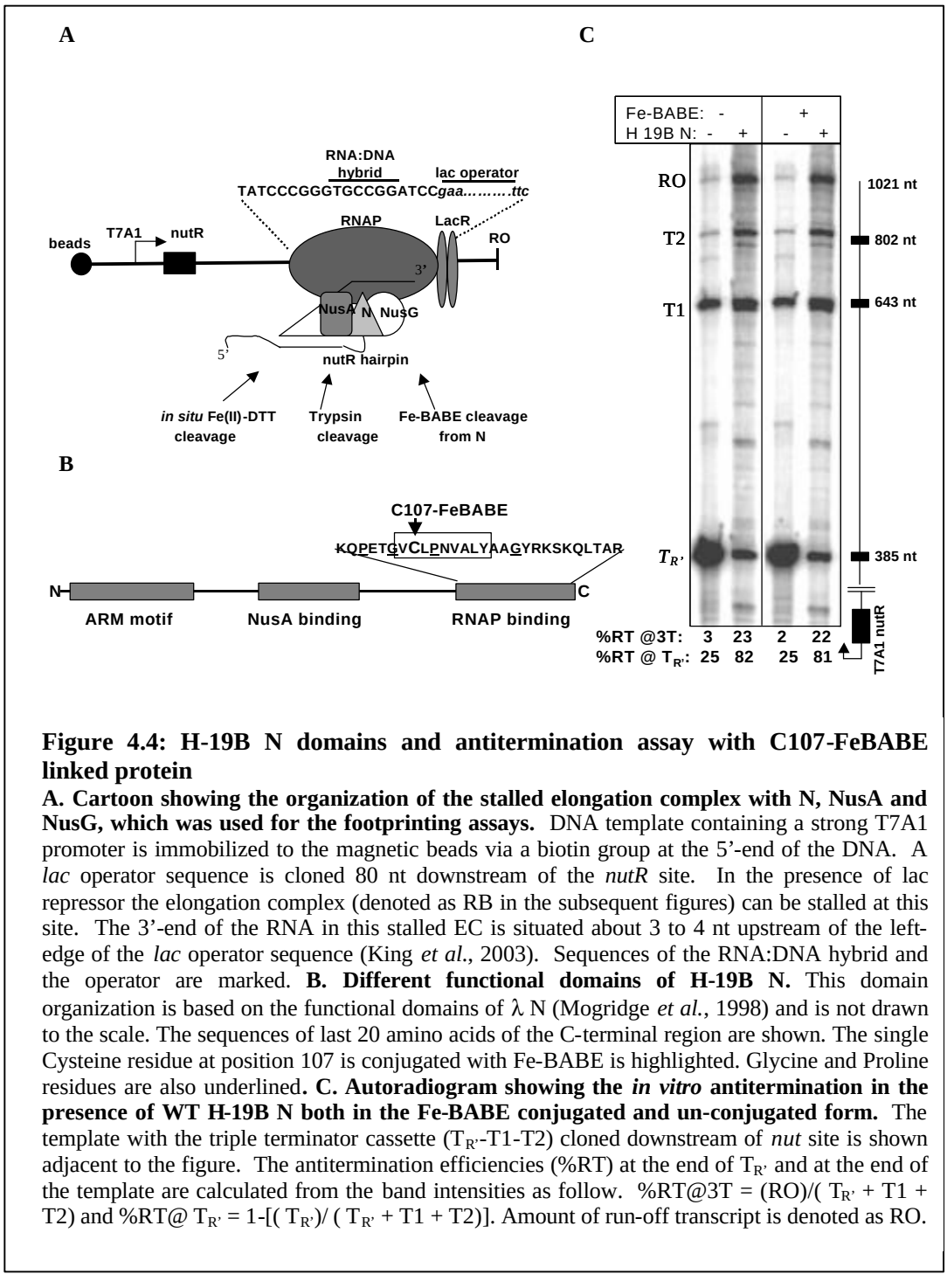


Figure 4.4: H-19B N domains and antitermination assay with C107-FeBABE linked protein

A. Cartoon showing the organization of the stalled elongation complex with N, NusA and NusG, which was used for the footprinting assays. DNA template containing a strong T7A1 promoter is immobilized to the magnetic beads via a biotin group at the 5'-end of the DNA. A *lac* operator sequence is cloned 80 nt downstream of the *nutR* site. In the presence of lac repressor the elongation complex (denoted as RB in the subsequent figures) can be stalled at this site. The 3'-end of the RNA in this stalled EC is situated about 3 to 4 nt upstream of the left-edge of the *lac* operator sequence (King *et al.*, 2003). Sequences of the RNA:DNA hybrid and the operator are marked. **B. Different functional domains of H-19B N.** This domain organization is based on the functional domains of λ N (Mogridge *et al.*, 1998) and is not drawn to the scale. The sequences of last 20 amino acids of the C-terminal region are shown. The single Cysteine residue at position 107 is conjugated with Fe-BABE is highlighted. Glycine and Proline residues are also underlined**. C. Autoradiogram showing the *in vitro* antitermination in the presence of WT H-19B N both in the Fe-BABE conjugated and un-conjugated form.** The template with the triple terminator cassette ($T_{R'}$-T1-T2) cloned downstream of *nut* site is shown adjacent to the figure. The antitermination efficiencies (%RT) at the end of $T_{R'}$ and at the end of the template are calculated from the band intensities as follow. %RT@3T = (RO)/($T_{R'}$ + T1 + T2) and %RT@ $T_{R'}$ = 1-[($T_{R'}$)/ ($T_{R'}$ + T1 + T2)]. Amount of run-off transcript is denoted as RO.

The *lac* operator sequence is placed about 80 nt downstream of the *nutR* site to facilitate the *nutR*-hairpin dependent binding of H-19B N to the stalled EC. This stalled EC remains active over the time period of *in vitro* assays (Cheeran et al., 2005; King et al., 2003). In all the protein foot printing assays, this stalled EC was formed using radiolabeled RNAP (see Materials and Methods). The single Cysteine residue (107th amino acid) at the C-terminal RNAP binding domain of H-19B N was conjugated with the hydroxyl radical generating reagent Fe-BABE (Figure 4.4 B; Meares et al., 2003). In order to know whether the interaction surface defined is physiologically relevant, we assayed the antitermination efficiency of the Fe-BABE labeled N at the end of a triple terminator cassette (Cheeran et al., 2005). We observed that Fe-BABE labeled N has a similar antitermination efficiency as the unlabeled N (Figure 4.4 C). Other than the wild type H-19B N with the a. a 107 to which the Fe BABE could be conjugated, we also generated A116C C107A single Cysteine mutant, to see if we could get a different cleavage pattern with Fe BABE. Here Fe-BABE conjugates to a Cysteine residue further towards the C terminal end of N protein. After processive antitermination assays, it was realized that the protein has very less processivity as far as processive antitermination is concerned (Figure 4.5). Very less amounts of antitermination (in the range of 8-10%) were obtained, compared to the Wt N, which yielded 20-35% antitermination.

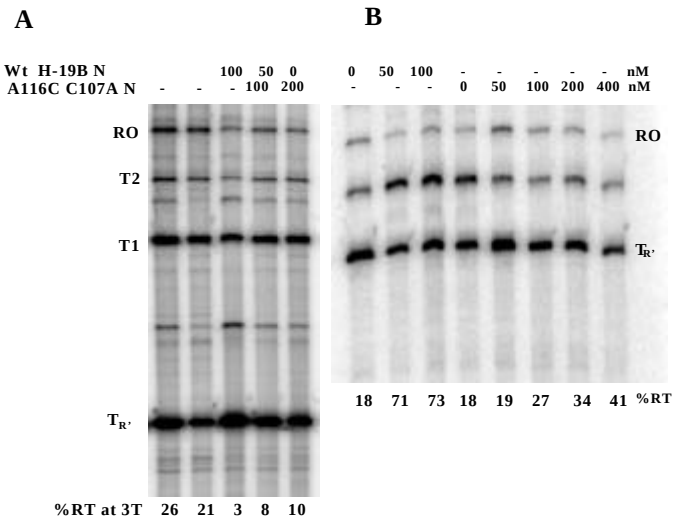


Figure 4.5: Antitermination assays on both triple terminator and single terminator templatesA. Antitermination on triple terminator (T7A1-*nutR*-3T) cassette. The transcription reactions contained 20 nM RNAP, 5 nM DNA template, 300 nM NusA, 200 nM NusG and 400 µM NTPs. Wild type/ A116C C107A H-19B N concentrations were as given in the figure. Antitermination rates obtained are given at the bottom. **B**. **Antitermination on single terminator ($T_{R'}$).** Transcription reaction conditions remained the same as above. Wild type/ A116C C107A H-19B N concentrations were as given in the figure. Antitermination rates obtained are given at the bottom

4.3.1.1 Fe-BABE labeled N generated cleavage patterns of b and b' subunits of RNA polymerase in the stalled EC

The N protein conjugated to Fe-BABE at Cys-107 generated specific cleavage products only in the β' subunit (lane 6, Figure 4.6 A), but not in the β subunit of RNAP of the stalled EC (Figure 4.6 B, compare lanes 5 with 3, 4 and 6). The cleavage was not observed when Fe-BABE was conjugated to a mutant N defective for binding to EC (ARM N; lane 7 of Figure 4.6 A), or with the un-conjugated Fe-BABE (lane 4 of Figure 4.6 A), which suggest that this cleavage pattern originated from specific binding of WT N to the EC. The Fe-BABE-mediated cleavage sites on β'were mapped close to amino acid residues 458 (conserved region D), 732 (conserved region F) and 926 (conserved region G). Cleavage at these three sites are usually obtained (lane 3 of Figure 4.6 A) when hydroxyl radical is generated from the active site by replacing Mg^{2+} with Fe^{2+} (Laptenko et al., 2003; Mustaev et al., 1997). We observed a major difference: no cleavage by Fe-BABE near amino acid 787 (compare lanes 3 and 6 of Figure 4.6 A) of β' subunit was detected.

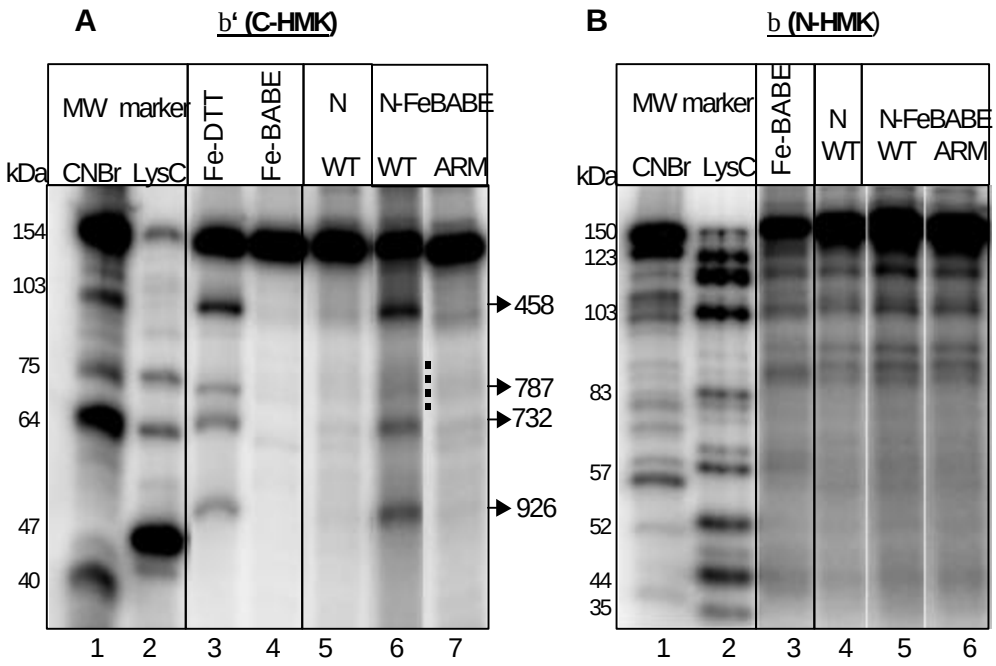


Figure 4.6: Protein cleavage profile of radiolabeled b'-subunit and b-subunit of RNAP in the stalled EC generated by Fe-BABE-conjugated H-19B N

A and **B**. Protein cleavage profile of radiolabeled **A)** β'-subunit and **B)** β-subunit of RNAP in the stalled EC generated by Fe-BABE-conjugated H-19B N. The EC is complexed with either un-conjugated (lane 5 of Figure 4.6 A and lane 4 of Figure 4.6 B) or Fe-BABE conjugated Wt N (lane 6 of 4.6 A and lane 5 of Figure 4.6 B). In lane 7 (Figure 4.6 A) and lane 6 (Figure 4.6 B), cleavage patterns are obtained in the presence of the H-19B N having mutations in N-terminal arginine rich motif (ARM). This mutant is unable to bind the elongation complex. Lane 4 (Figure 4.6 A) and lane 3 (Figure 4.6 B) are the non-specific cleavage patterns generated by free Fe-BABE. The positions of the cleaved products were obtained by comparing protein markers generated by CNBr (lane 1 of both the Figures 4.6 A and B) and LysC endoproteinease (lane 2 of both the Figures 4.6 A and B) cleavage of the same proteins.

Lane 3 of Figure 4.6 A shows the Fe-DTT cleavage pattern in β'-subunit from the active site. The concentrations of different components of the stalled EC were as follows: 1 μM RNAP, 100 nM Template DNA, 600 nM NusA, 300 nM NusG, 1 μM Wt H-19B N and ARM N. Fenton reactions were performed for 1 min. The locations of the cleavage products were identified from the polynomial fit of the calibration curves generated from the migrations of

We tested whether Fe-BABE-labeled N cleaves the RNA or DNA in the RNA:DNA hybrid using a similar experimental set-up as given earlier for FeBABE mediated footprinting of EC. However, we did not detect any such RNA or DNA cleavages (Figures 4.7 A and B).

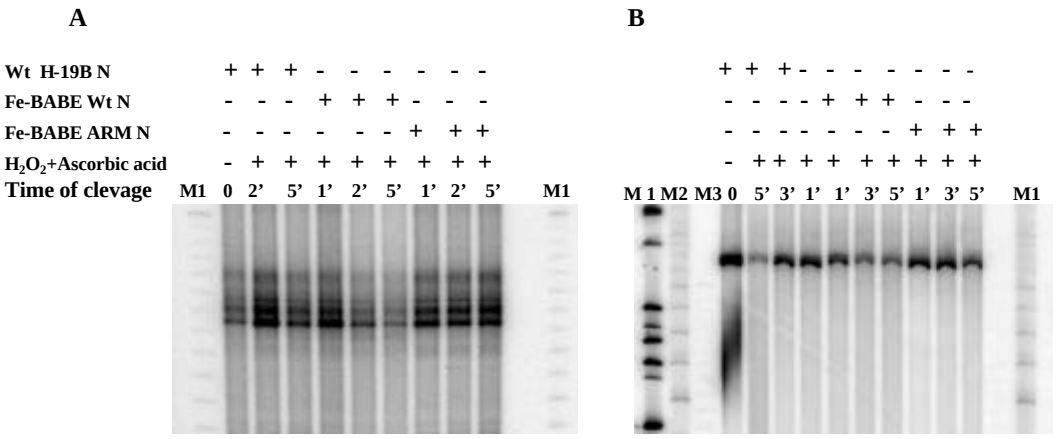


Figure 4.7: A. Fe-BABE clevage of RNA. The initial 22-mer RNA was labeled with a-P^{32}-CTP; M1 stands for 10 bp DNA marker (Invitrogen). The concentrations of different components of the stalled EC were as follows: 1 µM RNAP, 100 nM Template DNA, 600 nM NusA, 300 nM NusG, 1 µM Fe-BABE conjugated WT H 19B N and ARM N. The markers were **?**-P^{32}ATP labeled by PNK; refer Appendix III for protocol.
B. Fe-BABE clevage of template DNA M1, M2, M3 stands for bp (MBI), 25 bp and 10 bp (Invitrogen) DNA markers respectively; DNA was end labeled at the 3' end, see materials methods. Composition of the EC was same as above.

4.3.2 Probing the active site dynamics of N-modified elongation complex

Active site Mg (II) can be substituted with Fe (II) and *in situ* hydroxyl radicals can be generated in presence of DTT. These activated oxygen species are extremely reactive and can cause degradation of biopolymers within ~1 nm range (Mustaev et al., 1997). The changes observed in the cleavage pattern of the regions of beta and beta' subunits that are surrounding the active center should be an indication of the conformational changes in the active center. We monitored the cleavage pattern of the elongation complex in the presence and absence of WT or mutant N. All the other conditions were same as before in the case of Fe-BABE mediated footprinting, except that after the RB complex was formed, active site Mg (II) was replaced by Fe(II) by repeated washing of the complex with magnesium (Mg^{2+}) less transcription buffer and by subsequently adding Fe(II).

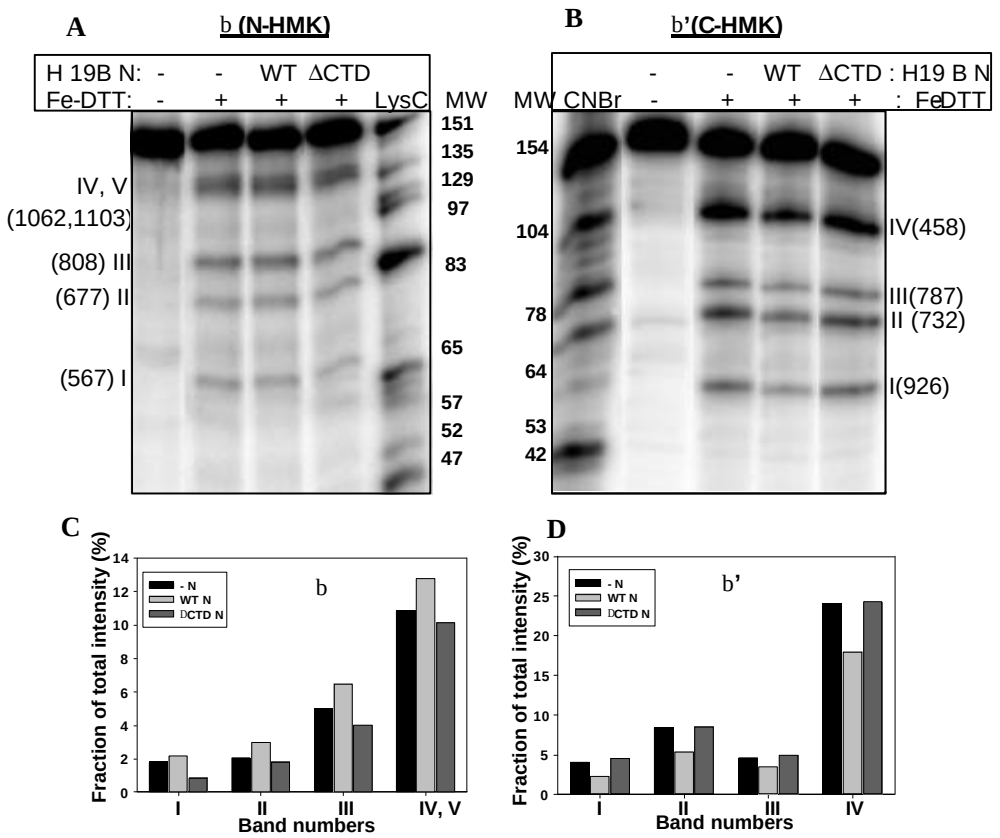


Figure 4.8: The protein cleavage profiles of A) b **and D)** b**'- subunits of RNAP in the stalled EC generated from the *in situ* Fe-DTT cleavage from the RNAP active site.**

A and **B. The protein cleavage profiles of A)** b **and D)** b**'- subunits of RNAP in the stalled EC generated from the *in situ* Fe-DTT cleavage from the active site of the RNAP** Stalled elongation complexes were made either in the absence or in the presence of Wt and mutant (ΔCTD) H-19B N. Position of the specific cleavage bands obtained in both the subunits are marked by comparing with the molecular weight markers generated from CNBr and LysC cleavages. Concentrations of different components of the elongation complex are same as above. Fe-DTT reactions were performed for 3 min.

C and **D**. **Bar diagrams depicting normalized intensities of cleavage products**

The intensity of the clevage products obtained from Fe-DTT footprinting are shown: C for β and D for β' subunit respectively. The intensities are expressed as the fraction of total intensity of all the cleavage products plus the un-cleaved band.

Characteristic cleavage patterns in the β and β' subunits (Mustaev *et.al.*, 1997 and Laptenko *et al.* and as shown in Figure 4.8 A and B respectively) were obtained, which matched with the earlier reports. The patterns were similar for both the free RNAP and the stalled elongation complex (Figure 4.8 A and B for elongation complex and Figure 4.6 A lane 3 for free RNAP). In the

presence of WT H-19B N there were modest changes in the intensity of the cleavage pattern. In general, the intensity of the bands in β' subunit was reduced when the EC was modified specifically with WT N (Figure 4.8 C and D), which may indicate structural rearrangement around the active center. Changes in cleavage pattern in β subunit were not significant.

4.3.3 Trypsin cleavage pattern of N-modified EC

The physical proximity of the C-terminal domain of H-19B N to the active center may induce some conformational changes in this region. To probe those changes, if any, we carried out limited Trypsin-cleavage of the β and β' subunits in N modified and unmodified EC (refer materials methods for cleavage protocol) and compared the Trypsin cleavage patterns obtained of these N-modified and unmodified stalled ECs. The Trypsin digestion patterns in the presence of WT N did not reveal protection of any regions of the β and β' subunits (Figure 4.9 A and B). However, binding of N induced enhanced cleavage at ß' positions R1123/K1132 and K731/R738, without changes in the digestion pattern of ß. The resolution of the gel was not high enough to identify the exact amino acids corresponding to the enhanced cleavage products. The enhancement in intensity was induced specifically by WT N, since it was not detected in the presence of Δ-CTD N, a mutant N that does not bind to the EC. These two enhanced cleavage positions are close to the active site (discussed later) which further supports the proposal that N induces conformational changes around the active center due to the physical proximity of its C-terminal domain to this region.

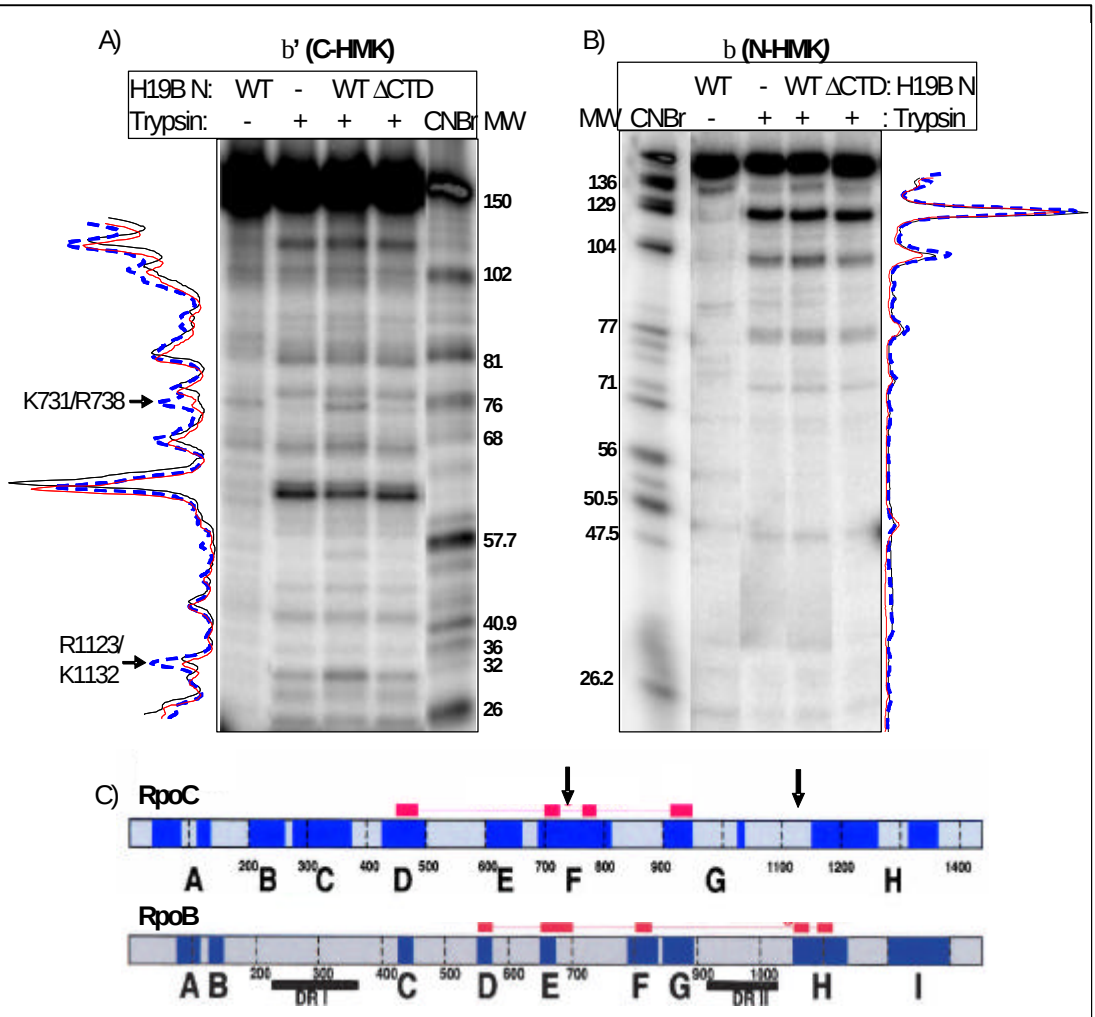


Figure 4.9: Limited Trypsin digestion pattern of A) ß' and B) ß subunits of RNAP in the stalled elongation complex in the absence and presence of either WT or mutant H-19B N (? CTD)

A, B The band intensities of each cleavage products are shown adjacent to each autoradiogram. The color codes are as follows. Blue – in the presence of WT N; Black – in the absence of N; Brown – in the presence of mutant N. In case of ß' subunit two cleavage products which are enhanced in the presence of WT N are indicated. The nearest two Arginine and Lysine residues corresponding to these bands are indicated. The positions of the cleavage products were identified from the molecular weight marker generated by CNBr. In these experiments Trypsin digestions were performed for 1 minute. Regions of ß and ß' subunits cleaved by Trypsin and Fe-DTT Arrows show location of a. a residue that had an increased sensitivity to Trypsin cleavage. Fe-DTT cleaved regions (Mustaev et al., 1997; Laptenko et al., 2003; Figure 4.8) are shown as red rectangular boxes.

4.3.4 Fe-EDTA mediated footprinting of N modified EC

Fe-EDTA Hydroxyl radical footprinting was carried out on Wt H-19 B N unmodified, modified and ARM/? CTD H-19B N modified roadblocked EC, as given in materials and methods section of this chapter.

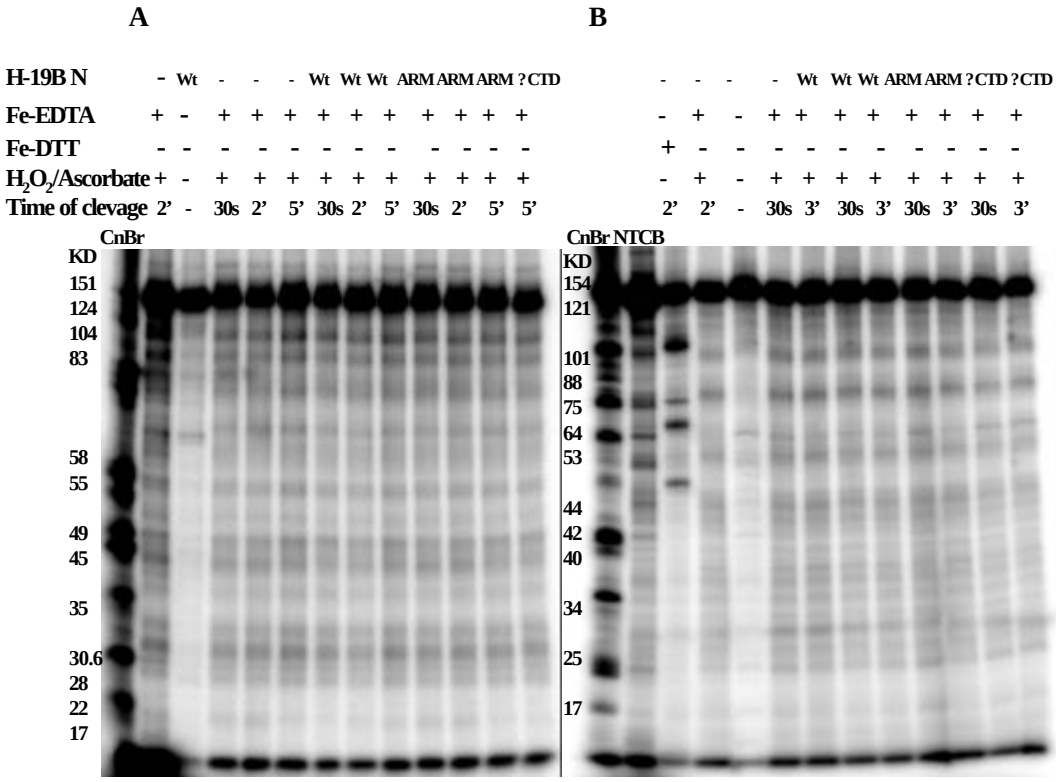


Figure 4.10: Fe-EDTA mediated clevage of *rpoB/C* in RB EC

A. **Fe-EDTA mediated cleavage of *rpoB* in a RB EC**
CnBr stand for *rpoB* molecular weight marker generated by cleaving RNAP (*rpoB* subunit radiolabelled) with CnBr. Formation of the EC (with no N modification, modified with Wt H-19B N, ARM N and ? CTD H-19B N as well as the Fe-DTT, Fe-EDTA cleavage was carried out as given earlier (refer Materials Methods). The concentrations of different components of the stalled EC were as follows: 1 µM RNAP, 100 nM Template DNA, 600 nM NusA, 300 nM NusG, 1 µM WT H-19B N, ARM N and ? CTD H-19B N. The clevage reactions were carried out for the time indicated.

B. **Fe-EDTA mediated clevage of *rpoC* in a RB EC**
CnBr and NTCB stand for *rpoC* molecular weight markers generated by cleaving RNAP (*rpoC* subunit radiolabelled) with CnBr and NTCB. Formation of the EC (with no N modification, modified with Wt H-19B N, ARM N and ? CTD H-19B N as well as the Fe-DTT, Fe-EDTA clevage was carried out as given earlier (refer Materials Methods). The concentrations of different components of the stalled EC were as follows: 1 µM RNAP, 100 nM Template DNA, 600 nM NusA, 300 nM NusG, 1 µM WT H-19B N, ARM N and ? CTD H-19B N. The clevage reactions were carried out for the time indicated.

Despite repeated attempts with varying concentrations of the constituents, we could not obtain a significant difference in the cleavage pattern or in the cleavage intensity between the H-19B N unmodified, Wt H-19B N modified, ARM H-19B N and ? CTD N modified EC using hydroxyl radical mediated cleavage as shown in Figure 4.10.

4.3.5 Mapping of the cleavage positions in the model structure of EC

In Chapter III, we have described H-19B N specific mutations RNAP mutations (P251S, P254L, R270C, G333S in β' and G1045D in β subunit) specific to H-19B N action, which are located in and around the RNA exit channel, and the lid and rudder elements of the elongation complex. In order to find the spatial relationship of the N-induced cleavage sites with the positions of those mutations, we localized both the cleavage sites and the positions of the mutations on the available model of EC, which has been developed on the basis of the crystal structures of *Thermus aquaticus* RNA polymerase (Zhang et al., 1999) and the yeast RNA polymerase II elongation complex (Gnatt et al., 2001), as well as the cross-linking data on protein-nucleic acids interactions in the EC of *E. coli* RNAP (Korzheva et al., 2000). We first determined the equivalent amino acids of these mutations and the cleavage sites in the model structure by structural alignment (Figure 4.11) using CLUSTALX program (Thompson et al., 1997). Figure 4.12 A shows the location of the cleavage sites together with H-19B N-specific mutations obtained earlier. Table 5 summarizes the molecular distances between the cleavage sites, N-specific mutations and different structural elements

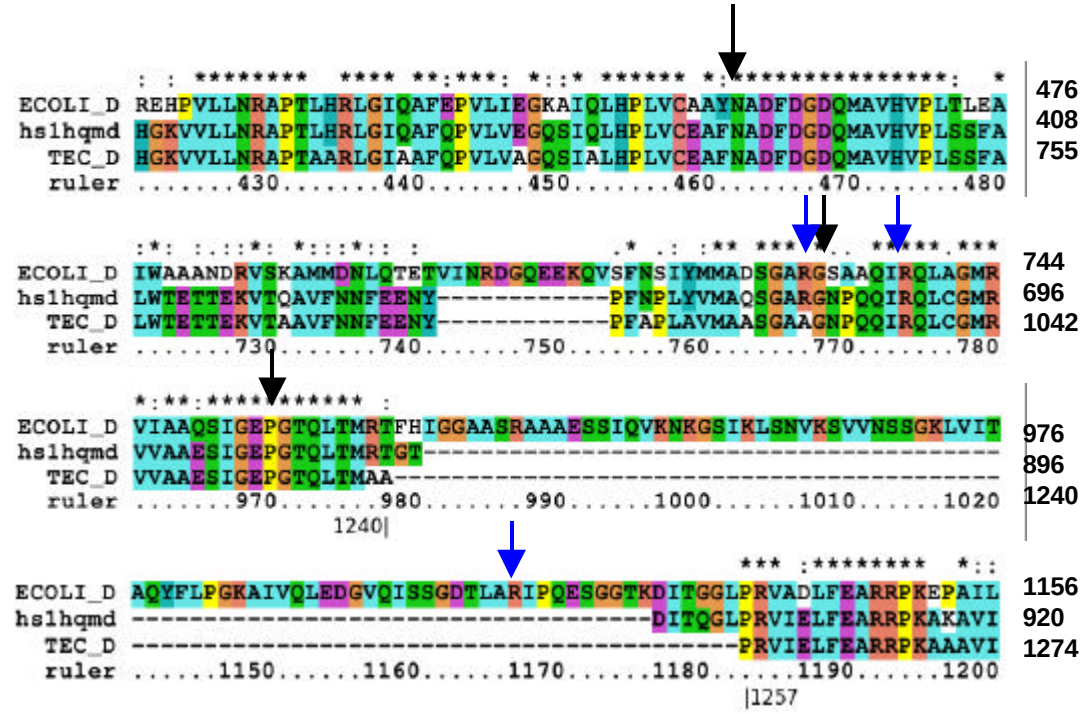


Figure 4.11: Location of cleavage sites and the H-19B N specific RNAP mutations on the model of EC

Structural alignment of the relevant regions of b **and** b**' subunits of *E.coli* RNAP with the equivalent regions of *Thermus acquaticus* RNAP using CLUSTALX**

Arrows indicate positions of the cleavage sites. Color codes are in accordance with the CLUSTALX notations. Hs1ddqc and hs1hqmd are the PDB identity numbers of *Thermus acquaticus* RNAP C (*rpoB*) and D chains (*rpoC*), respectively. TEC_C and TEC_D stand for ternary elongation complex C and D chains, respectively. Black arrows are for Fe-BABE-mediated cleavage sites, and blue arrows are for Trypsin-mediated cleavage sites.

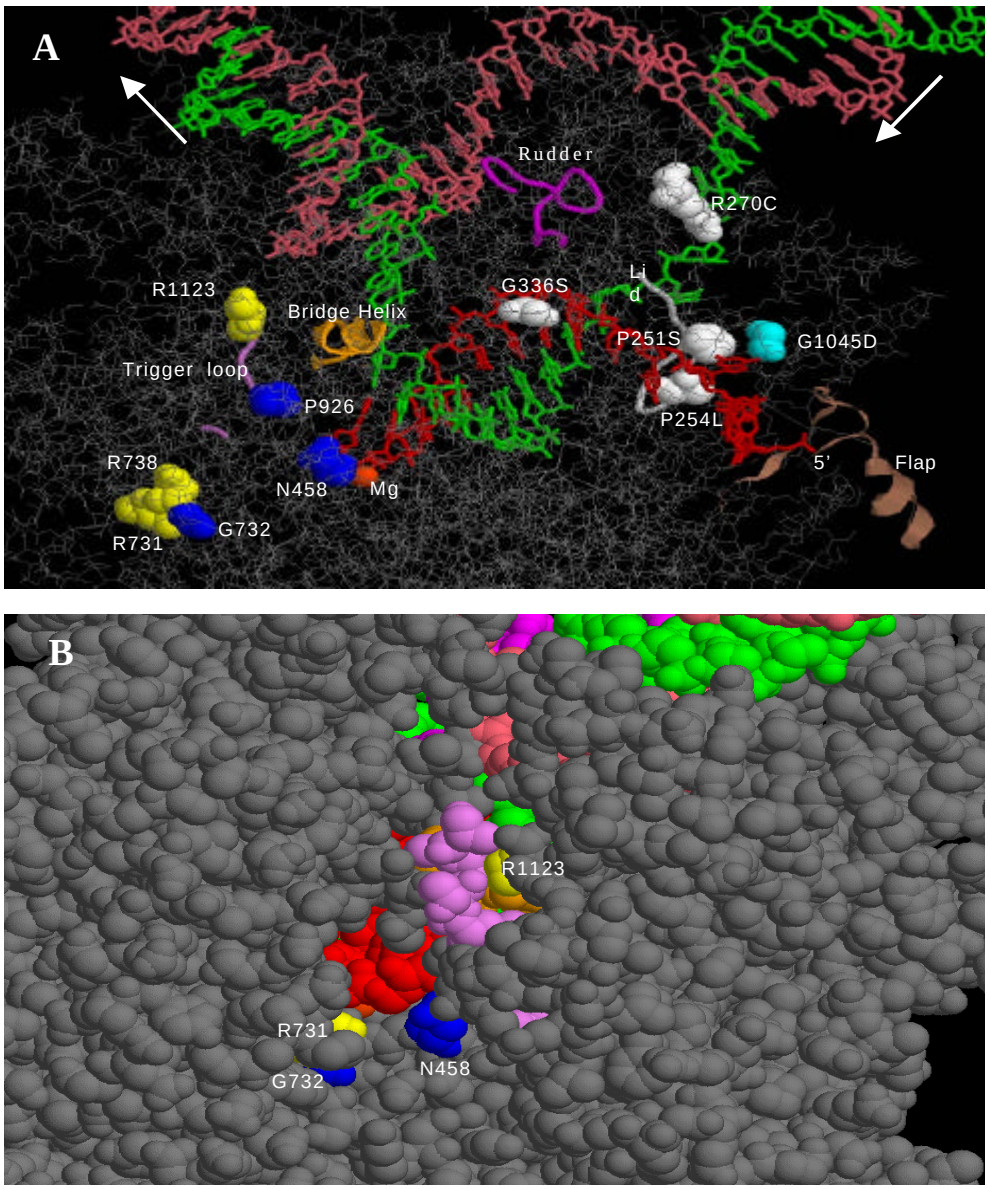


Figure 4.12: Location of cleavage sites and the H-19B N specific RNAP mutations on the model of EC

A. An expanded view of the nucleic acid framework in the of the model structure of EC using RASMOL. Locations of different mutations and cleavage sites are shown as space-filled in the wire-framed gray background of other parts of RNAP. The color-coding of template DNA, non-template DNA and RNA are green, pink and red, respectively. Active site Mg^{2+} is shown in orange. The *rpoC* mutants are shown as white space filled residues, while the *rpoB* mutant is colored cyan. The residues cleaved by Fe-BABE are colored blue, and the residues sensitive to Trypsin after N modification are colored yellow. Flexible loop structures of ß' rudder (colored pink), ß' lid (colored white), ß' bridge helix (colored orange), trigger loop (colored violet), and ß'flap domain (colored brown) are also indicated.

B. Space filled model of the EC surrounding the secondary channel area showing the surface accessibility of the cleavage sites. Color-coding is same as in B.

Table 5: Distances between the cleavage sites and N-specific mutations and different structural elements

A. Distances in Å between position of the H-19B N specific mutations and the cleavage sites.

***rpoC/B* mutant residues**	**N458**	**G732**	**P926**	**R731 (Trypsin)**	**R1123 (Trypsin)**
P251S	43.5	58.3	47.1	56.6	48.3
P254L	40.9	54.4	46.3	52.8	48.4
R270C	51.2	65.7	50.3	64.7	49.9
G336S	29.1	44.3	28.7	43.0	28.4
G1045D	53.2	66.1	57.7	64.7	58.2
***rpoC/B* mutant residues**	**Clevage residues obtained in *rpoB* subunit**				
	P567	**N677**	**N808**	**P1062**	**V1103**
P251S	35.3	47.4	55.6	40.1	54.5
P254L	32.3	42.6	50.6	34.7	49.8
R270C	39.9	53.7	68.3	55.7	63.5
G336S	21.9	35.1	47.5	37.7	43.1
G1045D	42.2	52.9	62.9	46.5	60.8

B. Distances in Å between the cleavage sites and the Active site Mg (II).

***rpoC* cleavage residues**	**Distances from Active site Mg in Å**
N458	8.6
G732	18.7
P926	17.8
R731 (Trypsin)	17.2
R1123 (Trypsin)	25.6
***rpoB* clevage** (Fe-DTT) **residues**	**Distance from Active site Mg in Å**
P567	14.3
N677	13.1
N808	22.2
P1062	17.2
V1103	17.2

C. Distances in Å between the cleavage sites and ß' bridge helix

rpoC/B mutant residues and clevage sites	Structural elements	
	ß Flap	**ß' bridge helix**
P251S	13.3	37.0
P254L	12.4	36.1
R270C	28.3	36.9
G336S	31.1	16.4
G1045D	14.5	46.4
N458	41.4	15.7
G732	52.8	24.3
P926	49.7	10.4
R731 (Trypsin)	55.2	22.9
R738 (Trypsin)	55.6	27.8
R1123 (Trypsin)	54.7	7.8

4.4 Discussion

It is well established that the specific interactions of the bacteriophage derived transcription antiterminators with the RNAP during elongation phase is mandatory to achieve the antitermination. Although several mutations in RNAP defective for N dependent antitermination has been reported earlier (Cheeran et al., 2005; Georgopoulos, 1971; Ghysen and Pironio, 1972; Jin et al., 1988a; Obuchowski et al., 1997; Schauer et al., 1996; Sternberg, 1976; Szalewska-Palasz et al., 2003), the interaction surface of this antiterminator on the EC remained to be elusive, knowledge of which is essential to understand the mechanism of antitermination.

In this report, we describe the region of EC that comes physically close to the C-terminal domain (the RNAP binding domain) of N from a lambdoid phage H-19B. The cleavage pattern derived from Fe-BABE conjugated N bound to a stalled EC revealed that the interaction surface to be close to the active site Mg^{2+} (Figure 4.6 A), which was further corroborated by the N-induced localized conformation changes nearby to this region of EC (Figure 4.8 A, B). The region of EC physically proximal to the C-terminal domain of H-19B N is close to and accessible easily from the secondary channel (Figure 4.12 B) but is quite distal to the RNA exit channel (~50 Å). Our results suggest that the C-terminal of N is likely to be inserted through this RNA exit channel. In order to cover this long distance, the C-terminal of N should be in an extended structure and interacting with the interior of the EC will stabilize this structure. Due to the presence of Proline and Glycine residues (Figure 5.13, discussed later) in this region it is not likely that it will able to form a long uninterrupted helix. Interestingly, the cleavage sites obtained from Fe-BABE or Trypsin (Figures 4.6, 4.9) are located close to the active site (refer Table 5 B for the distances), and did not overlap with the sites of the mutations. The space-filled version of the model (Figure 4.12 B) shows that both the Trypsin cleavage sites and two of the Fe-BABE-induced cleavage sites are visible through the secondary channel and located close to the surface. It is likely that binding of N makes the Trypsin cleavage sites more exposed to the surface. The *rpoC* clevage sites obtained were also quite close to structural elements like ß' bridge helix (see Table 5 C for distances). The Fe-BABE-induced cleavage sites are found to be ~30-60 Å away from the positions of the different mutations (see Table 5 A). Amino acid residue G333 comes closest (~ 29 Å) to N458 and P926, which are two of the Fe-BABE clevage sites obtained in *rpoC* subunit. Theoretically, hydroxyl radical generated from Fe-BABE tagged to Cysteine-107 of H-19B N can cleave the peptide backbone within a ~22 Å radius. Therefore, it is unlikely that the C-terminal domain of H-19B N binds to the region defined by the mutations as the hydroxyl radicals will not be able to travel over an average distance of ~45 Å to generate cleavages near active site Mg^{2+}.

Alternatively, mutations near the RNA exit channel may define the interaction site of NusA present in the N-NusA complex. We conclude that the C-terminal domain of H-19B N comes physically close to the active site Mg^{2+} and that this N-induced interaction near the active site exerts allosteric effects around the RNA-exit channel and the 5'-half of the RNA:DNA hybrid as defined by the location of the mutations.

CHAPTER V

Mechanism of Transcription Antitermination

5.1 Introduction

The isolation and characterization of RNAP mutants near the upstream edge of RNA:DNA hybrid and at the beginning of the RNA exit channel of the EC that are specifically defective for transcription antitermination mediated by N protein of a lambdoid phage, H-19B (Chapter III) as well as results from the foot printing of N modified EC pointed out that H-19B N could be exerting its effects in and around the active center (as discussed in Chapter IV). Our results did not show unequivocally that the mutant amino acids residues in the RNAPs defective for N mediated antitermination that we obtained are involved in the direct interaction with H-19B N; it is possible that the effect of these mutations are indirect. N can bind to a different part of RNAP and the effect might be exerted allosterically through this region. Due to the spatial location of the mutations close to the RNA:DNA hybrid and upstream part of the RNA exit channel, these altered amino acid residues could affect the interactions of RNAP with RNA:DNA hybrid and with RNA in the exit channel, which in turn could impair the process of antitermination. These is in accordance with the fact that a terminator hairpin brings about transcription termination by disrupting the RNA:DNA hybrid and by destabilizing the protein nucleic acid contacts in the EC and brings about conformational changes around the active center of RNAP. The regions in the TEC which we obtained as possible locations where H-19 BN could be exerting its effects from our biochemical studies involving the footprinting of the EC, define regions very near to the active center and structurally important elements like ß' bridge helix and trigger loop. From the location of the cleavage sites obtained, it could be seen that they were not on the surface of the EC. Two of the cleavage sites could only be visible through the secondary channel (Figure 4.12 B).

How can N access the regions buried inside the structure? In order to access this region, the C-terminal domain of H-19B N has to be inserted into the EC either through secondary channel or through the RNA-exit channel, which are the two openings through which the active center can be accessed and designed an experiment to address the question as to whether functional binding of N requires it to be inserted in the EC through the RNA exit channel or not. Other two experiments in this vein, to find if N comes close to the active center or if it brings about subtle changes around the active center was to find if the N modified EC can over come the effects of Streptolydigin and Tagetoxin, two antibiotics whose binding sites are close to the active site (Temiakov et al., 2005; Tuske et al., 2005; Vassylyev et al., 2005) and also to see if the differences in cleavage of the nascent RNA from the active sites of N modified and unmodified complexed using Fe-DTT clevage.

The other question to be addressed was if N modification of the EC can counter the destabilizing effects of the RNA terminator hairpin or not. To know about the elements/ regions of the EC that can get modified by the binding of N to the EC, we devised experiments, which looked into the relative stabilities of N modified and unmodified ECs, stalled at a terminator site. We also addressed the state of RNA hairpin in an N modified EC, to check if N modification disrupts hairpin formation or not. Assays using Class I (*his* pause) and Class II (*ops* pause) pause sequences (Artsimovitch and Landick, 2000; Artsimovitch and Landick, 1998; Chan et al., 1997) were carried out on appropriate DNA templates to see the effect of N modification on the paused EC, if any. We checked if N has any effect on the backtracking of RNAP on such paused templates and the effect of N modification on backtracked TEC complexes by checking for its sensitivity to GreB and PPi.

5.2 Materials and Methods

5.2.1 Stability of the N-modified EC stalled at a terminator

The fact that the N specific substitutions in RNAP exhibited a specific defect for H-19B N mediated antitermination as well as due to their localization close to the upstream edge of RNA:DNA hybrid and at the beginning of the RNA exit channel, made us hypothesize that, if N exerts its effect through this region, N-modification will affect either the formation of terminator hairpin in the RNA exit channel or stabilize the weak RNA:DNA hybrid or both. To test this, we investigated whether N-modification of the EC can impair RNA release efficiency from and increase stability to the EC that is artificially stalled at a terminator. Assays were compared between WT and G1045D mutant of RNAP. Furthermore, we attempted to probe the conformational state of the terminator hairpin in this N-modified stalled EC. We designed a template to stall the EC at the $T_{R'}$ terminator using lac repressor (*lacI*) as a roadblock (Figure 5.5 A). Previously it has been observed that, on such templates, *lac* repressor functions as an efficient roadblock and that the stalled EC retains its activity, with the 3'-end of the RNA remaining 3 to 4 nt upstream of the left edge of the *lac*-operator sequence (Sen et al., 2001; King et al., 2003). Hence, while designing the template we placed the *lac* operator sequence 4 nt downstream of the desired 3'-end of RNA:DNA hybrid. Placing a roadblock 4 nt downstream of the end of U-tract of the terminator did not alter the site of termination.

5.2.2 *In vitro* transcription

Transcription reactions were done at 32 °C in transcription (T-Glu) buffer containing, 20 mM Tris-Glutamate (Tris-Glu) pH 8.0, 10 mM Mg-Glutamate and 50 mM K-Glutamate. Reactions

were initiated with 175 µM ApU, 5 µM each of GTP and ATP and 2.5 µM of CTP to make a 22-mer EC. This EC was labeled with a-^{32} P-CTP. Then it was chased with 400 µM each of all the NTPS in presence of different concentrations of WT and mutant H-19B N proteins. Chasing reaction also contained 300 nM of NusA and 175 nM of NusG. Under this condition, reactions were essentially single round as the labeled CTP was diluted more than 1000-fold by the cold CTP. After 15 min, reactions were stopped by phenol extraction followed by ethanol precipitation in presence of glycogen. RNA pellets were then re-suspended in Formamide loading dye and loaded onto a 6% sequencing gel. Transcript intensities were quantified using Fuji 2500 phosphor imager (Fuji) or Typhoon 9200 (Amersham).

Transcription reactions with T7A1-*nutR -lacO* (Figure 5.1 A, 5.5 A) templates were done under same condition as above. In both the cases immobilized templates were used and about 100 nM Lac repressor was added when required, before chasing the 22-mer EC. For RNA release assays (Figure 5.5 B), reactions were chased for 3 min. and half of the supernatant was taken out for "S" lanes and the rest was for "S+P" lanes. For stability assays, after chasing for 3 min. beads were washed twice with Tris-Glu buffer + 100 mM NaCl, the pellets were resuspended in 10 µl T-Glu buffer and mixed with loading dye.

5.2.3 Probing the state of the terminator hairpin by RNAse footprinting and antisense oligo mediated cleavage

Stalled EC at the terminator was formed in a similar way as described in Figure 5.5 A, except that for RNAse footprinting experiments transcription reactions and washing of the beads were done in Tris-Cl buffer (50 mM Tris-Cl, pH 8.0 + 10 mM $MgCl_2$ +100 mM KCl). In absence of N footprinting was done before washing the beads whereas in presence of N stalled EC was washed extensively prior to footprinting. T1 and V1 reactions were done for 1 min and 45 sec, respectively. T1 clevage was done at 37 ^{o}C, whereas V1 was done at room temperature. Reactions were stopped by phenol extraction. For antisense oligo mediated cleavage assays we used a DNA oligo (oligo 6) antisense to the upstream part of the hairpin (Figure 5.7 B). Oligo6 was added to the reactions either during the chase of 22-mer EC ("to chase") or after the formation of the stalled complex ("to RB). 1 unit RNAseH was added and reactions were incubated for 3 min.

5.2.4 Pause Assays

Template DNA: DNA template encoding *ops* pause sequence downstream to H-19B *nut* site was amplified using RS83/RS248 primers (refer Appendix II), while the DNA template encoding the

his Pause was PCR amplified using primers RS58/83 and RS263, 264 and RS265 (refer Appendix II) in tandem, using each product as the template for the next round of amplification, using DeepVent (NEB) proof reading enzyme. A *lac* operator region encoded by an oligonucleotide was used to incorporate the *lac* operator downstream to the *ops* pause region for the roadblock assays with GreB and Ppi.

5.2.4.1 Transcription assays with pause templates

For pause assays, the templates with T7A1-*nutR-ops* and T7A1-nutR-*his* or T7A1-*nutR* (inherent H-19B pause*)* constructs were used. We found out that there is an inherent pause site in the H-19B DNA, just after the *nut* site. We did RNA sequencing (Figure 5.10) to locate the exact pause position brought about by the pause site inherent in H-19B DNA, just after the *nut R* region and designed the *ops* and *his* pause templates in such a way that this pause DNA sequences were incorporated to the H-19B DNA between the *nutR* sit, but before the inherent pause site. For assays with the pause template, at first a 22-mer elongation complex (EC_{22}) was made by initiating the transcription with 175 µM of ApU, 5 µM each of GTP and ATP, 2.5 µM of CTP, and α-^{32} P-CTP (3000 Ci/mmol, Amersham). For the kinetics of elongation on *ops* pause template the EC_{22} was chased in the presence of 100 µM each of UTP, CTP and ATP and 10 µM of GTP. When required NusA, NusG, WT and ΔCTD N were added to the chase solution. 5 µl aliquots were removed at indicated time and mixed with equal volume of Formamide loading dye. For assays on *his* pause template EC_{22} was chased in the presence of 150 µM of UTP, CTP, ATP and 10 µM of GTP. Products were analyzed on 8% sequencing gel.

For making the stalled complexes (RB) at *ops* and *his* pause sites, 5'-biotinylated templates with *ops* and *his* pause sequences fused to *lac* operator sequences were used. The stalled EC (RB) using Lac repressor as roadblock were formed essentially the same way as the protein-cleavage assays. Nus factors and N were added to this chase reaction when required. RB complexes were washed with transcription buffer to remove the NTPs. Indicated amount of GreB (Figure 5.8) or Na-Pyrophosphate (Figure 5.9) were then added to the respective RB complexes. Products were analyzed on 8% sequencing gels. For GreB/Ppi clevage of the *His* pause and *Ops* pause Roadblock complexes, a labeled 22- mer EC was made as earlier. This mix was aliquoted to fresh tubes and the different chase solutions having 100 µM each of all the NTPs with 1x Tris-Glu, 0.1 mg/ml BSA and proteins (300 nM NusA, 200 nM NusG, 200 nM Wt H-19B or CTD H-19B N), as per requirement of each reaction, were added to separate tubes. Transcription was carried out at 30 °C for 2 minutes. Under these conditions, the reactions were essentially single-round, as the

labeled CTP was diluted more than 1000-fold by the cold CTP. Three separate sets of EC s having NusA and NusG but no Wt H-19B N, with NusA NusG and Wt H-19B N, and NusA, NusG and CTD N were made. Each tube was washed 3x with 50 µl of 1x Tris-Glu buffer and resuspended in 8 µl in the same buffer. Required concentrations of GreB/ Ppi were added. For GreB clevage, the tubes were incubated for 2 minutes, while for Ppi different tubes with time points as required (between 3 min, 7 min and 10 min) were incubated at 37 ^{o}C. The reactions were stopped by extraction with phenol, Formamide loading dye was added, heated for 4 minutes at 95 ^{o}C and loaded onto 8% sequencing gels, exposed to a Phosphor imager screen (Amersham), and the transcript intensities were quantified using Typhoon 9200 (Amersham).

5.2.4.2 RNA sequencing

The DNA whose corresponding RNA was to be sequenced was used as the template in the *in vitro* transcription reaction. A 22-mer elongation complex (EC_{22}) was made by initiating the transcription with 175 µM of ApU, 5 µM each of GTP and ATP, 2.5 µM of CTP, and a-^{32}P-CTP (3000 Ci/mmol, Amersham) for 7 minutes. The short transcript was then aliquoted equally to four eppendorf tubes. Equal amounts of appropriate chase solutions were then added and incubated for 5 min for transcription to occur (here each of the chase solution contains 100 µM of three NTPs, 2.5 µM of another NTP and 100 µM dNTP corresponding to the NTP which was in lesser concentration). The reactions were stopped by addition of equal volumes of phenol and the RNA was extracted with NH_4OAc/Glycogen-Ethanol precipitation, and the washed, dried pellets were resupsended in Formamide loading dye. Products were analyzed on 8% sequencing gel.

5.3 Results

5.3.1 C-terminal of N maybe transferred to EC through RNA exit channel

5.3.1.1 Efficiency of antitermination is dependent on N binding in proximity to the RNA exit channel

The insights we obtained from the mutational study as well as the footprinting of EC helped us design experiments, which could give an insight into the molecular mechanization of N protein in the EC. From the location of the cleavage sites obtained, it could be seen that they are not on the surface of the EC. Two of the cleavage sites could only be visible through the secondary channel (Figure 4.13 B). How can N access the regions buried inside the structure? In order to access this region, the C-terminal domain of H-19B N has to be inserted into the EC either through secondary channel or through the RNA-exit channel, which are the two openings through which the active center can be accessed. We hypothesized that the N-terminal ARM motif of N will recognize the *nut* site of the nascent RNA as soon as the RNA emerges from the exit channel and due to the physical proximity, the C-terminal of N**,** with the help of NusA, can be inserted into the EC through the opening of the RNA exit channel. In order to test this hypothesis, we formed a 22-mer EC on a template described in Figure 5.1. We either added WT H-19B N to the 22-mer EC and chased it through the downstream $T_{R'}$ terminator (Figure 5.1 B, lanes 5, 6 and 7) or added N to a preformed stalled EC using Lac repressor as a road-block downstream of the *nut* site and chased the complex in the presence of IPTG through the same $T_{R'}$ terminator (Figure 5.1 B, lanes, 2, 3, 4). In the first case, N is present from the beginning of the elongation process and will bind to the *nut* site when it is very close to the RNA exit channel. When N is added after the stalled EC (RB) is formed, it will bind to the *nut* site which is now ~ 50 nt away from the RNA exit channel. If functional binding of N requires it to be inserted in the EC through the RNA exit channel, then we should observe efficient antitermination only when N is present from the onset of the reaction. Consistent with this hypothesis, we observed efficient antitermination (Figure 5.1, lanes 5 and 7) only when the N was added to the 22-mer EC ("to chase"). When N was added to the preformed RB ("to RB"; lane 4) and allowed the EC to be chased through the downstream terminator, we did not observe efficient antitermination (Figure 5.1B compare lane 4 with 1 and 3). In order to ascertain that this observation is due to specific binding of N to EC, we performed the same experiments with a RNAP mutant, G1045D, which is specifically defective for H-19B N-dependent antitermination. No antitermination was observed in this case under any condition (Figure 5.1B lanes 8 to 14).

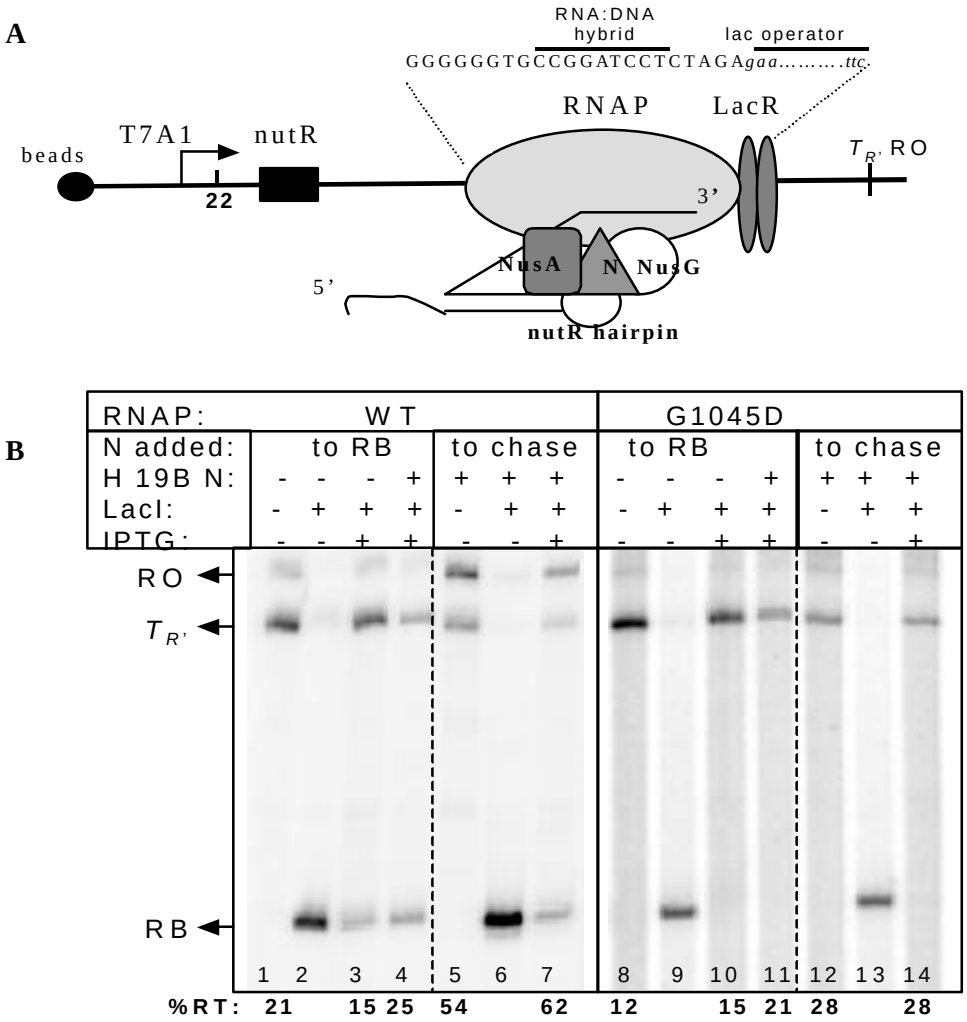


Figure 5.1: Roadblock assay for N binding

A. Cartoon of the N-modified stalled elongation complex road-blocked at a 22 bp *lac* operator site which is cloned 64 nt downstream of the *nutR* site and upstream of $T_{R'}$. On this template upon addition of IPTG the stalled complex can be chased through the downstream terminator and the antitermination efficiency can be measured under different conditions.

B. On this template at first a 22-mer-elongation complex (EC_{22}) was made with either with WT or G1045D RNAP. In one set (lanes 5-7 and 12-14) this EC_{22} was chased through the downstream terminator $T_{R'}$ in the presence of N either in the absence (lane 5 and 12) or in the presence (lanes 6, 7, 13 and 14) of lac repressor. In another set (lanes 2-4 and 9-11) at first the EC_{22} was chased in the presence of lac repressor to make the stalled complex (RB). Wt H-19B N was added to lanes 4 and 11 and the RB was then chased upon addition of 1 mM IPTG. In lanes 3 and 10 chasing were done in the absence of N. Run-off product is denoted as RO. The antitermination efficiencies (%RT) are shown below each lane and are calculated as: %RT = (RO)/ (RO + $T_{R'}$). In these reactions concentrations of different components were 10 nM RNAP, 5 nM DNA templates, 200 nM N, 300 nM NusA, 200 nM NusG, 100 nM LacI and 400 μM of NTPs.

5.3.1.2 Fe-DTT cleavage of H-19B N

The hypothesis that H-19B N can come close to the active center of the EC, made us think of designing an experiment where by we could probe whether H-19BN can be cleaved by Fe-DTT, when the active site Mg^{2+} is replaced with Fe and Fe-DTT cleavage reaction (as described in chapter 1V) carried out. Here HMK tagged H-19 B N (refer Materials methods of Chapter 1V) was used in the transcription reactions to make the RB. We could not get a Fe-DTT mediated cleavage of H-19B N, despite repeated attempts (Figure 5.2). This also points out to the fact that N may not approach the secondary channel of RNAP.

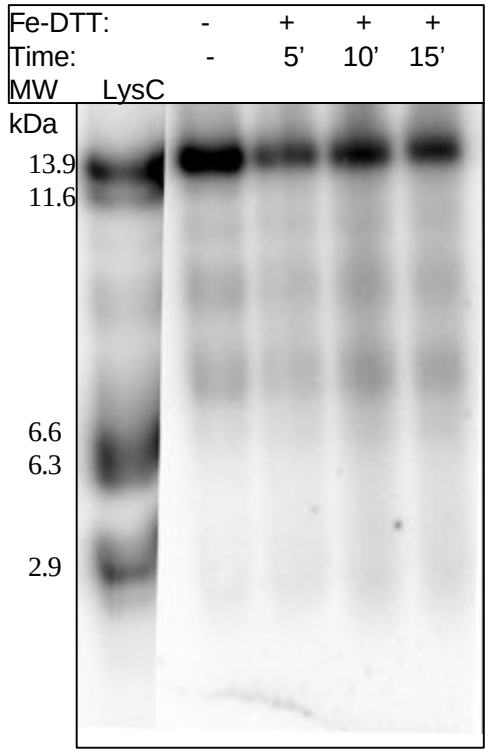


Figure 5.2: Cleavage of H-19B N present in the RB by Fe-DTT from the active center of the EC
The RB complex was made as described earlier for cleavage reactions, in the presence of Wt H-19B N, end labeled at C terminal. The RB complex was washed with transcription buffer with out $MgCl_2$, and resuspended in the same buffer. Fe-DTT cleavage reactions were initiated by the addition of 5 mM DTT and 40 mM of freshly prepared Ferrous Ammonium Sulphate. Reactions were terminated at indicated time points by addition of equal volume of 5x SDS-loading dye, heated to 95^{o}C for 4 min, and were loaded on to an 16% SDS-Tricine PAGE. LysC marker was generated by cleaving H-19B N with LysC protease as given earlier in Chapter IV. The concentrations of different components of the stalled EC were as follows: 1 mM RNAP, 100 nM Template DNA, 100 nM LacI, 600 nM NusA, 300 nM NusG, and 1µM ?-P^{32} ATP-labeled C terminal HMK Wt H-19B N.

5.3.1.3 N binding does not effect the actions of Tagetoxin (TGT) and Streptolydigin (STL)

To further rule out the possibility that N approaches the active site through the secondary channel, we assayed the effects of two antibiotics, Streptolydigin and Tagetoxin which has been found to act on residues close to the active site (Temiakov et al., 2005; Tuske et al., 2005; Vassylyev et al., 2005). The amino acid residues that this drugs interact with in the EC, were found to be quite close to the residues and structural elements that we obtained through our foot printing

experiments, which we postulated were the plausible sites were N interaction could occur. Streptolydigin should also get the access of the active site of the EC via secondary channel (Temiakov *et al.*, 2005; Tuske *et al.*, 2005).We wanted to see if the N modified ECs will be able to overcome the effects of this drugs. We did transcription assays for assaying the effects of this drugs, either as a competitor of N for binding (the antibiotics were added to the 22-mer initiation complex, incubated for 3 minutes and then chased with 100 µM NTP along with 300 nM NusA, 200 nM NusG and either in the presence or absence of 200 nM Wt H-19B N; here the 22-mer was chased for different time points ranging from 15 seconds to 5 minutes, Figure 5.3); or to assay whether N can have an effect on or alleviate the effects of the antibiotics on the road blocked EC which is already treated with the antibiotics (Figure 5.4). Here different RBs-RB EC formed without Wt N or with Wt N were made, washed twice and then either of the antibiotics were added to the EC, incubated for 3 minutes and then chased with IPTG for different time points.

The results we obtained clearly show that N couldn't overcome the action of both these antibiotics; either as a competitor or even on an N modified EC subjected to the action of Tagetoxin or Streptolydigin later on. It could be inferred that N addition does not block the secondary channel and though this drugs acts on amino acid residues and structural elements found quite close to the N interacting residue we found in our study, N as such does not have any direct interactions with these sites and the mechanistic aspects brought on in and around the active site by N binding could be the deciding factor towards the modification of RNAP to overcome the intrinsic terminators and towards processive antitermination.

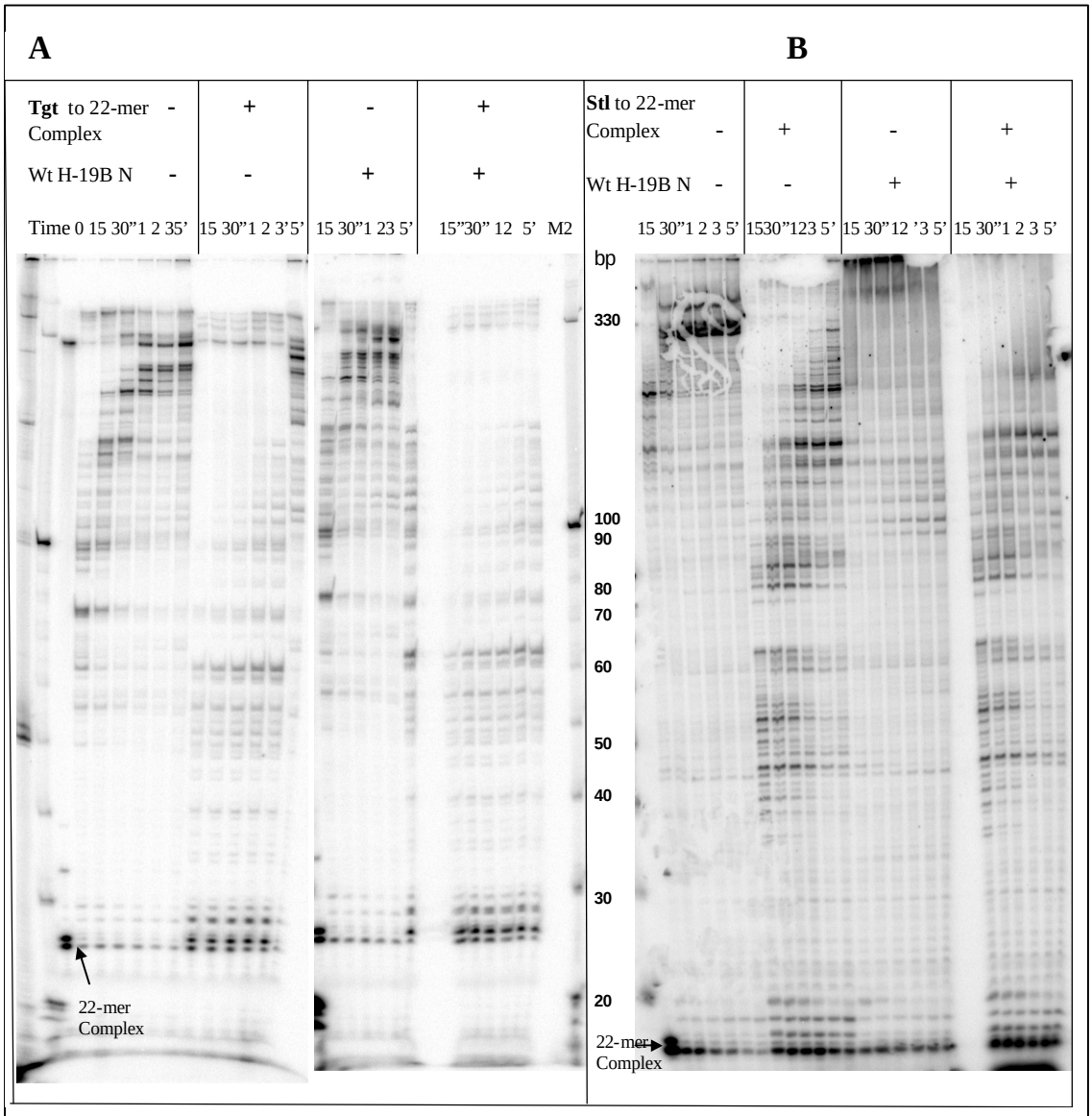


Figure 5.3: Tagetoxin (Tgt) and Streptolydigin (Stl) kinetics on *ops* pause template

A. Tagetoxin (Tgt) kinetics through *ops* pause template
B. Streptolydigin (Stl) kinetics through *ops* pause template
M denotes 10 bp (Invitrogen) DNA markers (the markers were **?**-P^{32}ATP labeled by PNK, refer Appendix III); 22-mer Initiation complex is shown with arrows. Concentrations of the components during the different stages of transcription were 20 nM RNAP, 5 nM DNA template, 100 nM LacI, 200 nM Wt H-19B N, 300 nM NusA, 200 nM, NusG and 100 µM NTPs. Tagetoxin and Streptolydigin were added to final concentrations of 60 µM and 0.5 mM respectively. These are saturating concentrations for these two antibiotics.

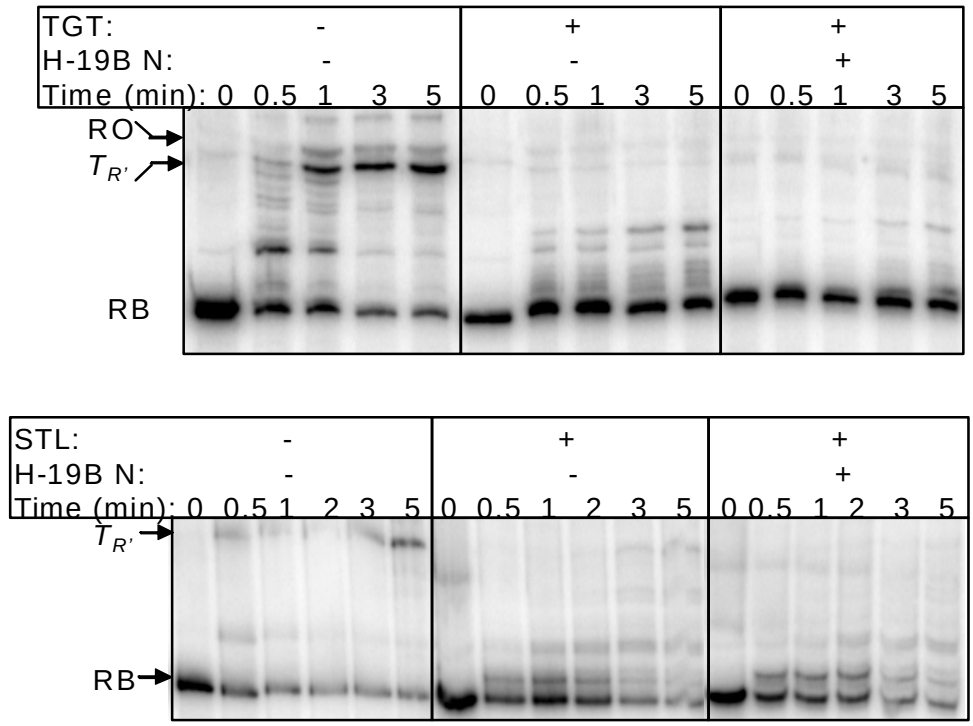


Figure 5.4: N binding effect on the actions of Tagetoxin (TGT) and Streptolydigin (STL).

A. Effects of N binding on a Tgt bound RB EC

B. Effects of N binding on a Stl bound RB EC

A stalled elongation complex (RB) was made as both in the absence and presence of WT H-19B N. The complex was washed thoroughly to remove NTPs and then was either incubated with Tagetoxin or Streptolydign or with buffer. The RB was then chased with IPTG and NTPs for indicated time periods. Concentrations of the components during the different stages of transcription were 20 nM RNAP, 5 nM DNA template, 100 nM LacI, 200 nM Wt H-19B N, 300 nM NusA, 200 nM, NusG and 100 µM NTPs. Tagetoxin and Streptolydigin were added to final concentrations of 60 µM and 0.5 mM respectively. These

5.3.2 N stabilizes a stalled Elongation Complex at a terminator

We could not obtain a cleavage of the C terminus of H-19B N with Fe^{2+} from the active center and the experiments with Stl and Tgt showed that N modification does not help an EC in overcoming the effects of these drugs (which binds very near to the active center) by conformational rearrangements in the active site. Hence we started thinking about ways as to find the locations that can get modulated by N binding and the effects it might bring on. Since the formation of a terminator hairpin and an unstable RNA-DNA hybrid are pre requisites for factor

independent termination to occur, we decided to have a close look at the stability of the RNA, in an N modified complex stalled at a terminator as well as to probe the state of RNA hairpin.

We formed stalled ECs on an immobilized template, modified with N, NusA and NusG (Figure 5.1A for the assembly of the RB complex) in presence of *nutR* site. The *-LacI* lanes describe amount of termination at the $T_{R'}$ terminator in the absence of any roadblock (Figure 5.1B). RNA release efficiency from the road blocked complex at $T_{R'}$ was almost 100% both in absence of WT N and in presence of ARM mutant of N. In presence of WT N this efficiency was reduced to around 30%. Similar assays with G1045D mutant showed 90 to 100% release of RNA either in presence or absence of H-19B N. Therefore, it can be inferred that N modification can impair the RNA release significantly from the stalled EC. In order to assess the stability of this N-modified EC stalled at the terminator, we washed the magnetic beads (EC is on the beads) extensively with transcription buffer supplemented with 100mM NaCl. Consistent with the reduced RNA release efficiency, a significant amount of N-modified WT-EC (~50%) survived the washing and was found to be retained in the pellet fraction, whereas under similar condition G1045D-EC was readily dissociated (Figure 5.2). Failure of G1045D in preventing RNA release as well as to impart stability to the stalled EC in presence of H-19B N, again support the proposition that this mutant is severely defective for N-mediated antitermination, as well as to the fact that N protein do function to stabilize the EC on terminators.

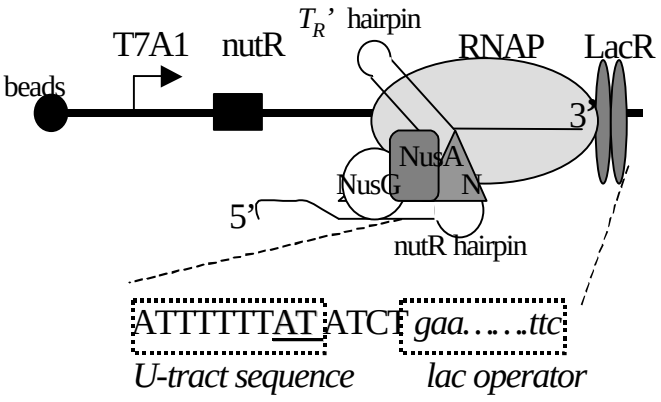


Figure 5.5: H-19B N modified RB EC
A. Cartoon of an N-modified EC stalled at $T_{R'}$ terminator using lac repressor (*LacI*) as roadblock (RB) *lac* operator sequence is cloned 4nt downstream of the termination point (marked by underline). Stalled ECs with WT and G1045D RNAP were formed in presence of NusA, NusG and H19B N (both WT and ARM mutant) proteins. 5'-end of the template was attached to streptavidin-coated magnetic beads through Biotin-Avidin linkage.

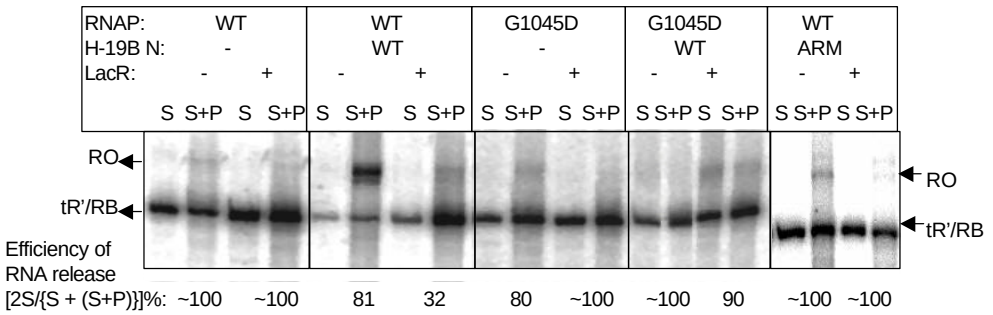


Figure 5.6: Efficiency of RNA release from N modified/unmodified EC stalled at a terminator

RNA released at T_R, terminator (S = half of the supernatant) and that remained with the beads (S+P) in the presence and absence of lac repressor roadblock. Elongation was performed either in the presence or absence of N (WT and ARM), NusA and NusG. Bands corresponding to run-off (RO), terminated ($T_{R'}$) and road-blocked (RB) transcripts are denoted. Efficiency of RNA release at the termination point in presence and absence of roadblock is shown below each lane. Concentrations of different factors were same as Figure 5.1.

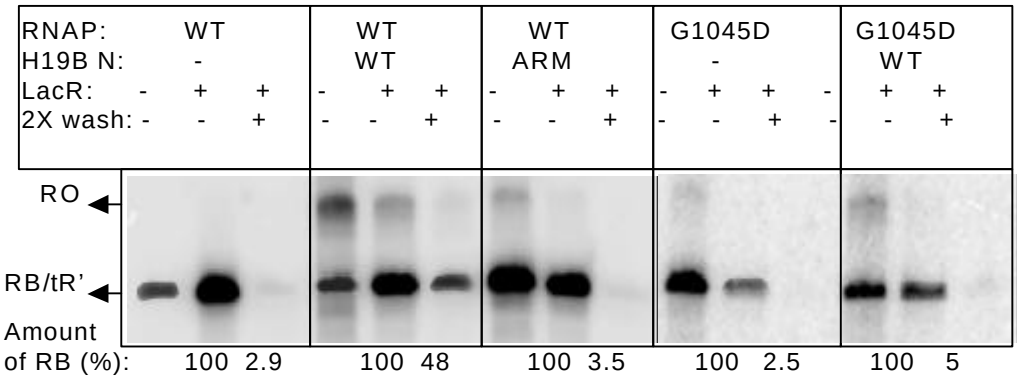


Figure 5.7: H-19B N modified road-blocked complexes at terminator

RBs were formed by transcribing WT or G1045D RNAP for 2 min in presence of Lac repressor protein and either in presence or absence of 200nM N. In the "+2X wash" lanes beads were washed twice with T-Glu buffer with 100mM NaCl, resuspended in T-Glu buffer and loaded onto the gel. Amount of EC retained with the beads were expressed relative to the unwashed lanes. – *LacI* lanes are to show that $T_{R'}$ and RB positions are same. Concentrations of different components were 10 nM RNAP, 5 nM DNA templates, 200 nM N, 300 nM NusA, 200 nM NusG, 100 nM LacI and 400 μM of NTPs.

5.3.3 State of the terminator hairpin in the N-modified stalled EC

As N modification imparts significant stability to EC, even when stalled at a terminator, we further investigated if H-19B N prevents the formation of terminator hairpin in this complex. We probed conformational state of the hairpin RNA of the stalled EC on an immobilized template described in figure 5.5 A, both in presence and absence of H 19B N, by using RNase T1 (specific for single stranded G residues) and RNase V1 (specific for double stranded RNA) clevage.

As can be inferred from Figure 5.8 A, lanes 1-7 represent reactions in absence of H-19B N, while lanes 8 -14 represent those in presence of it. In absence of N, more than 90% of the RNA was released from the EC when stalled at the terminator, as is evident from the non-occurrence of the RNA band upon washing the beads after the transcription reaction (Figure 5.8 A, lane 1). RNases were added before washing the beads and the cleavage pattern of the released RNA with stem-loop hairpin structure was analyzed. As evident from the Figure 5.8 A (lanes 2-7), the hairpin structure is more sensitive to T1 than V1 cleavage. Two G residues in the loop region of the hairpin were found to be sensitive to T1 (Figure 5.8 A, lanes 6 and 7). We observed a V1 sensitive band at the upper part of the left shoulder of the hairpin. Other cleavage sites outside the stem-loop region are indicated in the figure. In contrast to the above 90% release of RNA in the absence of N, in presence of N, about 50% of the EC was retained at the terminator even after extensive washing (Figure 5.8 A, compare lanes 8 and 9; also see Figure 5.7). The cleavage pattern obtained by T1 and V1 on the washed beads should represent the cleavage pattern of the RNA associated with the N-modified EC (lanes 10 to 14). Comparison of lanes 6, 7 and 13, 14 indicate that T1 cleavage pattern of the loop region does not change either in absence or presence of H-19B N. If N indeed had prevented the formation of hairpin in the modified EC, additional T1 sensitive regions were to be expected in the cleavage pattern of the RNA associated with the modified EC. No such additional T1 sensitive regions were observed in our experimental set up. However, the upper part of the left shoulder of the stem region showed even weaker sensitivity to V1 (lanes 10 to 12). These data suggest that upper part of the hairpin may still be intact in presence of N. At the same time, due to the lack of sensitivity of RNase V1 to the base part of the hairpin, we are not sure about the complete formation of the hairpin and desist from stating that N may prevent the formation of a terminator hairpin.

In order to further explore the state of the hairpin in presence of H-19B N, we measured the accessibility of the upstream part of the stem region to an antisense oligo (oligo 6, Figure 5.8 B). The accessibility was determined by RNAse H cleavage that specifically cleave RNA:DNA hybrid. Addition of the oligo during the transcription reaction (Figure 5.8 B, “to chase” lanes),

both in absence and presence of N, results in about 80% cleavage of the RNA:DNA hybrid. Under this condition oligo 6 can access the RNA before the hairpin is formed. We added the oligo after the stalling the EC at the terminator ("to RB" lanes), which allows sufficient time for the hairpin to form. (In the absence of N, as the majority of the complexes were dissociated after stalling at the terminator, we added the oligo before washing the beads). We observed that only 25% of the released RNA could be cleaved by RNaseH (lane 7), which indicated that the antisense oligo couldn't efficiently invade the preformed hairpin of the released RNA. In presence of N, we added the oligo after extensive washing of the beads to probe whether the oligo can access the upstream part of the stem of the hairpin associated with the N-modified stalled EC. About 50% (lane 14) of the hairpin RNA could be cleaved by RNaseH under this condition. This suggests that the antisense oligo can invade the hairpin more easily when it is associated with the N-modified EC. It should also be noted that, this efficiency (52%, lane 14) of RNaseH cleavage is still significantly lower than what observed (80%, lane 10) when oligo was added before the hairpin was formed which suggests that the state of the hairpin in the N-modified EC is not fully melted. Through these experiments, we present here a gross picture of the conformational state of the terminator hairpin associated with the N-modified EC. We concluded that even though a part of the stem of the hairpin may have formed, it might be unstable enough to be invaded by the antisense oligo. Therefore, in presence of H-19B N, it is likely that proper formation of terminator hairpin is affected and this in turn contributes to the stability of the EC at the terminator (as was shown in Figure 5.7).

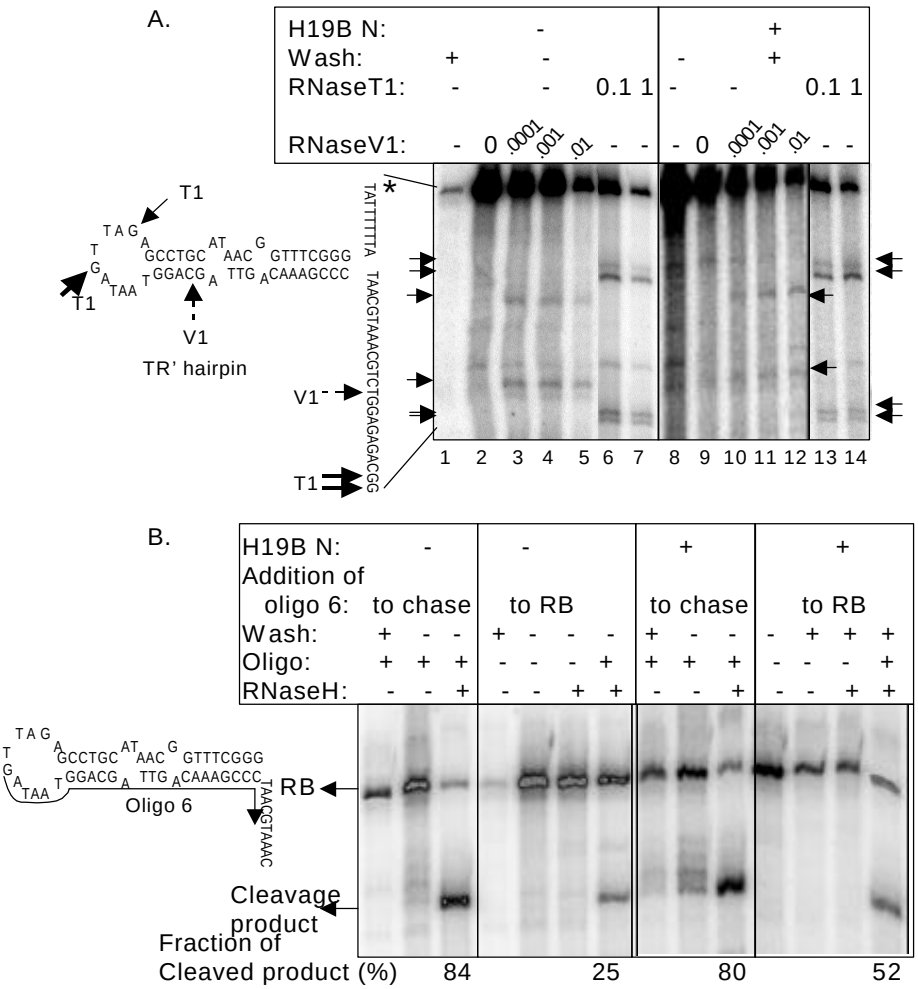


Figure 5.8: State of the terminator hairpin of the stalled EC

A. Stalled complexes at $T_{R'}$ terminator were formed in absence and presence of H 19B N in a similar way as described in Figure 5.5 A. RNaseT1 (lanes 6, 7, 13 and 14) and V1 (lanes 3 to 5 and 10 to 12) were added to the complexes before washing the beads when N is not present and added after washing when N is present. Amount of the RNases in units are indicated above the lanes. Cleavage products are shown by arrows. Cleavage positions are indicated on the sequence of $T_{R'}$ terminator. Thickness of the arrows on the sequence indicates the extent of sensitivity of that position to the respective RNases. Cleavage positions were assigned from parallel sequencing lanes. **B.** Oligo6, antisense to the upstream part of the stem of the hairpin (shown in the cartoon) was added either to the reaction during chasing the 22-mer EC ("to chase") or after the formation of RB complexes ("to RB"). In the "to RB" lanes oligo was added prior to washing the RB complexes when N was absent and added after washing when N was present. In the +RNaseH lanes 1 unit of the enzyme were added following the addition of oligo6. Concentration of oligo 6 was 15 μM. Fraction of cleaved product was calculated as: (Intensity of cleaved product)/([Intensity of cleaved product] + [Intensity of RB]). In both these assays NusA and NusG were present in all the reactions. Concentrations of different factors were same as Figure 5.1.

5.3.4 N prevents reversible backtracking at *ops* pause sequences

The antiterminators N and Q can suppress pausing during elongation (Ko et al., 1998; Nudler et al., 1998; Rees et al., 1993). Results from protein footprinting experiments (Figures 4.8, 4.9, 4.11 and 4.13) and from earlier mutational analyses (Chapter III; Cheeran et al., 2005) suggest that H-19B N may modulate interactions around the RNA:DNA hybrid as well as in the RNA exit channel. Therefore, we assayed the effect of H-19B N on two well-defined pause sequences, namely *ops* and *his* pauses, which involves altered interactions in these two regions. Pausing at *his* pause sequence is mediated by interaction of a RNA-hairpin with the flap domain of β-subunit located near the RNA exit channel (Artsimovitch and Landick, 1998; Artsimovitch and Landick, 2000; Chan et al., 1997; Toulokhonov et al., 2001; Toulokhonov and Landick, 2003). The sequence at the *his* pause codes for RNA that folds into a hairpin in the exit channel. At the *ops* pause sequence, pausing occurs due to the backtracking of the EC, which is dependent on the sequence of the RNA:DNA hybrid (Artsimovitch and Landick, 2000).

We followed the pausing kinetics through the *ops* pause sequence cloned downstream of the *nutR* site, both in the absence and presence of either WT or mutant H-19B N proteins (Figure 5.9 A). The amount of paused complex was plotted against time (Figure 5.9 B). In the presence of WT N, pausing efficiency (denoted as "a" in the equation of exponential decay; shown in the inset of Figure 5.9 B) was not affected, where as the rate of escape (denoted as "λ" in the equation) from the paused state was three times faster. This result suggests that N does not prevent the EC from entering the paused state, but reduces the half-life of this conformational state most likely by preventing the backtracking of EC that is an important component of this type of pausing (Artsimovitch and Landick, 2000). In order to ascertain that N prevents backtracking at this sequence, we then stalled the EC at this site by using Lac repressor as a road block (Figure 5.9 C). Backtracking of the stalled EC was monitored by GreB-induced cleavage of the RNA (Borukhov et al., 1993; Komissarova and Kashlev, 1997). In the presence of WT N, the sensitivity for GreB was diminished significantly (Figure 5.9 D). We ascertained this by Fe-DTT cleavage (Nudler et al., 1997) of the backtracked RNA in the presence and absence of H-19B N in the stalled RB complex in the same experimental set up, as explained earlier for GreB cleavage. In the absence of N, we observed that Fe-DTT mediated cleavage occurred at an internal position (3 nt upstream from the 3'-end and marked as "-3" in the Figure 5.9 E) which is a signature of backtracking. This cleavage was not evident in the presence of N (Figure 5.9 E).

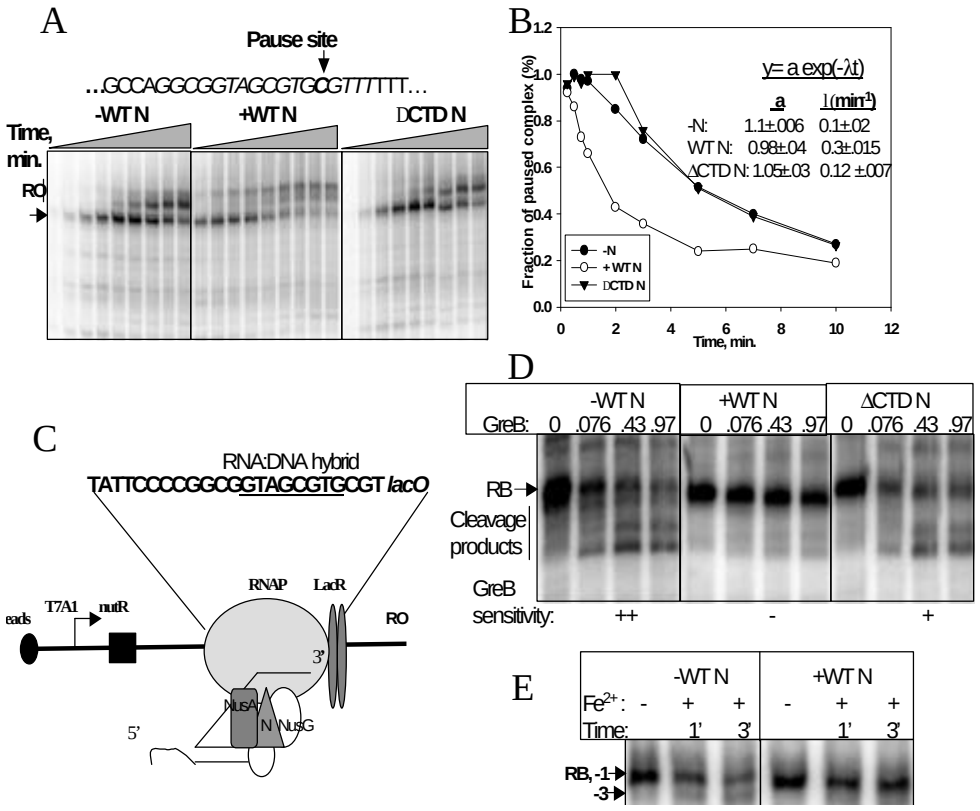


Figure 5.9: Anti-pausing activity at *ops* pause sequence

A. **Autoradiogram showing the time-course of transcription elongation through the *ops* pause sequence (indicated by arrow) both in the absence and presence of either WT or** Δ**CTD H-19B N.** The pause sequence and pausing site are shown above the autoradiogram. Run-off product is indicated as "RO". **B. Amount of fraction of paused complex obtained under different conditions is plotted against time.** The plots were fitted to the equation of exponential decay using SIGMAPLOT to calculate the rate (λ) of escape from the paused state and efficiency (a) of pausing. The equation and average values obtained from these calculations are shown in the inset of the Figure. **C. Cartoon showing the N-modified elongation complex stalled at *ops* using Lac repressor as a roadblock**. At this position the RNA:DNA hybrid contains the *ops* pause sequence. This stalled EC is used to probe backtracking at this sequence by GreB sensitivity. **D. Autoradiogram showing the GreB-induced cleavage of the RNA in the stalled elongation complex at the *ops* pause site.** Different concentrations of GreB were added to the stalled ECs (RB) either unmodified or modified with WT or ΔCTD H-19B N, and incubated for 3 min Cleavage products are indicated. The sensitivity towards GreB is indicated by "++" (highly sensitive), "+" (moderately sensitive) and "-"(not sensitive). Concentrations of different components are as follows: 20 nM RNAP, 5 nM DNA template, 100 nM LacI, 200 nM H-19B N, 300 nM NusA, 200 nM NusG and 100 μM NTPs. Concentrations of GreB are indicated. **E. Cleavage of 3'-end of the RNA in the RNA:DNA hybrid by Fe^{2+}.** A cleavage product 3nt upstream of the 3'-end is visible in the absence of N and is marked. Time of the reaction is indicated. Concentrations of different components are same as above.

5.3.5 N destabilizes the flap domain-RNA hairpin interactions at *his* pause sequence

Next we followed the pausing kinetics through the *his* pause site cloned downstream of the *nutR* sequence (see Materials methods transcription protocol; Figure 5.10 A). The amount of paused complex was plotted as described earlier for quantitating *ops* pause (Figure 5.9 B). As in the case of the *ops* pause, the pausing efficiency at the *his* pause did not change significantly in the presence of WT N, whereas the rate of escape from the paused state was increased by about 5-fold (see the inset of Figure 5.10 B for the rate constant values). The effect of N on the *his* pause was more pronounced than at the *ops* pause. The dwelling time in this paused state depends on the stability of the flap domain-RNA hairpin interaction (Toulokhonov and Landick, 2003). As binding of N only affected the pause half-life, it is likely that it did not affect the formation of the pause hairpin, but rather weakened the flap-hairpin interaction. It has been shown that the flap-hairpin interaction at this pause site leads to catalytic inactivation of the EC (Toulokhonov and Landick, 2003). Therefore, if the presence of H-19B N destabilizes this interaction it will also prevent the catalytic inactivation caused by the RNA hairpin. To test this hypothesis, we stalled the EC by Lac repressor at the pause site, removed the NTPs by washing, and tested the catalytic competence of the stalled EC by the pyrophosphorolysis reaction (Figure 5.10 C). Figure 5.10 D shows that the 3'-end of the RNA of this stalled EC exactly matched the pause site (compare the lanes 3 and 4, 10 and 11, 17 and 18). Formation of RNA–cleavage products (marked as "PPi cleavage") induced by pyrophosphate was only observed in the presence of WT H-19B N. This effect was specific to WT N as it was not observed either in the absence of N or in the presence of ΔCTD N, which does not bind to the EC. We conclude that the flap domain–RNA interaction is destabilized or significantly weakened in the N-modified EC and thereby preventing catalytic inactivation.

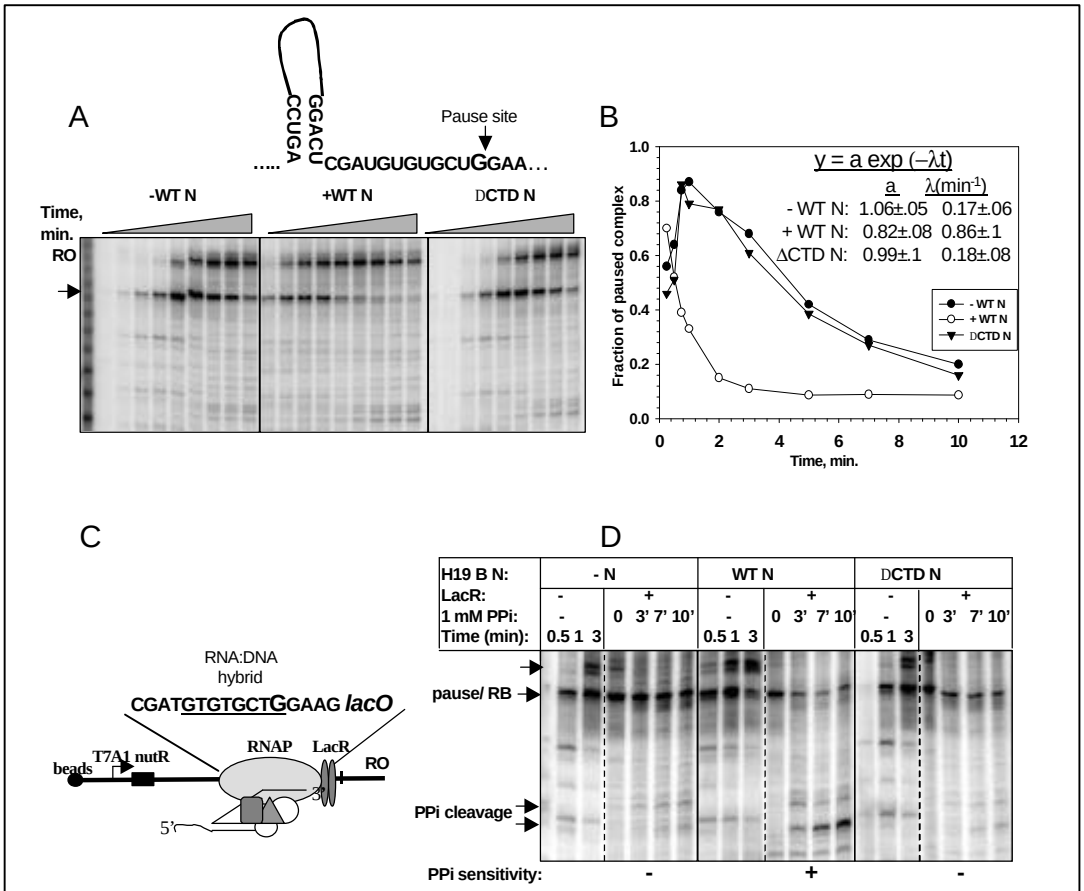


Figure 5.10: Anti-pausing activity at *his* pause sequence

A. **Autoradiogram showing the time-course of transcription elongation through the *his* pause sequence (indicated by arrow) both in the absence and presence of either WT or** Δ**CTD H-19B N.** Pause sequence, pausing site and pause-hairpin are shown above the autoradiogram. Run-off product is indicated as "RO".**B. Amount of fraction of paused complex obtained under different conditions is plotted against time.** The plots were fitted to the equation of exponential decay to calculate the rate (?) of escape from the paused state and efficiency (a) of pausing. The equation and average values obtained from these calculations are shown in the inset of the Figure. **C. Cartoon showing the N-modified elongation complex stalled at *his* pause site using Lac repressor as a roadblock**. At this position the RNA:DNA hybrid contains the *his* pause sequence and the 3'-end matches the pausing position (marked by an arrow). This set-up was used to probe the catalytic competence of the N-modified and un-modified ECs at the *his* pause. **D. Autoradiogram showing the RNA cleavage from pyrophosphorylosis reactions in the presence of Na-pyrophosphate from N-modified and un-modified stalled ECs at the *his* pause site.** Time courses of elongation through this pause site (lanes 1-3, 8-10 and 15-17) are shown along with each stalled complexes to compare the pausing and stalling positions. Pyrophosphorylosis reactions were performed by adding 1 mM Na pyrophosphate (PPi) to each of the stalled complexes, and incubated for indicated times. The cleavage products are indicated. Sensitivity towards PPi is marked as either "+" (sensitive) or "-"(not sensitive). Concentrations of different components are as in Figure 5.9, except that NTPs were present at 150 μM.

5.3.6 N can overcome an inherent pause sequence present after *boxB* in H-19B

5.3.6.1 Identification of the location of the H-19B inherent pause sequence

We noticed that during pause assays using the H-19 B templates having the T7AI- *nut R* sequence showed a pausing pattern, at a point just after the nutR region. We identified the pause region by RNA sequencing and running the sequence products side by side with a kinetic assay of the pause region (Figure 5.11) We identified that the pause position as 220 bp upstream of the T7 promoter and 43 bp ahead of *nutR* BOX B.

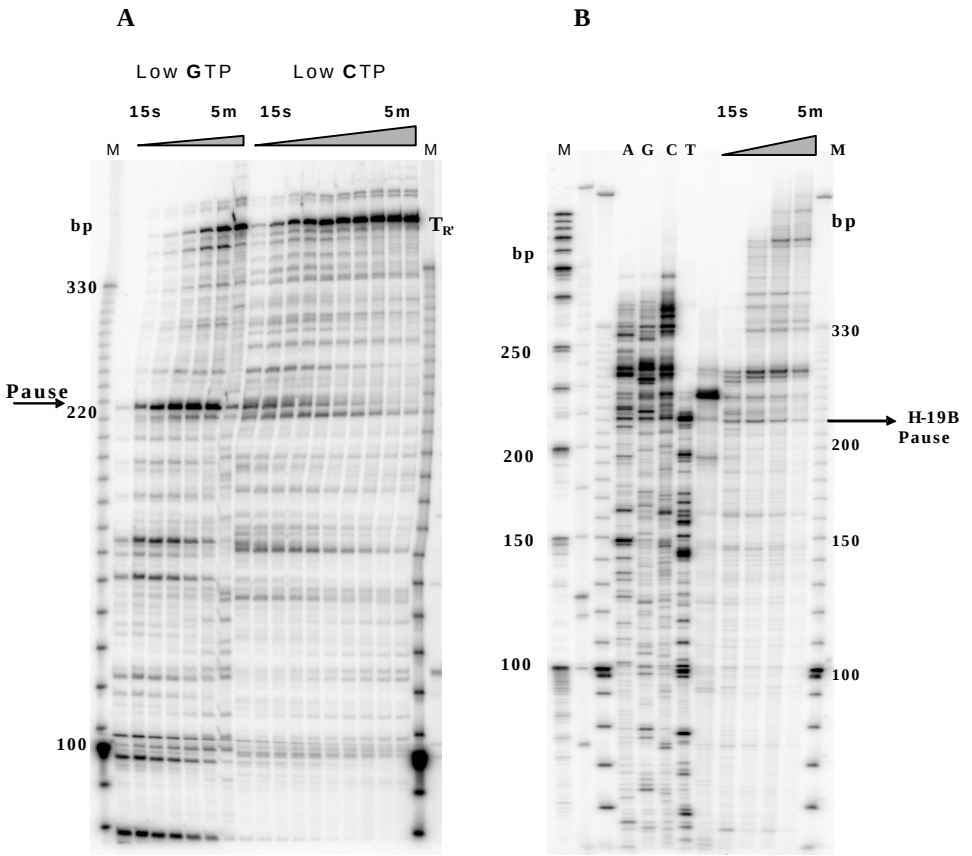


Figure 5.11 RNA sequencing and identification of H-19B Internal pause
A. Characterization of the inherent pause site. M stands for 10 bp DNA marker (Invitrogen); Pause kinetics with low amount of GTP and CTP as indicated (10 µM while all the other three NTPs were of 100 µM concentrations) in the chase solutions was carried out. **B. RNA sequencing of the Pause site, along with pause kinetics**. M- 50 bp (MBI), 25 bp and 10 bp (Invitrogen) DNA ladders are shown, in that order. RNA sequencing (see Materials methods) lanes corresponding to A, G, C and T; are also shown, along with an RB upstream to the pause site; Pause kinetics with low C concentration was carried out as in Figure 5.9 A.

5.3.6.2 Characterization of the H-19B inherent pause sequence by antipausing assays

We observed that H-19 B N modified EC has a very marked effect on the pausing at this position. N reduces the pausing efficiency at this internal pause site (Figure 5.12 B).

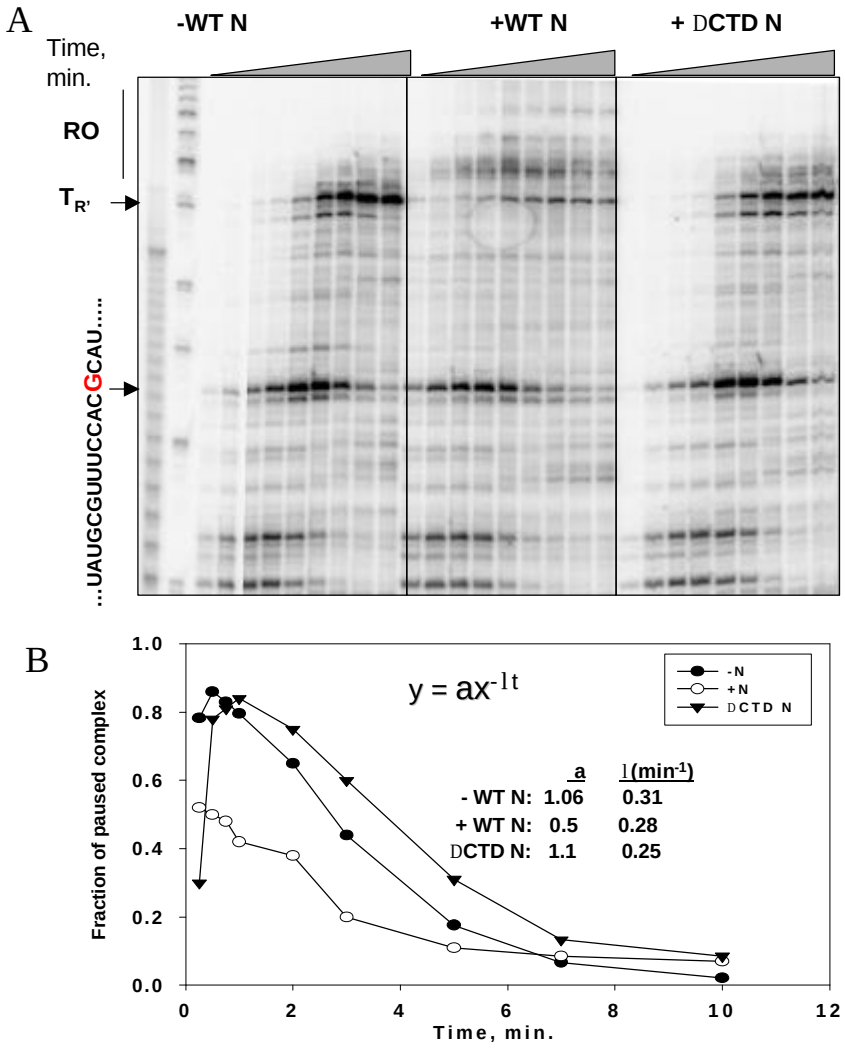


Figure 5.12: Anti-pausing activity at *H-19B* pause sequence

A. **Autoradiogram showing the time-course of transcription elongation through the *ops* pause.** The *ops* pause sequence is indicated by arrow. Time course both in the absence and presence of either WT or ΔCTD H-19B N. Time course was from 30s and 1 to 10 min at 1 min intervals. The pause sequence and pausing site (G, marked in bold color) are shown. Run-off product is indicated as "RO".

B. Amount of fraction of paused complex obtained under different conditions is plotted against time. The plots were fitted to the equation of exponential decay using SIGMAPLOT to calculate the rate (λ) of escape from the paused state and efficiency (a) of pausing. The equation and average values obtained from these calculations are shown in the inset of the figure.

From the analysis of the pause sequence, it was inferred that this sequence has a very low chance of forming a hairpin and hence this pause could be considered as a ClassII or hairpin independent pause. Since we had used another well characterized hairpin independent pause (*ops* pause Artsimovitch et al., 2000) for assaying the effects of N modification on Class II pauses, we did not attempt to further characterize this internal pause; but the insight helped us in designing the DNA templates required for *ops* and *his* pause assays in such a way that we could avoid the occurrence of this pause site in between those required pause sequences and the *nutR* upstream. The templates were so designed that the *his* or *ops* pause sequence was placed upstream to this internal pause sequence.

5.4 Discussion

What is the mechanism of H-19B N-mediated antitermination? Any transcription antitermination mechanism should involve one or all of the following.

1) An antiterminator may prevent or delay the terminator hairpin folding in the RNA exit channel.
2) It may stabilize weak interactions in the weak U-rich RNA:DNA hybrid.
3) It may alter the interactions of the clamp domain of RNAP with the downstream duplex DNA.

We have shown that the H-19B N modified EC is significantly stable even when stalled at a terminator (Figure 5.6, 5.7) and it appears that folding of the terminator hairpin in this modified EC is also affected (Figure 5.8). The altered interactions at the beginning of the RNA exit channel and upstream part of RNA:DNA hybrid can directly prevent the formation of the base of the hairpin. Alternatively, these altered interactions in the modified EC can stabilize the RNA:DNA hybrid which will in turn prevent the completion of the hairpin folding as the melting of the hybrid is essential for hairpin to form (Komissarova et al., 2002). The concept that H-19B N affects the proper formation of terminator hairpin is consistent with earlier proposal of prevention or delayed folding of hairpin in presence of λ N (Gusarov and Nudler, 2001). In addition to the effect on the hairpin formation, it is also possible that altered interactions surrounding the RNA:DNA hybrid could have an allosteric effect on the clamp domain that holds the downstream duplex DNA (Gnatt et al., 2001).

We have also shown that the interaction of H-19B N with the EC blocked the backtracking of the EC at a class II pause site (Artsimovitch and Landick, 2000; Figure 5.9) and prevented the catalytic inactivation caused by the RNA hairpin-flap domain interaction in the RNA exit channel at a class I pause site (Toulokhnov and Landick, 2003; Figure 5.10). The region of EC physically

proximal to the C-terminal domain of H-19B N is close to and easily accessible from the secondary channel (Figure 4.12 B, Table 5), but is quite distant to the RNA exit channel (~50 Å). However, lack of cleavage of C-terminal of N by Fe^{2+} from active site (Figure 5.2) and inability of N to suppress the effect of a secondary channel-binding antibiotic Tagetoxin (Figures 5.3, 5.4 A) suggest H-19B N may not use the secondary channel. Footprinting results and the locations of H-19B N-specific mutations obtained earlier (Figures 4.12 A, Cheeran et al., 2005) strongly suggest that the C-terminal of N is likely to be inserted through this RNA exit channel. In order to cover this long distance, the C-terminal of N should be in an extended structure, which will be stabilized by interacting with the interior of the EC.

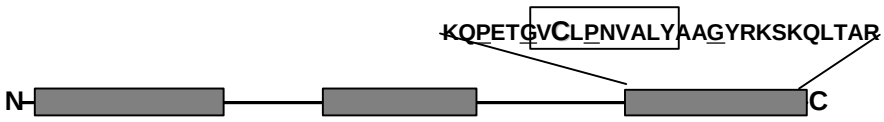


Figure 5.13: Cartoon representing H-19 B N protein
N and C stand for N terminus and C terminus of H-19B N. Amino acids at the C terminus of H-19B N is shown. The hydrophobic patch is highlighted in the box, with the C107 where Fe-BABE was linked in bold.

Due to the presence of Proline and Glycine residues (Figure 5.13, underlined a. a) in this region, it is not likely that H-19B N will able to form a long uninterrupted helix. Also, the presence of a hydrophobic patch in the C-terminal encompassing the Cysteine residue (Figure 5.13, sequences marked with a rectangle) will help this region to be inserted into the core of the EC. Alternatively, the presence of the N-NusA complex at the floor of the RNA exit channel, assuming that NusA in the NusA-N complex still binds to the N-terminal coiled-coil domain of β'-subunit (Borukhov et al., 2005), can widen the exit channel so that the regions close to the 3'-half of the RNA:DNA hybrid become more accessible to the free C-terminal end of N. It is likely that the presence of NusA at the exit channel guides the C-terminal domain of N into the EC by anchoring the central part of N onto its dimeric interface (Bonin et al., 2004). The presence of NusA complexed to the central part of N near the exit channel can physically block the formation of the RNA hairpin in the vicinity. And the role of C-terminal of N, which comes close to the active site, may be restricted to holding the 3'-end of the RNA in the weak RNA:DNA hybrid at the active site. Based on the protein footprinting data and the analysis of N-modified ECs stalled at class I and class II pause sites, we propose that the physical proximity of the C-terminal domain of N near the active site Mg^{2+} directly stabilizes the 3'-end of the weak U-rich

RNA:DNA hybrid and allosterically destabilizes the RNA hairpin interactions in the distal RNA exit channel, thus bringing about a highly procesive antitermination proficient EC, which reads through the termination signals and can transcribe kilo bases of RNA downstream to the transcription terminators.

CHAPTER VI

Conclusions and Future perspectives

6.1 Conclusions

The exact molecular mechanism by which N protein brings about transcription antitermination remains unknown, despite efforts to this end for the last three decades or so. The objective of this study was to get a better understanding of the interaction site of N protein on the EC and to understand more about the molecular mechanisms of N protein mediated transcription antitermination. The stability of an elongating complex can be attributed to the interactions between RNAP and the RNA-DNA hybrid. Intrinsic terminators destabilize the RNA-DNA hybrid, causes a pausing of RNAP at the terminator and along with the destabilization of the protein-nucleic acid interactions contributes to transcription termination. Antiterminators like N in ? and in related lambdoid phages convert RNA polymerase into a termination resistant form, during early phases of transcription elongation. N protein binds to RNA polymerase and together with host factors called Nus factors, transforms it into a highly processive, termination resistant transcription apparatus. We used the transcription antitermination system from a lambdoid bacteriophage H-19B, due to its reduced requirement of Nus factors for processive antitermination, which made it more suitable for *in vivo* and *in vitro* studies. The determination of the binding surface of N on RNAP is essential to understand the molecular mechanism of transcription antitermination. We addressed this by generating RNAP mutants, specifically defective for N mediated transcription antitermination and localized these substitutions in the structural model of EC. Attempts were also made to obtain suppressors of H-19B N for these RNAP mutants. Biochemical studies were carried out to determine the N binding surface by the footprinting of an N modified EC, using an array of biochemical tools. Experiments were also designed to see if N comes close to the active center of the RNAP or not and to analyze the subtle changes if any that it may bring about there. Taking into account the understanding that the terminator hairpin brings about transcription termination by destabilizing the RNA:DNA hybrid and protein nucleic acid contacts in the EC, we looked into the effects that N modulation of EC on the stability of an elongation complex, stalled at a terminator. We also addressed the state of RNA hairpin in an N modified EC, to check if N modification disrupts hairpin formation or not.

We were able to obtain definitive insights into most of these questions that we addressed.

1. As discussed in Chapter III, RNAP ß and ß’ mutants (ß G1045D, ß’ P251S, P254L, P251SP254L double, R270C and G336S) specifically defective for H-19B N mediated transcription antitermination both *in vivo* and *in vitro* were isolated. We also obtained a suppressor (L108F) in H-19B N for one of the RNAP mutant (ß’ P254L). In the presence

of this suppressor N protein, antitermination efficiency of ß' P254L mutant RNAP increased to the level of antitermination efficiency of Wt RNAP obtained in the presence of Wt N protein. This mutant N protein could also partially suppress the antitermination defects of P251S, P251S P254L double mutant and G336S *rpoC* mutants, where as it had no significant effects on R270C *rpoC* and G1045D *rpoB* mutant. Specific defects for N-mediated antitermination, increased concentration dependence for N, lack of processive antitermination of the mutant RNAPs; and the suppressor mutation in the RNAP binding domain of N together suggest that H-19B N exerts its effect on EC through the regions defined by the altered amino acids. These mutants define a very crucial region of RNAP, which are involved in the interactions important for the stabilization of the EC (Darst, 2001; Korzheva et al., 2000).

2. From the footprinting results as discussed in Chapter IV, it could be inferred that the C-terminal domain (the RNAP binding domain) of N from a lambdoid phage H-19B comes close to the active center of the EC. The cleavage pattern derived from Fe-BABE conjugated N bound to a stalled EC showed that the interaction surface must be close to the active site Mg^{2+} which was further corroborated by the N-induced localized conformation changes nearby to this region of EC. The lack of cleavage of C-terminal of N by Fe^{2+} from active site and inability of N to suppress the effect of a secondary channel-binding antibiotic Tagetoxin suggest H-19B N may not use the secondary channel. Footprinting results, the finding that the functional binding of N requires it to be inserted into the EC through the RNA exit channel (discussed in Chapter V) and the locations of H-19B N-specific mutations obtained earlier (Cheeran et al., 2005) implies that the C-terminal of N is likely to be inserted through this RNA exit channel. .

3. Our results discussed in Chapter V showed that H-19B N modified EC is significantly stable even when stalled at a terminator and that folding of the terminator hairpin in this modified EC is affected. In the presence of H-19B N, it is likely that proper formation of terminator hairpin is impaired and this in turn contributes to the stability of the EC at the terminator. The altered interactions at the beginning of the RNA exit channel and upstream part of RNA:DNA hybrid can directly prevent the formation of the base of the hairpin. Alternatively, these altered interactions in the modified EC can stabilize the RNA:DNA hybrid that will also prevent the completion of the hairpin folding as the melting of the hybrid is essential for hairpin to form (Komissarova et al., 2002). The concept that H-19B N affects the proper formation of terminator hairpin is consistent

with earlier proposal of prevention or delayed folding of hairpin in presence of λ N (Gusarov and Nudler, 2001).

Taking into account of all these results, we propose that the C-terminal domain of H-19B N approaches the regions around the active site of RNAP, quite probably through the RNA exit channel and that this physical proximity near the active site Mg^{2+} could stabilize the 3'-end of the weak U-rich RNA:DNA hybrid and along with the allosteric effects that this interaction brings about, destabilizes the RNA hairpin interactions in the distal RNA exit channel, preventing the RNAP pause at the U tract preceding the termination step. This, along with the improper formation of the terminator hairpin brought on as an effect of H-19B N action helps the TEC to go into a processive antitermination mode, efficiently transcribing through downstream terminators.

6.2 Future perspectives

Though the study presented here has given an indication as to where it is that H-19B N protein bind in a processive EC and the process by which it can modulate RNAP and the EC in general to convert it into processive antitermination machinery which can read through terminators, it also brings up some interesting questions.

1. Though the mutants that we obtained were found to be specifically defective for N mediated antitermination, the effects that host factor NusA plays on them is unclear. Future work in this regard can involve generation of suppressors (if any) in *E coli* NusA for the antitermination defective RNAP mutants that we have obtained in this study.
2. Another question that arises out of this study is whether our results define a generalized interaction surface for all the N proteins or not. Sequence alignment of the C-terminal amino acids from different N proteins revealed that this region is quite conserved among the majority of the known N proteins (Figure 6.1), except that from f -21 and phage L. It will be interesting to know whether C-terminal of ? ?N also get inserted into the core of the EC following the same mechanism as H-19B N. More work is needed to identify the exact molecular mechanisms of this process. Future work in this aspect can include the replacement of the CTD of H-19B N (a. a 100-127) with the CTD of ? ?phage (a. a 73-107). Such an N protein can used to assay the *in vivo/in vitro* antitermination rates. Biochemical studies like Fe-BABE cleavage (? N too has a single Cysteine at its CTD) using the purified domain swapped N protein can be carried out for checking if the CTD of this protein comes near to the RNAP active site or not. Biophysical studies of the H-

19B N CTD, by labeling the Cysteine residue with a suitable fluorophore to monitor the interaction of it with RNAP by fluorescence spectroscopy can also be thought about.

3. The fact that the CTD is the major domain required for RNAP binding and processive antitermination as well as the observation that the CTD of many of the antiterminator N proteins are conserved, also brings up the interesting question as to if one can design a short antitermination peptide from the CTD region of N protein, which can bring about processive antitermination by itself.

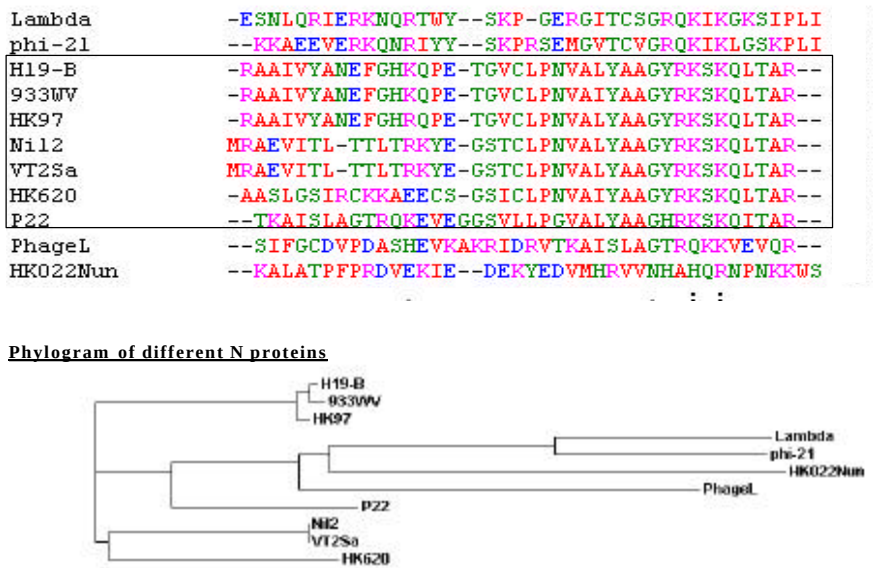


Figure 6.1: Clustal W analyses and phylogram of the N terminus of different N proteins from bacteriophages

Clustal W analyses of N proteins from different bacteriophages. 40 a. a from the CTD of each protein was subjected to ClustalW. The phylogram was obtained from ClustalW analysis. Phylogram of the various proteins was obtained from the ClustalW website, after analyses of the CTD.

4. It will be also interesting to see the exact molecular behavior of an N modified EC on factor dependent termination. *In vitro* assays involving an N modified EC on DNA templates having Rho utilization sites (*rut* sites) and Rho proteins can be carried out to assay processive antitermination through factor dependent terminators and stability of the EC at such terminators, in the presence of Rho protein.

CHAPTER VII

References

Aksoy, S., Squires, C. L., and Squires, C. (1984). Evidence for antitermination in Escherichia coli RRNA transcription. J Bacteriol *159*, 260-264.

Albrechtsen, B., Ross, B. M., Squires, C., and Squires, C. L. (1991). Transcriptional termination sequence at the end of the Escherichia coli ribosomal RNA G operon: complex terminators and antitermination. Nucleic Acids Res *19*, 1845-1852.

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*, 123-134.

Allison, L. A., Moyle, M., Shales, M., and Ingles, C. J. (1985). Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases. Cell *42*, 599-610.

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*, 3110-3122.

Artsimovitch, I., and Landick, R. (2000). Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. Proc Natl Acad Sci U S A *97*, 7090-7095.

Artsimovitch, I., and Landick, R. (2002). The transcriptional regulator RfaH stimulates RNA chain synthesis after recruitment to elongation complexes by the exposed nontemplate DNA strand. Cell *109*, 193-203.

Bai, L., Shundrovsky, A and Wang, M. D. (2004). Sequence-dependent kinetic model for transcription elongation by RNA polymerase. J Mol Biol *344*, 335-349.

Bailey, M. J., Hughes, C., and Koronakis, V. (1996). Increased distal gene transcription by the elongation factor RfaH, a specialized homologue of NusG. Mol Microbiol *22*, 729-737.

Bailey, M. J., Hughes, C., and Koronakis, V. (1997). RfaH and the ops element, components of a novel system controlling bacterial transcription elongation. Mol Microbiol *26*, 845-851.

Bailey, M. J., Hughes, C., and Koronakis, V. (2000). In vitro recruitment of the RfaH regulatory protein into a specialised transcription complex, directed by the nucleic acid ops element. Mol Gen Genet *262*, 1052-1059.

Banerjee, S., Chalissery, J., Bandey, I., and Sen, R. (2006). Rho-dependent transcription termination: more questions than answers. J Microbiol *44*, 11-22.

Banik-Maiti, S., King, R. A., and Weisberg, R. A. (1997). The antiterminator RNA of phage HK022. J Mol Biol *272*, 677-687.

Bar-Nahum, G., Epshtein, V., Ruckenstein, A. E., Rafikov, R., Mustaev, A., and Nudler, E. (2005). A ratchet mechanism of transcription elongation and its control. Cell *120*, 183-193.

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*, 885-899.

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*, 345-358.

Biggs, J., Searles, L. L., and Greenleaf, A. L. (1985). Structure of the eukaryotic transcription apparatus: features of the gene for the largest subunit of Drosophila RNA polymerase II. Cell *42*, 611-621.

Bonin, I., Muhlberger, R., Bourenkov, G. P., Huber, R., Bacher, A., Richter, G., and Wahl, M. C. (2004). Structural basis for the interaction of Escherichia coli NusA with protein N of phage lambda. Proc Natl Acad Sci U S A *101*, 13762-13767.

Borukhov, S., Laptenko, O., and Lee, J. (2001). Escherichia coli transcript cleavage factors GreA and GreB: functions and mechanisms of action. Methods Enzymol *342*, 64-76.

Borukhov, S., Lee, J., and Laptenko, O. (2005). Bacterial transcription elongation factors: new insights into molecular mechanism of action. Mol Microbiol *55*, 1315-1324.

Borukhov, S., Sagitov, V., and Goldfarb, A. (1993). Transcript cleavage factors from E. coli. Cell *72*, 459-466.

Borukhov, S., Severinov, K., Kashlev, M., Lebedev, A., Bass, I., Rowland, G. C., Lim, P. P., Glass, R. E., Nikiforov, V., and Goldfarb, A. (1991). Mapping of trypsin cleavage and antibody-

binding sites and delineation of a dispensable domain in the beta subunit of Escherichia coli RNA polymerase. J Biol Chem *266*, 23921-23926.

Burova, E., Hung, S. C., Chen, J., Court, D. L., Zhou, J. G., Mogilnitskiy, G., and Gottesman, M. E. (1999). Escherichia coli NusG mutations that block transcription termination by coliphage HK022 Nun protein. Mol Microbiol *31*, 1783-1793.

Busby, S., and Ebright, R.H. (1999). Transcription activation by catabolite activator protein (CAP). J Mol Biol. *22*, 199-213.

Bushnell, D. A., Westover, K. D., Davis, R. E., and Kornberg, R. D. (2004). Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Angstroms. Science *303*, 983-988.

Campbell, E. A., Korzheva, N., Mustaev, A., Murakami, K., Nair, S., Goldfarb, A., and Darst, S. A. (2001). Structural mechanism for rifampicin inhibition of bacterial rna polymerase. Cell *104*, 901-912.

Chamberlin, M. (1976). RNA Polymerase-an overview. In RNA Polymerase, R.Losick and M.Chamberlin, eds. (Cold Spring Harbor: Cold Spring Harbor Laboratory), pp-17-68.

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. (1995). Bipartite function of a small RNA hairpin in transcription antitermination in bacteriophage lambda. Proc Natl Acad Sci U S A *92*, 4061-4065.

Cheeran, A., Babu Suganthan, R., Swapna, G., Bandey, I., Achary, M. S., Nagarajaram, H. A., and Sen, R. (2005). Escherichia coli RNA polymerase mutations located near the upstream edge of an RNA:DNA hybrid and the beginning of the RNA-exit channel are defective for transcription antitermination by the N protein from lambdoid phage H-19B. J Mol Biol *352*, 28-43.

Conant, C. R., Van Gilst, M. R., Weitzel, S. E., Rees, W. A., and von Hippel, P. H. (2005). A quantitative description of the binding states and in vitro function of antitermination protein N of bacteriophage lambda. J Mol Biol *348*, 1039-1057.

Condon, C., Squires, C., and Squires, C. L. (1995). Control of rRNA transcription in Escherichia coli. Microbiol Rev *59*, 623-645.

Craig, M. L., Tsodikov, O. V., McQuade, K. L., Schlax, P. E., Jr., Capp, M. W., Saecker, R. M., and Record, M. T., Jr. (1998). DNA footprints of the two kinetically significant intermediates in formation of an RNA polymerase-promoter open complex: evidence that interactions with start site and downstream DNA induce sequential conformational changes in polymerase and DNA. J Mol Biol *283*, 741-756.

Cramer, P., Bushnell, D. A., Fu, J., Gnatt, A. L., Maier-Davis, B., Thompson, N. E., Burgess, R. R., Edwards, A. M., David, P. R., and Kornberg, R. D. (2000). Architecture of RNA polymerase II and implications for the transcription mechanism. Science *288*, 640-649.

d'Aubenton Carafa, Y., Brody, E., and Thermes, C. (1990). Prediction of rho-independent Escherichia coli transcription terminators. A statistical analysis of their RNA stem-loop structures. J Mol Biol *216*, 835-858.

deHaseth, P.L., Zupancic, M.L., and Record MT Jr. (1998). RNA polymerase-promoter interactions: the comings and goings of RNA polymerase. J Bacteriol. *180 (12)*, 3019-25.

Dalal, R. V., Larson, M. H., Neuman, K. C., Gelles, J., Landick, R., and Block, S. M. (2006). Pulling on the nascent RNA during transcription does not alter kinetics of elongation or ubiquitous pausing. Mol Cell *23*, 231-239.

Darst, S. A. (2001). Bacterial RNA polymerase. Curr Opin Struct Biol *11*, 155-162.

Das, A., Pal, M., Mena, J. G., Whalen, W., Wolska, K., Crossley, R., Rees, W., von Hippel, P. H., Costantino, N., Court, D., *et al.* (1996). Components of multiprotein-RNA complex that controls transcription elongation in Escherichia coli phage lambda. Methods Enzymol *274*, 374-402.

Datwyler, S. A., and Meares, C. F. (2000). Protein-protein interactions mapped by artificial proteases: where sigma factors bind to RNA polymerase. Trends Biochem Sci *25*, 408-414.

Davis, J. A., Takagi, Y., Kornberg, R. D., and Asturias, F. A. (2002). Structure of the yeast RNA polymerase II holoenzyme: Mediator conformation and polymerase interaction. Mol Cell *10*, 409-415.

Del Giudice, L., Manna, F., Massardo, D. R., Motto, M., Alifano, P., and Wolf, K. (1990). The Mu1 transposable element of maize contains two promoter signals recognized by the Escherichia coli RNA polymerase. Mol Gen Genet *222*, 71-76.

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*, 8660-8664.

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*, 5565-5577.

Downing, W., and Dennis, P. P. (1991). RNA polymerase activity may regulate transcription initiation and attenuation in the rplKAJLrpoBC operon in Escherichia coli. J Biol Chem *266*, 1304-1311.

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

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*, 945-951.

Fink, A.L. Natively unfolded proteins. (2005). Curr.Opin.Structl.Biol. *15*, 35-41.

Fish, R. N., and Kane, C. M. (2002). Promoting elongation with transcript cleavage stimulatory factors. Biochim Biophys Acta *1577*, 287-307.

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., and Doelling, J. H. (1989). Overexpression of N antitermination proteins of bacteriophages lambda, 21, and P22: loss of N protein specificity. J Bacteriol *171*, 2513-2522.

Friedman,D.I. (1971). A bacterial mutant affecting lambda development, p.733-738. *In* A.D.Hershey (ed.), The bacteriophage lambda.Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Friedman,D.I. (1992). Interaction between bacteriophage lambda and its *Escherichia coli* host. Curr Opin Genet Dev. *2 (5)*,727-38.

Friedman, D.I., and Gottesman, M. (1983). Lytic mode of lambda development. *In* Lambda II. Hendrix,R.W., Roberts, J.W., Stahl,F.W., and Weisberg, R.A.(eds) Cold Spring Harbor Laboratory Press, pp 21-51.

Friedman, D. I., and Court, D. L. (1995). Transcription antitermination: the lambda paradigm updated. Mol Microbiol *18*, 191-200.

Friedman, D. I., and Court, D. L. (2001). Bacteriophage lambda: alive and well and still doing its thing. Curr Opin Microbiol *4*, 201-207.

Friedman, D. I., Olson, E. R., Georgopoulos, C., Tilly, K., Herskowitz, I., and Banuett, F. (1984). Interactions of bacteriophage and host macromolecules in the growth of bacteriophage lambda. Microbiol Rev *48*, 299-325.

Georgopoulos, C. P. (1971). 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*, 2977-2981.

Ghysen, A., and Pironio, M. (1972). Relationship between the N function of bacteriophage lambda and host RNA polymerase. J Mol Biol *65*, 259-272.

Gill, S. C., Yager, T. D., and von Hippel, P. H. (1991). *Escherichia coli* sigma 70 and NusA proteins. II. Physical properties and self-association states. J Mol Biol *220*, 325-333.

Geiselmann, J. and von Hippel, P. H. 1992. Functional interactions of ligand cofactors with *Escherichia coli* transcription termination factor Rho. I. Binding of ATP. Protein Sci. *1*, 850-60.

Geiselmann, J., Seifried, S. E., Yager, T. D., Liang, C. and von Hippel, P. H. 1992b. Physical properties of the *Escherichia coli* transcription termination factor Rho. 2. Quaternary structure of the Rho hexamer. Biochemistry *31*, 121-32.

Gnatt, A. (2002). Elongation by RNA polymerase II: structure-function relationship. Biochim Biophys Acta *1577*, 175-190.

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

Gogol, E. P., Seifried, S. E. and von Hippel, P. H. 1991. Structure and assembly of the *Escherichia coli* transcription termination factor Rho and its interaction with RNA. I. Cryoelectron microscopic studies. J. Mol. Biol. *221*, 1127-38.

Gopal B, Haire LF, Gamblin SJ, Dodson EJ, Lane AN, Papavinasasundaram KG, Colston MJ, Dodson G. (2001). Crystal structure of the transcription elongation/anti-termination factor NusA from Mycobacterium tuberculosis at 1.7 A resolution. J Mol Biol. *14*, 314(5): 1087-95.

Gopal, V., Brieba, L. G., Guajardo, R., McAllister, W. T., and Sousa, R. (1999). Characterization of structural features important for T7 RNAP elongation complex stability reveals competing complex conformations and a role for the non-template strand in RNA displacement. J Mol Biol *290*, 411-431.

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*, 421-428.

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., Nodwell, J. R., and Mason, S. W. (1993). Transcriptional antitermination. Nature *364*, 401-406.

Gross, C. A., Chan, C., Dombroski, A., Gruber, T., Sharp, M., Tupy, J., and Young, B. (1998). The functional and regulatory roles of sigma factors in transcription. Cold Spring Harb Symp Quant Biol *63*, 141-155.

Gross, C. A., Chan, C. L., and Lonetto, M. A. (1996). A structure/function analysis of Escherichia coli RNA polymerase. Philos Trans R Soc Lond B Biol Sci *351*, 475-482.

Gruber, T. M., and Gross, C. A. (2003a). Assay of Escherichia coli RNA polymerase: sigma-core interactions. Methods Enzymol *370*, 206-212.

Gruber, T. M., and Gross, C. A. (2003b). Multiple sigma subunits and the partitioning of bacterial transcription space. Annu Rev Microbiol *57*, 441-466.

Guajardo, R. and Sousa, R.A. (1997). A model for the mechanism of polymerase translocation. J Mol Biol *265*, 8-19.

Gusarov, I., and Nudler, E. (2001). Control of intrinsic transcription termination by N and NusA: the basic mechanisms. Cell *107*, 437-449.

Harada, K., Martin, S. S., Tan, R., and Frankel, A. D. (1997). Molding a peptide into an RNA site by in vivo peptide evolution. Proc Natl Acad Sci U S A *94*, 11887-11892.

Head, S.C., Peitric, M., Richardson S., Roscoe, M., and Karmali M.A. (1988). Purification and characterization of Verocytotoxin 2. FEMS Microbiol Lett *51*, 211-216.

Heisler, L.M., Feng, G., Jin, D.J., Gross, C.A., Landick, R. (1996). Amino acid substitutions in the two largest subunits of Escherichia coli RNA polymerase that suppress a defective Rho termination factor affect different parts of the transcription complex. J Biol Chem.*271*(24):14572-83.

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

Heyduk, T., Heyduk, E., Severinov, K., Tang, H., and Ebright, R. H. (1996). Determinants of RNA polymerase alpha subunit for interaction with beta, beta', and sigma subunits: hydroxyl-radical protein footprinting. Proc Natl Acad Sci U S A *93*, 10162-10166.

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*, 631-641.

Huang, A., de Grandis, S., Friesen, J., Karmali, M., Petric, M., Congi, R., and Brunton, J. L. (1986). Cloning and expression of the genes specifying Shiga-like toxin production in Escherichia coli H19. J Bacteriol *166*, 375-379.

Huang, A., Friesen, J., and Brunton, J. L. (1987). Characterization of a bacteriophage that carries the genes for production of Shiga-like toxin 1 in Escherichia coli. J Bacteriol *169*, 4308-4312.

Hudson, G. S., Holton, T. A., Whitfield, P. R., and Bottomley, W. (1988). Spinach chloroplast rpoBC genes encode three subunits of the chloroplast RNA polymerase. J Mol Biol *200*, 639-654.

Ishihama, A. (2000). Functional modulation of Escherichia coli RNA polymerase. Annu Rev Microbiol *54*, 499-518.

Jin, D. J., Cashel, M., Friedman, D. I., Nakamura, Y., Walter, W. A., and Gross, C. A. (1988a). Effects of rifampicin resistant rpoB mutations on antitermination and interaction with nusA in Escherichia coli. J Mol Biol *204*, 247-261.

Jin, D. J., Walter, W. A., and Gross, C. A. (1988b). Characterization of the termination phenotypes of rifampicin-resistant mutants. J Mol Biol *202*, 245-253.

Juhala, R. J., Ford, M. E., Duda, R. L., Youlton, A., Hatfull, G. F., and Hendrix, R. W. (2000). Genomic sequences of bacteriophages HK97 and HK022: pervasive genetic mosaicism in the lambdoid bacteriophages. J Mol Biol *299*, 27-51.

Kandel, E. S., and Nudler, E. (2002). Template switching by RNA polymerase II in vivo. Evidence and implications from a retroviral system. Mol Cell *10*, 1495-1502.

Kashlev, M., and Komissarova, N. (2002). Transcription termination: primary intermediates and secondary adducts. J Biol Chem *277*, 14501-14508.

Kashlev, M., Nudler, E., Severinov, K., Borukhov, S., Komissarova, N., and Goldfarb, A. (1996). Histidine-tagged RNA polymerase of Escherichia coli and transcription in solid phase. Methods Enzymol *274*, 326-334.

Kato, M., Ito, T., Wagner, G., Richardson, C. C., and Ellenberger, T. (2003). Modular architecture of the bacteriophage T7 primase couples RNA primer synthesis to DNA synthesis. Mol Cell *11*, 1349-1360.

Kettenberger, H., Armache, K. J., and Cramer, P. (2003). Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. Cell *114*, 347-357.

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.

King, R. A., Markov, D., Sen, R., Severinov, K., and Weisberg, R. A. (2004). A conserved zinc binding domain in the largest subunit of DNA-dependent RNA polymerase modulates intrinsic transcription termination and antitermination but does not stabilize the elongation complex. J Mol Biol *342*, 1143-1154.

King, R. A., Sen, R., and Weisberg, R. A. (2003). Using a lac repressor roadblock to analyze the E. coli transcription elongation complex. Methods Enzymol *371*, 207-218.

Ko, D. C., Marr, M. T., Guo, J., and Roberts, J. W. (1998). A surface of Escherichia coli sigma 70 required for promoter function and antitermination by phage lambda Q protein. Genes Dev *12*, 3276-3285.

Komissarova, N., Becker, J., Solter, S., Kireeva, M., and Kashlev, M. (2002). Shortening of RNA:DNA hybrid in the elongation complex of RNA polymerase is a prerequisite for transcription termination. Mol Cell *10*, 1151-1162.

Komissarova, N., and Kashlev, M. (1997). 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*, 1755-1760.

Korzheva, N., Mustaev, A., Kozlov, M., Malhotra, A., Nikiforov, V., Goldfarb, A., and Darst, S. A. (2000). A structural model of transcription elongation. Science *289*, 619-625.

Landick, R. (2001). RNA polymerase clamps down. Cell *105*, 567-570.

Landick R, Wang D, Chan CL. (1996). Quantitative analysis of transcriptional pausing by Escherichia coli RNA polymerase: his leader pause site as paradigm. Methods Enzymol. *274*, 334-53.

Laptenko, O., Lee, J., Lomakin, I., and Borukhov, S. (2003). Transcript cleavage factors GreA and GreB act as transient catalytic components of RNA polymerase. EMBO J *22*, 6322-6334.

Laspia, M. F., Wendel, P., and Mathews, M. B. (1993). HIV-1 Tat overcomes inefficient transcriptional elongation in vitro. J Mol Biol *232*, 732-746.

Leeds, J. A., and Welch, R. A. (1996). RfaH enhances elongation of Escherichia coli hlyCABD mRNA. J Bacteriol *178*, 1850-1857.

Leffers, H., Gropp, F., Lottspeich, F., Zillig, W., and Garrett, R. A. (1989). Sequence, organization, transcription and evolution of RNA polymerase subunit genes from the archaebacterial extreme halophiles Halobacterium halobium and Halococcus morrhuae. J Mol Biol *206*, 1-17.

Lesnik, E. A., Sampath, R., Levene, H. B., Henderson, T. J., McNeil, J. A., and Ecker, D. J. (2001). Prediction of rho-independent transcriptional terminators in Escherichia coli. Nucleic Acids Res *29*, 3583-3594.

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*, 6012-6019.

Li, X., Hansen, P. A., Xi, L., Chandraratna, R. A., and Burant, C. F. (2005). Distinct mechanisms of glucose lowering by specific agonists for peroxisomal proliferator activated receptor gamma and retinoic acid X receptors. J Biol Chem *280*, 38317-38327.

Linderoth, N. A., and Calendar, R. L. (1991). The Psu protein of bacteriophage P4 is an antitermination factor for rho-dependent transcription termination. J Bacteriol *173*, 6722-6731.

Link, G. (1984). DNA sequence requirements for the accurate transcription of a protein-coding plastid gene in a plastid in vitro system from mustard (Sinapis alba L.). EMBO J *3*, 1697-1704.

Little, J,W. The SOS regulatory system (1995). *In* Regulation of gene expression in *Escherichia coli.* Lynn,E.C.C., and Lynch,A.S. (eds). Georgetown, TX: pp 453-479.

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*, 150-161.

Lorenzi, H. A., Robledo, G., and Levin, M. J. (2006). The VIPER elements of trypanosomes constitute a novel group of tyrosine recombinase-enconding retrotransposons. Mol Biochem Parasitol *145*, 184-194.

Lloyd, S. G., Niu, W., Tebbutt, J., Ebright, H. R., and Busby, S.W.J. (2003). Requirement for two copies of RNA polymerase a subunit C-terminal domain for synergistic transcription activation at complex bacterial promoters. Genes Dev *16*, 2557–2565.

Luzzati, D. (1970). Regulation of lambda exonuclease synthesis: role of the N gene product and lambda repressor. J Mol Biol *49*, 515-519.

Magyar, A., Zhang, X., Abdi, F., Kohn, H., and Widger, W. R. (1999). Identifying the bicyclomycin binding domain through biochemical analysis of antibiotic-resistant rho proteins. J Biol Chem *274*, 7316-7324.

Mah, T. F., Kuznedelov, K., Mushegian, A., Severinov, K., and Greenblatt, J. (2000). The alpha subunit of E. coli RNA polymerase activates RNA binding by NusA. Genes Dev *14*, 2664-2675.

Mah, T. F., Li, J., Davidson, A. R., and Greenblatt, J. (1999). Functional importance of regions in Escherichia coli elongation factor NusA that interact with RNA polymerase, the bacteriophage lambda N protein and RNA. Mol Microbiol *34*, 523-537.

Marr, M. T., and Roberts, J. W. (2000). Function of transcription cleavage factors GreA and GreB at a regulatory pause site. Mol Cell *6*, 1275-1285.

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*, 1504-1512.

McDowell, J. C., Roberts, J. W., Jin, D. J., and Gross, C. (1994). Determination of intrinsic transcription termination efficiency by RNA polymerase elongation rate. Science *266*, 822-825.

Meares, C. F., Datwyler, S. A., Schmidt, B. D., Owens, J., and Ishihama, A. (2003). Principles and methods of affinity cleavage in studying transcription. Methods Enzymol *371*, 82-106.

Miller, J.H. (1992). A short course in bacterial genetics: A laboratory manual and handbook for *Escherichia coli* and related bacteria. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA.

Miller, J.H., Claeys, I.V., Kirschbaum, J.B., Nasi, S., Van den Elsacker, S., Molholt, B., *et al.* *(*1976). Altered RNA polymerases resulting from temperature-sensitive mutations in the rif

region of the *E. coli* chromosome, in *RNA Polymerase.* (ed, R. Losick, and M. Chamberlin). pp. 519–538. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Minakhin L, Bhagat S, Brunning A, Campbell EA, Darst SA, Ebright RH, Severinov K. (2001). Bacterial RNA polymerase subunit omega and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. Proc Natl Acad Sci U S A. *98(3)*: 892-7.

Mizuguchi, K., Deane, C. M., Blundell, T. L., Johnson, M. S., and Overington, J. P. (1998). JOY: protein sequence-structure representation and analysis. Bioinformatics *14*, 617-623.

Modrak, D. and Richardson, J. P. 1994. The RNA-binding domain of transcription termination factor Rho: isolation, characterization, and determination of sequence limits. Biochemistry *33*, 8292-9.

Mogridge, J., and Greenblatt, J. (1998). Specific binding of Escherichia coli ribosomal protein S1 to boxA transcriptional antiterminator RNA. J Bacteriol *180*, 2248-2252.

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*, 265-275.

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*, 2831-2845.

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*, 4143-4148.

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

Mooney, R. A., Darst, S. A., and Landick, R. (2005). Sigma and RNA polymerase: an on-again, off-again relationship? Mol Cell *20*, 335-345.

Murakami, K. S., and Darst, S. A. (2003). Bacterial RNA polymerases: the wholo story. Curr Opin Struct Biol *13*, 31-39.

Murakami, K. S., Masuda, S., Campbell, E. A., Muzzin, O., and Darst, S. A. (2002). Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. Science *296*, 1285-1290.

Mustaev, A., Kozlov, M., Markovtsov, V., Zaychikov, E., Denissova, L., and Goldfarb, A. (1997). Modular organization of the catalytic center of RNA polymerase. Proc Natl Acad Sci U S A *94*, 6641-6645.

Naryshkin, N., Revyakin, A., Kim, Y., Mekler, V., and Ebright, R. H. (2000). Structural organization of the RNA polymerase-promoter open complex. Cell *101*, 601-611.

Naryshkina, T., Kuznedelov, K., and Severinov, K. (2006). The role of the largest RNA polymerase subunit lid element in preventing the formation of extended RNA-DNA hybrid. J Mol Biol *361*, 634-643.

Nechaev, S., Yuzenkova, Y., Niedziela-Majka, A., Heyduk, T., and Severinov, K. (2002). A novel bacteriophage-encoded RNA polymerase binding protein inhibits transcription initiation and abolishes transcription termination by host RNA polymerase. J Mol Biol *320*, 11-22.

Neely, M. N., and Friedman, D. I. (1998a). Arrangement and functional identification of genes in the regulatory region of lambdoid phage H-19B, a carrier of a Shiga-like toxin. Gene *223*, 105-113.

Neely, M. N., and Friedman, D. I. (1998b). Functional and genetic analysis of regulatory regions of coliphage H-19B: location of shiga-like toxin and lysis genes suggest a role for phage functions in toxin release. Mol Microbiol *28*, 1255-1267.

Neely, M. N., and Friedman, D. I. (2000). N-mediated transcription antitermination in lambdoid phage H-19B is characterized by alternative NUT RNA structures and a reduced requirement for host factors. Mol Microbiol *38*, 1074-1085.

Nehrke KW, Platt T. (1994). A quaternary transcription termination complex. Reciprocal stabilization by Rho factor and NusG protein. J Mol Biol *243 (5)*: 830-9.

Neuman, K. C., Abbondanzieri, E. A., Landick, R., Gelles, J., and Block, S. M. (2003). Ubiquitous transcriptional pausing is independent of RNA polymerase backtracking. Cell *115*, 437-447.

Nickels, B. E., and Hochschild, A. (2004). Regulation of RNA polymerase through the secondary channel. Cell *118*, 281-284.

Nickels, B. E., Roberts, C. W., Sun, H., Roberts, J. W., and Hochschild, A. (2002). The sigma(70) subunit of RNA polymerase is contacted by the (lambda)Q antiterminator during early elongation. Mol Cell *10*, 611-622.

Nie, X. M., Xiao, B. Y., Li, X. L., Zhang, B. C., Li, W. F., Wang, R., Cao, L., and Li, G. Y. (2003). [Expression of a new cloned nitroreductase gene NOR1 and purification of expressed product]. Ai Zheng *22*, 136-139.

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*, 2141-2151.

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*, 261-268.

Nudler, E., Avetissova, E., Markovtsov, V., and Goldfarb, A. (1996). Transcription processivity: protein-DNA interactions holding together the elongation complex. Science *273*, 211-217.

Nudler, E., Gusarov, I., Avetissova, E., Kozlov, M., and Goldfarb, A. (1998). Spatial organization of transcription elongation complex in Escherichia coli. Science *281*, 424-428.

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.

Obuchowski, M., Wegrzyn, A., Szalewska-Palasz, A., Thomas, M. S., and Wegrzyn, G. (1997). An RNA polymerase alpha subunit mutant impairs N-dependent transcriptional antitermination in Escherichia coli. Mol Microbiol *23*, 211-222.

Olins, P. O., Erickson, B. D., and Burgess, R. R. (1983). Overproduction of Escherichia coli NusA protein. Gene *26*, 11-18.

Opalka, N., Chlenov, M., Chacon, P., Rice, W. J., Wriggers, W., and Darst, S. A. (2003). Structure and function of the transcription elongation factor GreB bound to bacterial RNA polymerase. Cell *114*, 335-345.

Palangat, M., and Landick, R. (2001). Roles of RNA:DNA hybrid stability, RNA structure, and active site conformation in pausing by human RNA polymerase II. J Mol Biol *311*, 265-282.

Pani, B., Banerjee, S., Chalissery, J., Abhishek, M., Malarini, L. R., Suganthan, R., and Sen, R. (2006). Mechanism of inhibition of Rho-dependent transcription termination by bacteriophage P4 protein Psu. J Biol Chem *281*(*36)*, 26491-500.

Pasman, Z., and von Hippel, P. H. (2000). Regulation of rho-dependent transcription termination by NusG is specific to the Escherichia coli elongation complex. Biochemistry *39*, 5573-5585.

Patel, D. D., and Pickup, D. J. (1989). The second-largest subunit of the poxvirus RNA polymerase is similar to the corresponding subunits of procaryotic and eucaryotic RNA polymerases. J Virol *63*, 1076-1086.

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*, 217-228.

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*, 91-93.

Platt, T. (1994). Rho and RNA: models for recognition and response. Mol Microbiol *11*, 983-990.

Polyakov, A., Severinova, E., and Darst, S. A. (1995). Three-dimensional structure of E. coli core RNA polymerase: promoter binding and elongation conformations of the enzyme. Cell *83*, 365-373.

Prasch, S., Schwarz, S., Eisenmann, A., Wohrl, B. M., Schweimer, K., and Rosch, P. (2006). Interaction of the intrinsically unstructured phage lambda N Protein with Escherichia coli NusA. Biochemistry *45*, 4542-4549.

Price, M.A., and Tullius, T.D. (1992). Using hydroxyl radical to probe DNA structure. Methods Enzymol., *212*, 194-219.

Quan, S., Zhang, N., French, S., and Squires, C. L. (2005). Transcriptional polarity in rRNA operons of Escherichia coli nusA and nusB mutant strains. J Bacteriol *187*, 1632-1638.

Record, M. Jr., W.Reznikoff, M., Craig, K., McQuade and P.Schlax (1996). *Escherichia coli* RNA Polymerase ($E\ s^{70}$) promoters and the kinetics of the steps of transcription initiation.p.792-821. F.C. Neirhardt., R. Curtiss III., J.L. Ingraham., E.C.C. Lin., K.B. Low, B. Magasanik., W.S.Reznikoff., M.Riley., M.Schaecher and H.E.Umbarger (ed), *In Escherichia coli* and *Salmonella*: cellular and molecular biology, 2end ed. ASM press. Washington. D.C.

Reeder, T. C., and Hawley, D. K. (1996). Promoter proximal sequences modulate RNA polymerase II elongation by a novel mechanism. Cell *87*, 767-777.

Rees, W. A., Yager, T. D., Korte, J., and von Hippel, P. H. (1993). Betaine can eliminate the base pair composition dependence of DNA melting. Biochemistry *32*, 137-144.

Richardson, J. P. (2002). Rho-dependent termination and ATPases in transcript termination. Biochim Biophys Acta *1577*, 251-260.

Richardson, J.P., and Greenblatt, J.(1996). Control of RNA chain elongation and termination, p.822-848. *In* F.C.Neidhardt, R.C. Curtiss, J.L. Ingraham., E.C.C. Lin., K.B. Low, B. Magasanik., W.S.Reznikoff., M.Riley., M.Schaecher and H.E.Umbarger (ed), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2end ed. ASM press.Washington.D.C.

Richardson, J.P., and Greenblatt, J.(1996). Control of RNA chain elongation and termination, p.822-848. *In* F.C.Neidhardt, R.C. Curtiss, J.L. Ingraham., E.C.C. Lin., K.B. Low, B. Magasanik., W.S.Reznikoff., M.Riley., M.Schaecher and H.E.Umbarger (ed), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2end ed. ASM press.Washington.D.C.

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*, 21597-21603.

Riddles, P, W., Blakeley, R.L. and Zerner, B. (1979) Elleman's reagent: 5,5' dithiobis (2-nitro benzoic acid)- a reexamination. Anal. Biochem. *94*, 75-81.

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*, 485-493.

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*, 483-492.

Roberts, J., and Park, J. S. (2004). Mfd, the bacterial transcription repair coupling factor: translocation, repair and termination. Curr Opin Microbiol *7*, 120-125.

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

Roberts, J. W. (1988). Phage lambda and the regulation of transcription termination. Cell *52*, 5-6.

Rosenberg, M., and Court, D. (1979). Regulatory sequences involved in the promotion and termination of RNA transcription. Annu Rev Genet *13*, 319-353.

Salstrom, J. S., and Szybalski, W. (1978). Coliphage lambdanutL-: 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.

Sambrook, J. and Russel, D.W (2001). *In* Molecular cloning, A laboratory manual. Third edition; Cold Spring Harbor Laboratory Press. Cold Spring Harbor, New York.

Santangelo, T. J., Mooney, R. A., Landick, R., and Roberts, J. W. (2003). RNA polymerase mutations that impair conversion to a termination-resistant complex by Q antiterminator proteins. Genes Dev *17*, 1281-1292.

Santangelo, T. J., and Roberts, J. W. (2004). Forward translocation is the natural pathway of RNA release at an intrinsic terminator. Mol Cell *14*, 117-126.

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 Microbiol *21*, 839-851.

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*, 809-818.

Scotland, S.M., Smith, H.R., Willshaw, G.A., and Rowe, B. (1983). Vero cytotoxin production in strain *Escherichia coli* is determined by genes carried on bacteriophages. Lancet *2*, 216.

Selby, C. P., and Sancar, A. (1995). Structure and function of transcription-repair coupling factor. I. Structural domains and binding properties. J Biol Chem *270*, 4882-4889.

Sen, R., King, R. A., Mzhavia, N., Madsen, P. L., and Weisberg, R. A. (2002). Sequence-specific interaction of nascent antiterminator RNA with the zinc-finger motif of Escherichia coli RNA polymerase. Mol Microbiol *46*, 215-222.

Sen, R., King, R. A., and Weisberg, R. A. (2001). Modification of the properties of elongating RNA polymerase by persistent association with nascent antiterminator RNA. Mol Cell *7*, 993-1001.

Sharp, J. A., Guterman, S. K., and Platt, T. (1986). The rho-115 mutation in transcription termination factor rho affects its primary polynucleotide binding site. J Biol Chem *261*, 2524-2528.

Shin, D. H., Nguyen, H. H., Jancarik, J., Yokota, H., Kim, R., and Kim, S. H. (2003). Crystal structure of NusA from Thermotoga maritima and functional implication of the N-terminal domain. Biochemistry *42*, 13429-13437.

Sidorenkov, I., Komissarova, N., and Kashlev, M. (1998). Crucial role of the RNA:DNA hybrid in the processivity of transcription. Mol Cell *2*, 55-64.

Skordalakes, E. and Berger, J. M. 2003. Structure of the rho transcription terminator: mechanism of mRNA recognition and helicase loading. Cell *114*, 135-46.

Skordalakes, E. and Berger, J.M. (2006). Structural insights into RNA-dependent ring closure and ATPase activation by the Rho termination factor. Cell *127(3)*, 553-64.
Simons, R.W., Houman, F. and Kleckner, N. (1987). Improved single and multicopy *lac*- based cloning vectors for protein and operon fusions. Gene *53*, 85-96.

Smith, H.W., and Linggood, M.A. (1971). The transmissible nature of enterotoxin production in a human enteropathogenic strain of *Escherichia coli. J Med Microbiol 4*, 301–305.

Smith, H.R., Day,N.P., Scotland,S.M., Gross,R.J., and Rowe,B.(1984). Phage determined production of vero ctytoxin in strains of *Esherichia coli* serogroup 0157. Lancet *1*, 1242-1243.

Smith,H.W., Green,P. and Parsell, Z.(1983). Vero cell toxins in *Escherichia coli* and related bacteria:transfer by phage and conjugation and toxic action in laboratory animals, chickens and pigs. J Gen Microbiol *129*, 3121-3137.

Sosunova, E., Sosunov, V., Kozlov, M., Nikiforov, V., Goldfarb, A., and Mustaev, A. (2003). Donation of catalytic residues to RNA polymerase active center by transcription factor Gre. Proc Natl Acad Sci U S A *100*, 15469-15474.

Steiner T, Kaiser JT, Marinkovic S, Huber R, Wahl MC. (2002). Crystal structures of transcription factor NusG in light of its nucleic acid- and protein-binding activities. EMBO J. *21 (17)*: 4641-53.

Squires, C. L., and Zaporojets, D. (2000). Proteins shared by the transcription and translation machines. Annu Rev Microbiol *54*, 775-798.

Steege, D. A., Cone, K. C., Queen, C., and Rosenberg, M. (1987). Bacteriophage lambda N gene leader RNA. RNA processing and translational initiation signals. J Biol Chem *262*, 17651-17658.

Sternberg, N. (1976). A class of rifR RNA polymerase mutations that interferes with the expression of coliphage lambda late gene. Virology *73*, 139-154.

Stevens, M. P., Clarke, B. R., and Roberts, I. S. (1997). Regulation of the Escherichia coli K5 capsule gene cluster by transcription antitermination. Mol Microbiol *24*, 1001-1012.

Sullivan, S. L., and Gottesman, M. E. (1992). Requirement for E. coli NusG protein in factor-dependent transcription termination. Cell *68*, 989-994.

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*, 1339-1344.

Sweetser, D., Nonet, M., and Young, R. A. (1987). Prokaryotic and eukaryotic RNA polymerases have homologous core subunits. Proc Natl Acad Sci U S A *84*, 1192-1196.

Szalewska-Palasz, A., Strzelczyk, B., Herman-Antosiewicz, A., Wegrzyn, G., and Thomas, M. S. (2003). Genetic analysis of bacteriophage lambdaN-dependent antitermination suggests a possible role for the RNA polymerase alpha subunit in facilitating specific functions of NusA and NusE. Arch Microbiol *180*, 161-168.

Temiakov, D., Zenkin, N., Vassylyeva, M. N., Perederina, A., Tahirov, T. H., Kashkina, E., Savkina, M., Zorov, S., Nikiforov, V., Igarashi, N., *et al.* (2005). Structural basis of transcription inhibition by antibiotic streptolydigin. Mol Cell *19*, 655-666.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res *25*, 4876-4882.

Toulokhonov, I., Artsimovitch, I., and Landick, R. (2001). Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins. Science *292*, 730-733.

Toulokhonov, I., and Landick, R. (2003). The flap domain is required for pause RNA hairpin inhibition of catalysis by RNA polymerase and can modulate intrinsic termination. Mol Cell *12*, 1125-1136.

Toulokhonov, I., and Landick, R. (2006). The role of the lid element in transcription by E. coli RNA polymerase. J Mol Biol *361*, 644-658.

Travers, A.A., and Burgess, R.R. (1969). Cyclic re-use of the RNA polymerase sigma factor. Nature *222*, 537-540.

Tullius, T. D., and Greenbaum, J. A. (2005). Mapping nucleic acid structure by hydroxyl radical cleavage. Curr Opin Chem Biol *9*, 127-134.

Tuske, S., Sarafianos, S. G., Wang, X., Hudson, B., Sineva, E., Mukhopadhyay, J., Birktoft, J. J., Leroy, O., Ismail, S., Clark, A. D., Jr., *et al.* (2005). Inhibition of bacterial RNA polymerase by streptolydigin: stabilization of a straight-bridge-helix active-center conformation. Cell *122*, 541-552.

Van Gilst, M. R., and von Hippel, P. H. (1997). Assembly of the N-dependent antitermination complex of phage lambda: NusA and RNA bind independently to different unfolded domains of the N protein. J Mol Biol *274*, 160-173.

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

Vassylyev, D. G., Svetlov, V., Vassylyeva, M. N., Perederina, A., Igarashi, N., Matsugaki, N., Wakatsuki, S., and Artsimovitch, I. (2005). Structural basis for transcription inhibition by tagetitoxin. Nat Struct Mol Biol *12*, 1086-1093.

Vogel, U., and Jensen, K. F. (1995). Effects of the antiterminator BoxA on transcription elongation kinetics and ppGpp inhibition of transcription elongation in Escherichia coli. J Biol Chem *270*, 18335-18340.

von Hippel, P. H. (1998). An integrated model of the transcription complex in elongation, termination, and editing. Science *281*, 660-665.

Vo, N.V., Hsu, L.M., Kane, C.M., and Chamberlin MJ. 2003. In vitro studies of transcript initiation by Escherichia coli RNA polymerase. 2. Formation and characterization of two distinct classes of initial transcribing complexes. Biochemistry *8; 42(13*), 3787-97.

Wang, D., Bushnell, D. A., Westover, K. D., Kaplan, C. D., and Kornberg, R. D. (2006). Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. Cell *127*, 941-954.

Wang, Y., Severinov, K., Loizos, N., Fenyo, D., Heyduk, E., Heyduk, T., Chait, B. T., and Darst, S. A. (1997). Determinants for Escherichia coli RNA polymerase assembly within the beta subunit. J Mol Biol *270*, 648-662.

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*, 3612-3616.

Weilbaecher, R., Hebron, C., Feng, G., and Landick, R. (1994). Termination-altering amino acid substitutions in the beta' subunit of Escherichia coli RNA polymerase identify regions involved in RNA chain elongation. Genes Dev *8*, 2913-2927.

Weisberg, R. A., and Gottesman, M. E. (1999). Processive antitermination. J Bacteriol *181*, 359-367.

Wei, R. R. and Richardson, J. P. 2001a. Identification of an RNA-binding Site in the *ATPbinding* domain of *Escherichia coli* Rho by H_2O_2/Fe-EDTA cleavage protection studies. J. Biol. Chem. *276*, 28380-7.

Wei, R. R. and Richardson, J. P. 2001b. Mutational changes of conserved residues in the Q-loop region of transcription factor Rho greatly reduce secondary site RNA-binding. J. Mol. Biol. *314*, 1007-15.

Westover, K. D., Bushnell, D. A., and Kornberg, R. D. (2004). Structural basis of transcription: nucleotide selection by rotation in the RNA polymerase II active center. Cell *119*, 481-489.

Wilson, C., and Dombroski, A. J. (1997). Region 1 of sigma70 is required for efficient isomerization and initiation of transcription by Escherichia coli RNA polymerase. J Mol Biol *267*, 60-74.

Wilson, H. R., Zhou, J. G., Yu, D., and Court, D. L. (2004). Translation repression by an RNA polymerase elongation complex. Mol Microbiol *53*, 821-828.

Worbs, M., Bourenkov, G. P., Bartunik, H. D., Huber, R., and Wahl, M. C. (2001). An extended RNA binding surface through arrayed S1 and KH domains in transcription factor NusA. Mol Cell *7*, 1177-1189.

Xu, Y., Kohn, H. and Widger, W. R. 2002. Mutations in the Rho transcription termination factor that affect RNA tracking. J. Biol. Chem. *277*, 30023-30.

Yager, T. D., and von Hippel, P. H. (1991). A thermodynamic analysis of RNA transcript elongation and termination in Escherichia coli. Biochemistry *30*, 1097-1118.

Yang, L. Y., Jiang, H., Rangel, K. M., and Plunkett, W. (2003). Cisplatin-induced ubiquitination of RNA polymerase II large subunit and suppression of induction by 7-hydroxystaurosporine (UCN-01). Oncol Rep *10*, 1489-1495.

Yang, X. J., Goliger, J. A., and Roberts, J. W. (1989). Specificity and mechanism of antitermination by Q proteins of bacteriophages lambda and 82. J Mol Biol *210*, 453-460.

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*, 5301-5305.

Yanofsky, C. (2000). Transcription attenuation: once viewed as a novel regulatory strategy. J Bacteriol *182*, 1-8.

Yanofsky, C., Konan, K. V., and Sarsero, J. P. (1996). Some novel transcription attenuation mechanisms used by bacteria. Biochimie *78*, 1017-1024.

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*, 1181-1189.

Yarnell, W. S., and Roberts, J. W. (1999). Mechanism of intrinsic transcription termination and antitermination. Science *284*, 611-615.

Yin, Y.W. and Steitz, T.A. (2004). The structural mechanism of translocation and helicase activity in T7 RNA polymerase. Cell *116*, 393-404.

Yuzenkova, J., Delgado, M., Nechaev, S., Savalia, D., Epshtein, V., Artsimovitch, I., Mooney, R. A., Landick, R., Farias, R. N., Salomon, R., and Severinov, K. (2002). Mutations of bacterial RNA polymerase leading to resistance to microcin j25. J Biol Chem *277*, 50867-50875.

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*, 107-109.

Zellars, M., and Squires, C. L. (1999). Antiterminator-dependent modulation of transcription elongation rates by NusB and NusG. Mol Microbiol *32*, 1296-1304.

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

Zhou, Y., Mah, T. F., Greenblatt, J., and Friedman, D. I. (2002). Evidence that the KH RNA-binding domains influence the action of the E. coli NusA protein. J Mol Biol *318*, 1175-1188.

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*, 4514-4518.

APPENDIX I

Bacterial Strains, Phages and Plasmids used

A.1 Bacterial strains

All the bacterial strains that were used in this study are derivatives of *Escherichia coli* and their genotypes have been listed in Table. Bacterial strains were routinely stored on solid agar plates at 4 °C and as thick suspensions in 20% glycerol either at −20 °C or at −70 °C. Plasmid harboring strains, were reconstructed whenever necessary by fresh transformations.

Bacterial strains, plasmids and phages used in this study

Strains/ plasmid/ phages	Relevant Genotype	References
Strains		
RS20	K3093 (K37 sup° Tn5::LacZ LacI $_{Q1}$).	Neely and Friedman, 1998.
RS31	RS20, ?RS45 lysogen carrying pLac-*nutR*-3T-lacZYA fusion.	This study.
DJ354	MG1655, *rpoC*120 btuB::Tn10(tet^r, ts).	From Din Jin.
DJ3401	MG1655, *rpoB* 3401(ts).	From Din Jin.
RS36	RS31, *rpoC*120 btuB::Tn10(tet^r, ts)	This study.
RS38	RS36, with pK8601 spec^R.	This study.
RS44	MG1655 with pBAD18M-*rpoC*WT.	This study.
RS62	RS20, ?RS45 lysogen carrying pLac-*nutR*-lacZYA.	This study.
RS68	RS62, *rpoC*120 btuB::Tn10(tet^r, ts).	This study.
RS69	RS68 with pK8601, spec^R.	This study.
RS78	RS36 with pRS63.	This study.
RS80	RS31, *rpoB*(ts) btuB::Tn10(tet^r, ts).	This study.
RS81	RS62, *rpoB*(ts) btuB::Tn10(tet^r, ts).	This study.
RS82	RS80 with pK8601.	This study.
RS83	RS81 with pK8601.	This study.
RS90	MG1655 with pHYD 534.	This study.
RS99	K9774 (K37 strain with H-19BStx::Cam lysogen).	From Dr. David Friedman.
RS158	K3093 *rpoB*(ts) btuB::Tn10(tet^r, ts)	This study.
RS215	K3093 *rpoC*120 btuB::Tn10(tet^r, ts)	This study.
RS223	K3093 *rpoB*G1045D.	This study.
RS385	DH5α with pRS385; amp^R	This study
RS422	DH5α with pRS422; amp^R.	This study
RS464:	DH5α with pRS464; amp^R.	This study
RS485	MG1655 with pTRC99A amp^R	From Dr. Sergei Borukhov
RS514	RS38 with pRS513; amp^R , spec^R , tet^R.	This study

Plasmids

pK8641	pTL61T with plac – H-19B *nutR*-T_R'T1T2-lacZYA fusion. T_R'T1T2 is a triple terminator cassette, amp^R.	Neely and Friedman, 1998.
pK8628	pTL61T with plac – H-19B *nutR*-lacZYA fusion, amp^R.	Neely and Friedman, 1998.
pK8601.	pGB2 with plac- H-19B N, $spec^R$.	Neely and Friedman, 1998.
pBAD18 M	*rpoC* gene cloned at HindIII-SacI site of pBAD18, amp^R.	Sen *et al.*, 2002.
pHYD534	*rpo*B cloned BamH1-HindIII site of pBR322.	From Dr.J.Gowri shankar.
pET28b-NusA	NusA, his tag at N-terminal cloned in pET28b.	From Dr. Irina Artsimovitch
pRS8	H-19B N cloned at NdeI/XhoI site of pET21a, amp^R.	This study.
pRS18	pTL61T with pT7A1-T1T2-LacZYA fusion cloned at EcoRI/BamHI site, amp^R.	This study.
pRS22	pTL61T with pT7A1-H-19B *nutR*-T_R'T1T2-lacZYA, amp^R.	This study.
pRS25	pTL61T with pT7A1-H-19B *nutR* (ΔcII) T1T2-LacZYA, amp^R.	This study.
pRS63	pBAD18M-*rpo*CP254L, amp^R.	This study.
pRS89	ARM mutant of H-19B N cloned at NdeI/XhoI site of pET21a, amp^R.	This study.
pRS111	pBAD18M-*rpo*CG333S, amp^R	This study.
pRS114	pHYD534-*rpo*B G544S, G1045D, amp^R	This study.
pRS117	pHYD534-*rpoB*G1045D, amp^R.	This study.
pRS119	WT NusG cloned at NdeI/XhoI site of pET21a, his-tag at C-terminal, amp^R.	This study.
pRS159	pHYD534-*rpoB* D185N,G1045D, amp^R	This study.
pRS160	pHYD534-*rpoB* V296M,G1045D, amp^R	This study.
pRS163	pHYD534-*rpoB* G162S, G1045D, amp^R	This study.
pRS196	pGB2 with plac- H-19B N L108F, $spec^R$	This study.
pRS243	H-19B N L108F cloned at NdeI/XhoI site of pET21a, amp^R.	This study.
pRS245	pHYD534, *rpoB*G1045D with his tag at C-terminal, amp^R.	This study.
pRS281	pBAD18M-*rpoC* R270S, amp^R.	This study.
pRS283	pBAD18M-*rpoC* P251S P254L, amp^R.	This study.
pRS306	pBAD18M-*rpoC* P251S, amp^R.	This study.
pRS385	pRS25 with T7A1-*nutR*- *lac*O -T_R· fusion, amp^R	This study
pTRC99A	*E.coli rpoB* with N terminal HMK, His tag ß S531Y- NPH *rpoB*, amp^R	From Dr. Sergei Borukhov
pRS422	CTD H-19BN cloned at NdeI/XhoI site of pET21b, amp^R	This study
pRS464	N terminal HMK tagged H-19BN cloned into XhoI/NdeI site of pET33b (+), kan^R	This study
pRS484	pTRC99A- *E.coli rpoB* with N terminal HMK, His tag ß S531Y- NPH *rpoB*, amp^R	This study
pRS513	pBAD18M- *E.coli rpoC* with C terminal HMK and His tag, amp^R	Laptenko *et al.*, 2003.
pRS523	N terminal HMK, Histag NusA cloned into cloned into XhoI/NdeI site of pET33b (+), kan^R	This study This study

Phages	
?RS45, ?*c*I857,	Dr.J.Gowrishankar
P1*kc* (P1 phage)	CDFD
?C1B2	
Y1090 (? *Nin*5)	From Dr. R.A.Weisberg.

APPENDIX II

Oligos used in this study

Oligos		
Name	***Sequence***	***Description***
RS2	CTTGCATGCCTGCAGGTCGACTC	Downstream primer after $T_{R'}$ on T7A1-*nutR*-$T_{R'}$ in pK8641, pRS25#.
BioRS2	CTTGCATGCCTGCAGGTCGACTC	Biotinylated RS2
RS8	CGCGCGTTCCACCGGTTCTTACAGCC	*rpoC* gene, between 50216 to 50244
RS10	TTACTCGTTATCAGAACCGCCCAGACCTGCGTTC	C terminal end of *rpoC*
RS11a	GTGAAAGATTTATTAAAGTTTCTG	N terminal oligo of *rpoC*
RS11b	CGTTTTTCGTTACGTACGATGATGTCC	*rpoC* 864-890
RS11c	ATCATCCACTCTGGTTTGTTACC	*rpoC* 691-713
RS12	CTGCGCCCGTATCTTTGGGCCGG	*rpoC* 50761 to 50784
RS13	CGCGCTCAACCATTTTCTTCGCAGC	*rpoC* 51785-51801
RS14	GGTGGTCGTTTCGCGACTTCTGACC	*rpoC* 51358-51384
RS15	CCCAGCGCCTGGTTGACGATGG	*rpoC* 52360-52380
RS16	CCGGTACTGATCGAAGGTAAAGC	*rpoC* 51905-51928
RS17	GGGTGGAGATGAAGTACTGGAGTACG	*rpoC*, 52895-52920
RS18	GCGGACCAGATCATCATGTACACACCGGC	*rpoC*, 52451-52475
RS19	CCGACTTCACGTTGCTGAGCTTGATGC	*rpoC*, 53460-53486
RS20	CCCATGAAGGTATCATGATGACTCCGG	*rpoC*, 53037-53053
RS21	GCGAGGATTGCCGGCTCTTTCGGACG	*rpoC*, 54035-54060
RS22	CCTTACGGTGCGGTACTGGCGAAAGGCG	*rpoC* 53581- 53609
RS23	CGGCTGCTTCGGTCAGCACGCGAGTGG	*rpoC* 53573-54600
RS24	CCGCAGGTGGTAAAGATCTGCGTCGTCCGG	*rpoC* 53794-53820
RS24a	AGTACAGGACGTATACCGTCTGC	*rpoC* 3445-3470
RS58	ATAAACTGCCAGGAATTGGGGATCG	Upstream oligo of pTL61T (and all its derivatives like pK8641, pRS22, pRS25) vector sequence.
RS59	TCATCTTGCTGTCAGTTGTTTGGATTTCCG	H19-B N gene specific primer (C-terminal)
RS63	GATTATTTGCACGGCCTCACA	Upstream to for *rpoC* gene in pBAD vector.
RS64	GCGCCTCGAGTCTTGCTGTCAGTTGTTTGG	Downstream primer of H-19BN gene with XhoI site without stop codon.
RS65	GCGCCTCGAGTCATCTTGCTGTCAGTTG	Downstream primer of H19B N gene with XhoI site and stop codon.

RS66a	GCGCGCCATATGACACGCAGAACTCAGTTC	Upstream primer for H-19B N gene with NdeI site
RS67	GCGCGAATTCGGATCCAGATCCCGAAAATTTATC	Upstream primer with EcoRI site for amplifying T7A1 promoter
RS68	GCGCAAGCTTCGCCGTGTCCCTCTCGATGG	Downstream primer with HindIII site for amplifying T7A1 promoter
RS70	GCTCTAGAGAATTGTGAGCGCTCACAATTCTCTAG AGC	*Lac* Operator palindrome flanked by XbaI sites on either side.
RS75	CCCGGGTGCCGGATCCTCTAGAGTCG	Upstream primer of $T_{R'}$ with SmaI/BamHI/XbaI sites
RS76	GATCTCCTAGGCCGTGGGCCCAATAGCTAGTTGAA AGGTATGCGTGG	Downstream primer of *nutR* site with SmaI/BamHI/XbaI sites.
RS76a	GCGCGCGATCTCCTAGGCCGTGGGCCCAATAGCTA GTTGAAAGGTATGCGTGG	Downstream primer of *nutR* site with SmaI/ BamHI/ XbaI sites.
RS77	GCGCGCGATCTCCTAGGCCGTGGGCCCGCTCTACG CGATACGAACACTAGG	Downstream primer for making *nutR* deletion, upstream to *nut* site with SmaI/BamHI/XbaI sites.
RS78	GCTAGTTATTGCTCAGCGGT	T7 terminator sequence
RS79	TAATACGACTCACTATAGGG	T7 promoter sequence
RSRK-1	CGCCAGGGTTTTCCCAGTCACGAC	Downstream primer in the *lacZ* gene.
RS99	GAATTGTGAGCGCTCACAATTCGGATCCGGCACCC GGGAATAGC	T7A1-*nutR-LacO*fusion.
RS101	GCGCGCCATATGACACGCAGAACTCAGTTCAAAGGC AATTCTCGTTCTGCAGCAGCAGAGCGTTTAAAG G.	For replacing ARM Arg motif of H-19B N, by with 3 Ala, will be used as upstream primer with NdeI site.
RS107	CCGGCACCCGGGGAATAGCTAG	Downstream primer to make *nutR* RNA together with upstream primer RS108.
RS108	GGAATTCCTAATACGACTCACTATAGGGATCGAGA GGGACACGGG	Upstream primer withT7 promoter.

RS114	GCGCGCGCCATATGTCTGAAGCTCCTAAAAAGCGC	*E.Col* NusG upstream primer with NdeI site
RS115	GCGCGCGCCTCGAGTTAGGCTTTTTCAACCTGGC	*E.Coli* NusG downstream primer with stop codon and XhoI site.
FORB1	ATAGGATCCAAACTGTCCGCTCAATGGAC	*rpoB*, positioned ~ 50 bp upstream of the start codon
REVB1	ATA*GGATCC*GTTAGCACAATCCGCTGCCG	*rpoB*, positioned ~ 50 bp downstream of the stop codon
SeqB1	CGTTAGTTGGGTCGACACG	*rpoB* 1080 region.
SeqB2	CTGCTGGCTAAGCTGAGCC	*rpoB* 960 region
SeqB3	CGCAGAGTCGGAACGGCC	*rpoB* 2070 region
SeqB4	GTGGTATCCGTCGGTGCG	*rpoB* 1980 region
SeqB5	GCGGCGTTTCGCTTCGAG	*rpoB* 3090 region
SeqB5a	AAAGACAAACGTGCGCTGGAAATCGAAGA	*rpoB* 2881 region
SeqB6	GAGCTGGGCCTGACAGACG	*rpoB* 3000 region
SeqB7	GATCTGGGCGCTGACGTTCGTC	*rpoB* 3448- 3460
SeqB8	GATCTGCTCTGTGGTGTAGTTCAGG	*rpoB* 570-590, reverse
SeqB9	CGACATCATTGATGTTATGAAAAAGCTC	*rpoB* 1271-1294
SeqB10	GGTCAGACCGCCTGGGCCGAGTGC	*rpoB* 15985-1615
SeqB11	CATGCGCGTAGCGTTCATGCCGTGG	*rpoB* 2340-2321
SeqB12	CGTCGACCAGGTGGTTCAGTTTCAGC	*rpoB* 3695- 3718
RS125	GCG CGC GCCATATGAAACCAGTAACGTTATACG	Forward primer for *lac* repressor gene.
RS126	GCG CGC CTC GAG TCA TGCCCGCTTTCCAGTCGGG	Reverse primer for *lac* repressor gene, with stop codon.
RS128	ATGGC AACCATGAAA G*AAGCTT*CGGCTGGC	Forward primer for amplification of *rpoB* from pHYD534.
RS129	GCGCGCGCGGATCCTTAGTGGTGGTGGTGGTGGTG CTC GTCTTCCAGTTCGATGTTGATACCC	Reverse primer with C-terminal His-tag, BamH1 site for *rpoB* amplification.
RS130	GCGCGCGCATCGAT A ATG TGC CAA TCG CGG GGG G.	Upstream primer for lambda N gene with Cla1 site
RS131	GCGCGCGCGGATCCCTAGATAAGAGGAATCGATTT TCCC	Downstream oligo for lambda N gene with BamH1 site with stop codon.

RS132	TAAGGAGGTATATCGATAATGACACGCAGAACTCAG	H-19B N forward primer with ClaI site to be amplified from pK8601.
RS133	GCTGCAGGTCGACGGATCCTTAGTTACTTACCCGG	H-19B N reverse primer with BamH1 site to be amplified from pK8601
RS135	ACGATCTGGCACCGGACGTGCTGAAGATTG	Top strand primer for SDM with 3134 mutation of *rpoB* G0145D.
RS136	CAATCTTCAGCACGTCCGGTGCCAGATCGTC	Bottom strand primer for SDM with 3134 mutation of *rpoB* G1045D
RS137	ACGATCTGGCACCGGGCGTGCTGAAGATTG	WT top strand primer of 3134 region of *rpoB*.
RS138	CAATCTTCAGCACGCCCGGTGCCAGATCGTC	WT bottom strand primer of 3134 region of *rpoB*
RS143	CCCGAAACCGTTATGCAGGC	Anti sense to downstream part of $T_{R'}$ stem will be used for oligo induced release RNA release from $T_{R'}$ RB.
RS144	AGATCTGCGTTCGCTGGTTCCGCTGGATGG	Forward primer, P251S change, used for SDM to make P251S from P254L-P251S *rpoC* double mutant.
RS145	CCATCCAGCGGAACCAGCGAACGCAGATCTGG	Reverse primer, P251S change, used for SDM to make P251S from P254L-P251S *rpoC* double mutant.
RS172	ACGTGATTTTACTTTTTCGGGGATC	Reverse primer, downstream to *Nin*, in H-19B Q gene.
RS173	CTTTTTAATGGTGAAAGCCATGC	Forward primer, upstream to *Nin* orf 58B of H-19B.
RS175	GTGCAACCAACATAAAAGCGGAAATCTCG	Upstream to *Nin* orf 59 of H-19B.
RS191	GAATTGTGAGCGCTCACAATTCGCTAGTTGAAAGGTATGCG	*lacO* fusion just upstream of SmaI site of T7A1-*nutR* used for antisense oligo release
RS192	GAATTGTGAGCGCTCACAATTCAAAGCAAAAAAGGTATGCGTGG	T7A1-*nutR*-*lac*Ofusion, RNA:DNA hybrid changed to HK pause sequence.

RS200	CGAAAGTAGC GCGCAGATGA GTGTGA	Upstream to H-19B *Nin*, after *Ren* gene
RS201	CAGCCTCCCCTGTTACTTTAAGCATT	Reverse primer, downstream to *roi* of H-19B.
RS202	GCGCGCCTCGAGTCACTTATGTCCAAACTCATTCGC GTACACAAT	Downstream primer for making del-CTD of H-19B N.
RS203	GTTGAAAGGTATGCGTGGAAACGC	Primer between *box B* and RB used for cleaving the tether of N-modified RB.
RS204	GCGCGCGGATCCTCAGGATGAACGGCAATGCTGCT C	Upstream primer for amplifying *his* pause sequence from pRK771a
RS228	GCGCGCGCGAGCTCTTAGTGGTGGTGGTGGTGGTG AACAGATGCACGACGCTCGTTATCAGAAC CGCCCAG	C-term C terminal *rpoC* with HMK, 6-His tag and SacI site.
RS247	GAATTGTGAGCGCTCACAATTCCTTGCATGCCTGCA GGTCGACTC	*LacO* after $T_{R'}$, used to collect N-modified RB after passing through the terminator.
RS248	ACGTCGACTCTAGACGCACGCTACCGCCGGGGAAT AGCTAGT	Downstream *ops* pause site primer on pRS25.
RS262	GGAATTGTGAGCGCTCACAATTCACGCACGCTACC GCCGGGGAATA	Downstream primer used to make *nutR-ops-LacO* using RS58/RS248 as template
RS263	CTGAAAGACTAGTCAGGATGATGGTTGGCCTTAGT TGGTCAGATATATTGGG	Downstream oligo1 for making *His* pause template.
RS264	CTGAATGTCTTCCAGCACACATCGCCTGAAAGACT AGTCAGGATGATGGTTG	Downstream oligo 2 for making *His* pause template
RS265	TTCATGCACCACTGGAAGATCTGAATGTCTTCCAGC ACACATCGC	Downstream oligo 3 for making *His* pause template
RS267	GGTATGCGTGGAAAAAACGCACGCTACCGCCTGGC CTTAGTTGGTCAGATATATTGGG	Downstream primer for amplifying *ops* pause sequence just before the H-19B internal pause seq.
RS274	GGAATTGTGAGCGCTCACAATTCCTTCCAGCACAC ATCGCCTGAAAGACTAG	RB after 3nt of the His pause position
RS275	GGAATTGTGAGCGCTCACAATTCAAC*G*CACGCTAC CGCCTGGCC	RB after 4nt of the *His* pause position
RS276	GGAATTGTGAGCGCTCACAATTCAAACGCACGCTA	RB after 3nt of *ops* pause

	CCGCCTGGCC	position
RS277	GGAATTGTGAGCGCTCACAATTCAAACGCACGCTACCGCCTGGCC	Road Block after 4nt of *ops* pause position
RS283	CCGGTAGCCTGCCGCGTAAAGACATACATTTGGAAGAGCTACACCAGTTTCTGG	Downstream primer 1 for generating A112C C107A Cys mutant in CTD of H19B-N by overlap PCR
RS284	GCGCGC*CTCGAG*TCATCTTGCTGTCAGTTGTTTGGATTTCCGGTAGCCTGCCGCGTAAAG	Downstream primer 2 for generating A1162 C107A Cys mutant in CTD of H19B-N by overlap PCR, with XhoI site and stop codon
RS285	TTTCCGGTAGCCACACGCGTAAAGAGCTACATTTGGAAGAGCTACACCAGTTTCTGG.	Downstream primer 1 for generating A116C C107A Cys mutant in CTD of H19B-N by overlap PCR
RS286	GCGCGC*CTCGAG*TCATCTTGCTGTCAGTTGTTTGGATTTCCGGTAGCCACACGCGTAAAG	Downstream primer 2 for generating A116C C107A Cys mutant in CTD of H19B-N by overlap PCR, with XhoI site and stop codon

APPENDIX III

Molecular Biology Techniques Employed

A 3.1 Isolation of plasmid DNA

The rapid alkaline lysis method of plasmid isolation, as described by Sambrook (Sambrook et al., 2001) was followed with minor modifications. Bacterial pellet from 3 ml of stationary-phase culture was resuspended in 200 µl of ice-cold solution I (50 mM glucose, 25 mM Tris-Cl pH 8.0, 10 mM EDTA pH 8.0 containing 1 mg/ml lysozyme) by vortexing. After 5 min incubation at room temperature, 400 µl of freshly prepared solution II (0.2 N NaOH, 1% SDS) was added and the contents were mixed, by gently inverting the tube several times. This was followed by the addition of 300 µl of ice-cold solution III (5 M potassium acetate, pH 4.8) and gentle mixing. The tube was incubated on ice for 10 min and centrifuged at 12,000 rpm for 15 min at 4°C. The clear supernatant was removed into a fresh tube and, if required, was extracted with an equal volume of phenol: chloroform mixture. The supernatant was precipitated with either two volumes of cold 95% ethanol or 0.6 volumes of isopropanol at room temperature for 30 min. The nucleic acids were pelleted by centrifugation, washed with 70% ethanol, vacuum dried, and dissolved in appropriate volume of TE buffer. If required, the sample was treated for 30 min with DNase free RNase at a final concentration of 20 µg/ml. The plasmid DNA was checked on a 0.8% agarose gel and stored at –20 °C. The plasmid DNA thus isolated was suitable for procedures such as restriction digestion, ligation, and preparation of radiolabeled probes. Alternatively, in cases where high purity of the plasmid preparation is required (for example, for procedures such as automated DNA sequencing), plasmid DNA was also prepared by using the commercial kit for plasmid DNA isolation (mini prep and midi prep) from Qiagen, following the manufacturer's protocols.

A.3.2 Agarose gel electrophoresis

The DNA samples were mixed with the appropriate volumes of the 6X loading dye (0.25% bromophenol blue, 0.25% xylene cyanol and 30% glycerol in water) and subjected to electrophoresis through agarose gel (0.8% for plasmid DNA and 1%-1.6% for PCR products, as per the size variations of the DNA fragments) in either 1x TBE or 1x TAE buffer. The gel was stained in 1 µg/ml of Ethidium Bromide solution prepared in Ix TAE for 10 min at room temperature and the bands were visualized by fluorescence under UV-light.

A.3.3 PCR Reactions

Polymerase Chain Reactions (PCR) was carried out for amplifying DNA on PCR machines (Mastercycler, Eppendorf). The concentrations of the components were as follows: DNA 20-50 ng, Primers (forward and reverse)- 0.5 µM; dNTPs - 0.25 mM; $MgCl_2$- 2.5 mM; Taq DNA Pol

(Roche), as per manufacturer's instructions. Proof reading DNA Pols (Pfu-Stratagene, DeepVent-NEB and Dnazyme-Finnzymes) were used for amplifying DNA used in cloning and for templates in transcription reactions, concentration of which ranged from 1-2.5 U/50 µl, as per the length of DNA to be amplified.. For amplifying *rpoC/B* subunit genes, we have used 0.3 mM $MgCl_2$. The initial denaturation was done at 95 °C for 2 min; denaturation, annealing (at 55 °C) and extension (72 °C) were for 30 s, 30 s and 2 min respectively, reaction was for 30 cycles. The final extension was done at 72 °C for 10 minutes. For larger DNA products (above 1 kb), we have used the following cycle: The initial denaturation was at 95 °C for 2 min; denaturation, annealing (at 55 °C) and extension (72 °C) were for 30 s, 1 min and 2 min respectively and reaction was for 30 cycles. The final extension was at 72 °C for 15 minutes. For amplifying *rpoC/B* subunit genes, we used the following cycle: The initial denaturation at 95 °C for 2 min; denaturation, annealing (at 55 °C) and extension (72 °C) were for 30 s, 1 min and 6 min respectively, 30 cycles. The final extension was at 72 °C for 15 minutes.

A.3.4 Restriction enzyme digestion and analysis

0.5-1 µg of DNA was used for each restriction enzyme digestion. 2-4 units of the restriction enzymes with the appropriate 10X buffers supplied by the manufacturers and BSA (as when required) were used in the reactions. The digestion was allowed to proceed for 3-6 hour at the temperature recommended by the manufacturer. The DNA fragments were visualized by Ethidium Bromide staining following electrophoresis on agarose gels. Commercially available DNA size markers were run along with the digestion samples to compare with and to estimate the sizes of the restriction fragments. The restriction digested DNA samples that were to be used for cloning purposes were phenol extracted and ethanol precipitated or purified and concentrated by passing through Centricon centrifugal filters before the next step.

A.3.5 Purification of DNA fragments by elution from the gel

DNA fragments to be used for specific purposes such as ligation or radioactive labeling were eluted from the agarose gel, after the electrophoresis. The gel piece containing the desired band of DNA was sliced out using afresh scalpel blade and the DNA was purified from the gel using the purification kits available for this purpose, following the manufacturer's protocols. Qiagen gel extraction and purification kits were used through out this study. The efficiency of elution was determined by checking a small aliquot of DNA sample on the gel.

A.3.6 Klenow fill-in reaction

The Klenow fill-in reaction was carried out after DNA fragments were digested with restriction enzyme(s) that leave incompatible recessed 3′ ends, in order to convert them into blunt-ended products for ligation. After digestion, the restriction enzyme was heat inactivated by incubation at 65 °C for 10 min; the fill-in reaction was then set up by supplementation with 33 µM (final concentration) of each dNTP and addition of Klenow DNA polymerase to a final concentration of 1 unit/µg of DNA. (The Klenow enzyme is 50% active in standard restriction enzyme buffers.) The reaction was carried out for 5 to 10 min at 37°C, following which the enzyme was heat inactivated by 10 min incubation at 75°C. In the case of DNA fragments eluted from gel, the fill-in reaction was carried out in 1x Klenow buffer (10 mM Tris-HCl (pH 7.5), 5 mM $MgCl_2$, 7.5 mM DTT).

A.3.7 Ligation of DNA

Typically, 100-200 ng of DNA was used in each ligation reaction. The ratio of vector to insert was maintained between 1:3 and 1:5. The reactions were usually done in a 10 µl volume containing ligation buffer (provided by the manufacturer) and 0.05 Weiss units of T4 DNA ligase, and the reactions were incubated at 16 °C for 12-16 h.

A.3.8 End labeling of DNA markers

DNA markers were radiolabelled with T4 Polynucleotide Kinase (NEB). Reactions were carried out essentially as per manufacturer's protocols, with the buffer provided along with the enzyme and ? P^{32} ATP (Amersham/BRIT). The labeled DNA was purified by passing through G-25 spin columns (Amersham).

A.3.9 Transformation protocols

A.3.9.1 Calcium chloride method: For routine plasmid transformations, where high efficiency is not required, the following method, which is a modification of that described by Sambrook et al. was used. An overnight culture of the recipient strain was subcultured in fresh LB with the appropriate antibiotics as required and grown till mid-exponential phase. The culture was chilled on ice for 15 min, and the steps hereafter were done on ice or at 4 °C. The culture was centrifuged, and the pellet was resuspended in one-third volume of cold 0.1 M $CaCl_2$. After 15 min incubation on ice, the cells were again recovered by centrifugation. If the competent cells prepared were to be stored for later use, storage purposes, $CaCl_2$ with 20% glycerol was added

and stored in ⁻70 °C as 100 µl aliquots. Alternatively, if directly using for transformation, the pellet was resuspended in one-tenth volume of cold 0.1 M $CaCl_2$. The suspension was incubated on ice for 5 minutes and aliquoted 100 µl each, in sterile epppendorf tubes; incubated for 30 minutes and then DNA was added (~10-100 ng of DNA in less than 10 µl volume). The mixture was again incubated on ice for 30 min, and then was subjected to a heat shock for 90 seconds at 42 °C. The tubes were plunged into ice immediately after the 90 min heat shock, incubated in ice for 2 minutes and then 0.9 ml of LB broth was added to the tube, incubated at 37 °C for 45 min for phenotypic expression of the antibiotic marker before being plated on selective medium at various dilutions or as a whole, after centrifuging the cells at 4000 rpm and plating the whole pellet resuspended in 100 µl LB. A negative control tube (with no plasmid DNA addition) was also routinely included in each of the experiments.

A.3.9.2 Preparation of high efficiency competent cells (ultra competent cells)

Competent cells for high efficiency transformation were prepared as follows. An overnight culture of the recipient strain (usually, strain DH5α) was subcultured into fresh LB broth at 1:100 dilution and allowed to grow with good shaking at 37 °C till an A_{600} of ~0.5. The cells were chilled on ice for 15 min and centrifuged at 4000 rpm at 4°C. The pellet was resuspended in one-third volume of ice-cold RF1 buffer (30 mM potassium acetate, 100 mM Rubidium Chloride (RbCl), 50 mM $MnCl_2.4H_2O$, 10 mM $CaCl_2.2H_2O$, and 15% glycerol, pH 5.8) and incubated on ice for 20 min. The cells were recovered after centrifugation at 4000 rpm at 4 °C and the pellet was resuspended in one-twelfth volume of ice-cold RF2 solution (10 mM MOPS pH 6.8, 10 mM RbCl, 75 mM $CaCl_2.2H_2O$, and 15% glycerol pH 6.8). The cell suspension was again incubated on ice for 15 min, and 100 µl aliquots were then distributed into pre-chilled microfuge tubes. The aliquots were flash-frozen in liquid nitrogen and stored at ⁻70 °C till further use.

Alternatively, Inoue's method (Sambrook and Russel, 2001) was also employed; the difference from the above method being that ice cold Transformation Buffer (10mM PIPES (free acid), 15 mM $CaCl_2.2H_2O$, 250 mM KCl, and 55 mM $MnCl_2.2H_2O$; pH adjusted to 6.8 with 1 N KOH and filter sterilized) was used, instead of RF1 buffer. In addition, during the final re-suspension step, DMSO (to a final concentration of 7%) could be added as the cryo-protectant, instead of glycerol. For transformation, the tube was thawed on ice, DNA was added to it and the subsequent protocol followed was as described early, for transformation.

A.3.9.3 Electroporation

A.3.8.3.1 Preparation of electro competent cells

E.coli cell cultures were grown in LB with good aeration to an A_{600} of ~ 0.6 to 0.8. The culture was chilled on ice for 20 min and all subsequent steps were carried out on ice or at 4 °C. The cells were harvested by centrifugation, washed three times with ice-cold sterile distilled water, then once with ice-cold 10% glycerol, and finally resuspended in the latter in $1/500^{th}$ the volume of the original culture. The thick cell suspension was dispensed into aliquots of 50 µl, flash frozen using liquid nitrogen and stored at ⁻70 °C. Aliquoted tubes were taken when required, was thawed on ice and used for electroporation as described below.

A.3.9.3.2 Electroporation: A 20 µl aliquot of electro competent cells was transferred into an electroporation cuvette, the DNA to be electroporated was added to it, and the mix was kept on ice 45 min. The amount of DNA that can be added is limited by the salt concentration in the DNA used (typically, from a ligation mix, 1-2 µl of DNA was used). Electroporation was carried out in an electroporator (Biorad) following the manufacturers instructions. Generally, a pulse, at a resistance of 200 O and at 1800 V was used for electroporation of the *E.coli* electro competent cells. Immediately after electroporation, SOB medium was added to the cells in the cuvette. The cells were transferred to a sterile glass tube, was incubated for 30 min to 1 hr with shaking for aeration and then plated into the appropriate selection plates.

A.3.10 Glycerol Stock Preparation:

Single colonies of the required bacteria were inoculated into 3 ml LB with the appropriate antibiotics and grown overnight in a shaker incubator at the required temperature. The overnight culture was centrifuged at 4000 rpm and the supernatant was discarded without disturbing the bacterial pellet. The pellet was then resuspended in 750 µl of sterile 10Mm $MgSO_4$. This was then transferred into sterile cryovials with proper labeling. 250 µl of sterile 80% glycerol was added into this (final concentration of glycerol becomes 20%), mixed well and was stored in ⁻70 °C deep freezer.

All the media and buffers (except those which were heat labile) were sterilized by autoclaving for 15 minutes at 121°C. Media and buffers used in this study are described below.

LB medium:

Tryptone	10 g
Yeast extract	5 g
NaCl	10 g
Water to	1000 ml, pH adjusted to 7.0 - 7.2 with 1 N NaOH.

LB agar:

LB medium	1000 ml
Bacto-agar	15 g

LB soft agar:

LB medium	100 ml
Bacto-agar	0.6 g

MacConkey agar:

MacConkey Agar (Difco)	51.5 g
Water to	1000 ml

Citrate buffer: (0.1 M; pH 5.5)

Citric acid (0.1 M)	4.7 volumes
Sodium citrate (0.1 M)	15.4 volumes

Z Buffer IM, for IL

$Na_2HPO_4.7H_2O$ (O.06M)	16.10 gms/lt.
$NaH_2PO_4.7H_2O$ (0.04M)	5.50 gms/lt.
KCl (0.01M)	0.75 gms/lt.
$MgSO_4 .7H_2O$ (0.001M)	0.25 gms/lt.

pH was adjusted to 7.0 and stored at 4°C.Prior to use, added ßME (270 µl for 100ml).

ONPG

80 mg O- Nitro Phenyl-ß-D-Galactosidase to 20 ml milliQ water (4mg/ml)

1.0 M Na_2CO_3

Na_2CO_3 - 10.6 gms/lt; stored at room temperature.

TE buffer: 10 mM Tris-Cl (pH 8.0) and 1 mM EDTA.

TBE and TAE buffers:

TBE: 90 mM Tris-borate, 2 mM EDTA (pH 8.0) and TAE: 40 mM Tris-acetate, 2 mM EDTA (pH 8.0) were used as standard electrophoresis buffers. TBE and TAE were prepared as 10X and 50 x concentrated stock solutions, respectively, and used at 1x concentration. (Alternatively, 10X solution of TBE was made by dissolving 10x TBE concentrate powder from Ambion and dissolving in 1L RO water).

MNNG (n-Methyl, n`-Nitro, n-Nitrosoguanidine; Sigma): 1mg/ml stock solution was used. 35-50 mg of MNNG was dissolved in Acetone to make a 10 mg/ml solution of MNNG. Diluted 1:10 in

citrate buffer (0.5ml of MNNG in 4.5 ml of citrate buffer) for making the 1 mg/ml working standard

Sequencing Gel mix (For 500 ml)
10X TBE – 1X- 50 ml

40% Acrylamide (19:1 Acrylamide: BisAcrylamide) mix to the required percentage.

7M Urea - 210 g

Double distilled water-to make up till the final amount (1L)

Sequencing gel mix was also purchased from Ambion.

For casting sequencing gels, to the pre-made mix of 35 ml, 150 µl of 10% APS and 35 µl of TEMED were added.

Antibiotics

Antibiotics were used at the following final concentrations (µg/ml):

Ampicillin (for plasmids)	100
Chloramphenicol (for plasmids)	50
Kanamycin	50
Rifampicin	100
Spectinomycin	50
Tetracycline	15

Chemicals

Most chemicals were obtained from commercial sources. The sources for some of the fine chemicals used in this study are given below. Most of the chemicals such as antibiotics, sugars, IPTG, ONPG, MNNG, NTPs, DTT, DMF and X-gal were obtained from Sigma Chemical Co. The media components for the growth of bacteria were mostly from Difco laboratories. The materials used in the recombinant DNA experiments such as restriction endonucleases, T4 DNA ligase, DNA polymerase and DNA size markers were obtained from companies including New England Biolabs, MBI, Sigma, Finnzyme, Gibco-BRL, Promega and Amersham (USB). Kits used for plasmid isolation, purification of DNA fragments from agarose gel were purchased from Quiagen. The oligonucleotide primers used in this study were mainly synthesized from Sigma Aldrich and MWG Biotech. The radioactive NTPs were from BRIT, Mumbai as well as from Amersham, UK.