TARGET(S) OF TRANSCRIPTIONAL REGULATORS ppGpp AND DksA

By

NIMISH GUPTA

A thesis submitted to the Graduate School-New Brunswick

Rutgers, The state University of New Jersey

In partial fulfillment of the requirements

For the degree of

Master of Science

Graduate Program in **Chemical and Biochemical Engineering**

Written under the direction of

Dr. Richard H. Ebright and Dr. Charles M. Roth

And approved by

____

____

____

New Brunswick, New Jersey

January, 2009

ABSTRACT OF THE THESIS

Target(s) of transcriptional regulators ppGpp and DksA

By NIMISH GUPTA

Thesis Director:

Dr. Richard H. Ebright and Dr. Charles M. Roth

Bacterial RNA polymerase is a highly regulated enzyme, altering the metabolic and functional state of bacteria in response to their environment. rRNA synthesis is a major metabolic expense and is, thus, highly regulated. One such response is the ppGpp-induced stringent response which changes the metabolic state of the cell, particularly inhibiting rRNA synthesis. ppGpp, guanosine tetraphosphate, is produced rapidly during amino acid starvation. ppGpp acts along with an accessory molecule, a protein DksA. The mechanism and binding site of ppGpp/DksA remains unclear and disputed. The objective of this study is to determine the binding site of ppGpp and DksA and understand the mechanism by which they regulate the cell's metabolic state. The strategy is to randomly mutagenize RNAP genes, specifically *rpoC* which codes for β' subunit, and perform genetic screens to identify mutants resistant to toxic effects of ppGpp/DksA. The mutants were isolated, sequenced, tested and characterized by *in vivo* and *in vitro* assays. The mutants isolated in this study belong to two distinct clusters: the

switch region, which controls the opening and closing of active center clamp allowing dsDNA entry, and the secondary channel, which is the channel through which NTP substrates enter. The mutants in the switch region could imply a conformational determinant for ppGpp/DksA. The secondary channel mutants could be binding/conformational determinants for ppGpp/DksA. The *in vitro* results indicate that two mutants in the secondary channel interact with DksA which could confirm DksA's putative binding site.

Acknowledgements

I would like to thank first of all, Dr. Richard Ebright, who gave me this opportunity to work in his lab and for his guidance throughout.

I would like to thank Dr. Charles M. Roth for his support and cooperation; Dr. Henrik Pedersen for being on the committee.

I would like to thank all my lab members for their support, assistance and help whenever I needed it. It would have been a much tougher ride without their help.

And above all,

This thesis is dedicated to my Dad, Mom, brother and Nisha Madam, who are the driving force for all my work, a constant source of inspiration and are the ones who taught me the most important things in life and made me what I am.

Table of Contents

Chapter 1: Introduction

1.1 Transcription

Inside the cells, the DNA or the genetic code of any organism has to be transcribed into RNA which eventually translates into proteins. These proteins perform the important functions within the cell. The process of formation of RNA from DNA is known as transcription.

The cells respond to their environment and adjust their internal state by regulating key steps in the gene expression route. Transcription is the first step in gene expression and is the most regulated step. RNA polymerase is the enzyme responsible for this process.

1.2 RNA Polymerase

RNA Polymerase (RNAP) recognizes short sequences of DNA called promoters and transcribes the genes under control of that promoter. The promoter is located upstream of the coding sequence of every gene. The RNA polymerase is a primary target for regulation (Young, Gruber, & Gross, 2002). The essential component of RNAP is evolutionarily conserved in all living organisms from bacteria to eukaryotes and the structural similarities are striking (Ebright, 2000).

In Bacteria, RNA polymerase performs synthesis of all RNAs (mRNAs, tRNAs, rRNAs). The bacterial RNAP exists in two forms: Core and Holo (Vassylyev, et al., 2002). Core Bacterial RNAP (Figure 1) is a five subunit enzyme (ββ’α’α’’ω) (Darst, 2001). Bacterial RNA polymerase has a molecular mass of around 400 kDa. The eukaryotic enzyme has more than a dozen subunits and has a molecular mass greater than 500 kDa (Zhang, Campbell, Minakhin, Richter, Severinov, & Darst, 1999).

The core enzyme can perform catalysis but does not have any specificity for promoters. When this core enzyme associates with σ subunit, the holoenzyme is formed (Darst, 2001). The holoenzyme (Figure 2) has specificity for promoters (Murakami & Darst, 2003). (Figure 3)

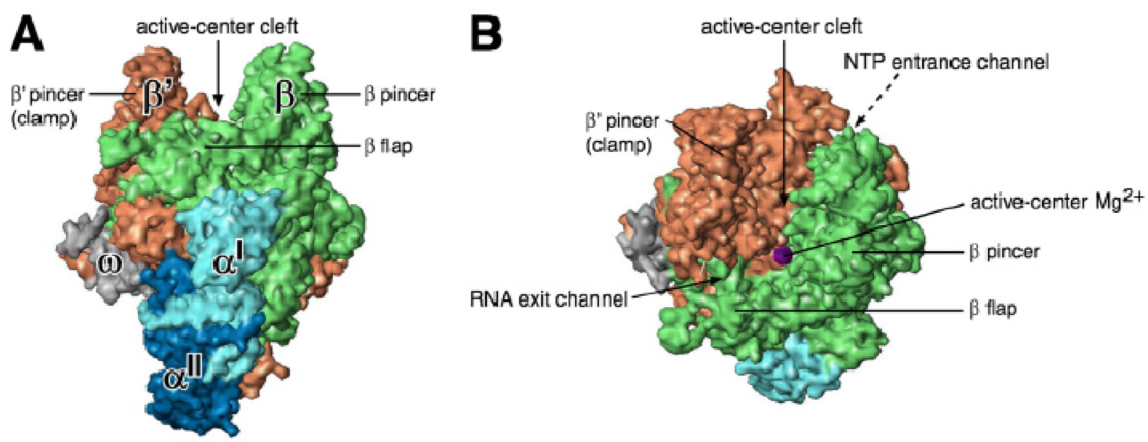


Figure 1. RNAP core. β' is in orange; β is in green; αI is in light blue; αII is in dark blue; ω is in gray; the active-center Mg2+ is in violet. (A) "Upstream" face (B) "Top" face (Zhang, Campbell, Minakhin, Richter, Severinov, & Darst, 1999)

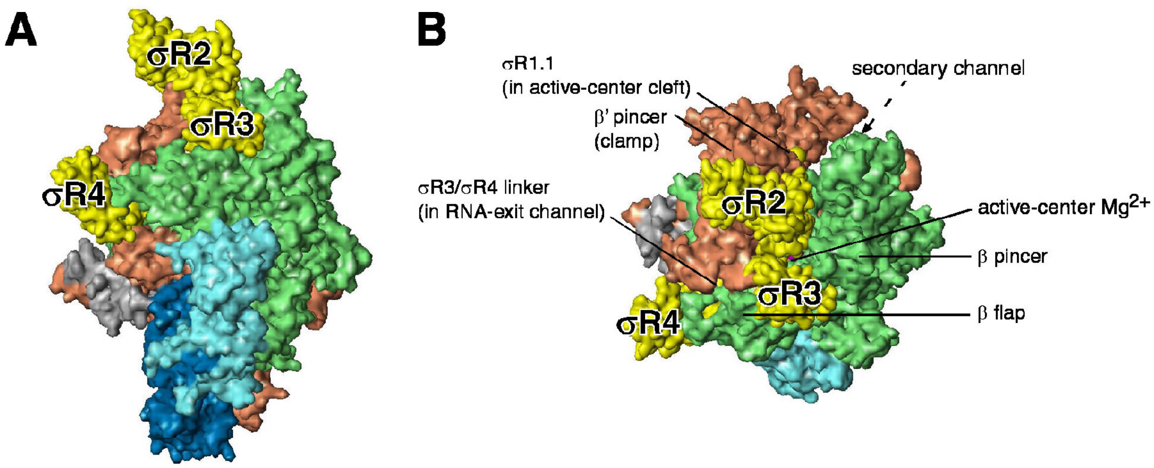


Figure 2. RNAP holoenzyme: (A) Upstream face (B) Top Face (Vassylyev, et al., 2002)

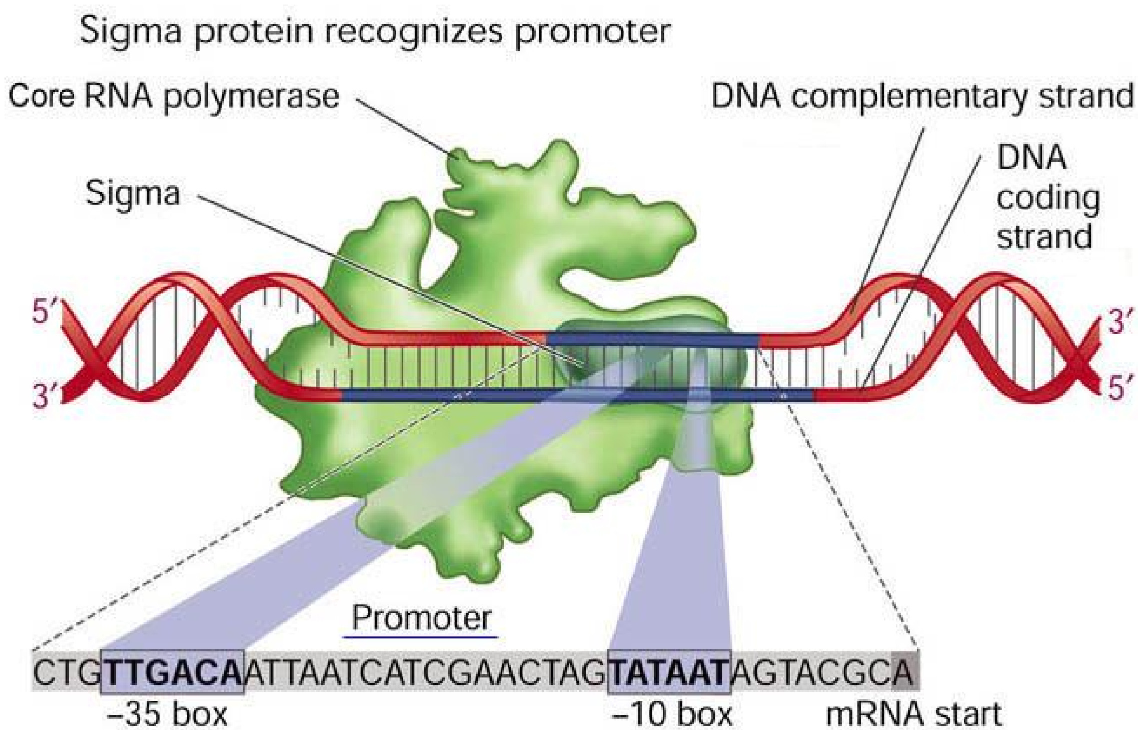


Figure 3. Bacterial RNAP holoenzyme recognizes promoters. (Source: http://porpax.bio.miami.edu/~cmallery/150/gene/mol_gen.htm)

1.2.1 RNAP structure

The first structure of core RNA polymerase at high resolution was determined for *Thermus aquaticus (Taq)* (Zhang, Campbell, Minakhin, Richter, Severinov, & Darst, 1999). There is abundant functional and sequence similarity between *E. coli* RNAP and *Taq* RNAP for this structure to be directly relevant to *E. coli*. The core RNAP (Figure 1) has a structure reminiscent of a "crab claw" formed by the two largest subunits in the RNAP: β' and β. These two subunits ("pincers") form a cleft which accommodates the double-stranded DNA. The bottom of the cleft has the active center Mg^{2+} ion. (Ebright, 2000) In addition to the active-center cleft, RNAP core contains two other distinct channels: the secondary channel, which mediates access of NTP substrates to the active center cleft and the RNA-exit channel, which mediates exit of nascent RNA from the active-center cleft (Zhang, Campbell, Minakhin, Richter, Severinov, & Darst, 1999).

The two sequence-identical α subunits interact with β' and β (C-terminal domain); and with promoter sequence on the DNA and activators and repressors upstream to the promoter (N-terminal domain). The N-terminal domain (NTD) and C-terminal domain (CTD) of α subunit is connected by a flexible linker (13-36 amino acids). The NTD is involved in dimerization, interaction with RNAP and interaction with activators and repressors. The CTD can also take part in dimerization but it is a weaker interaction than NTD dimerization. The CTD is mainly responsible for interaction with upstream DNA, which can be weak non-sequence-specific interactions or strong sequence-specific interactions, and interactions with transcriptional factors. (Ebright & Busby, 1995)(Ebright, 2000)

The ω subunit, the smallest one, is located at the base of β' subunit. The ω subunit interacts with β', helps in assembly of RNAP and restores denatured core RNAP to its active form. The ω subunit is, however, not known to be essential for the function of the enzyme (Vrentas, Gaal, Ross, Ebright, & Gourse, 2005). The ω subunit simultaneously binds to the NTD and CTD of β' subunit. Assembly of *E. coli* RNA polymerase proceeds via the pathway (Ishihama, 1981):

$\alpha + \alpha \rightarrow \alpha_2 + \beta \rightarrow \alpha_2\beta + \beta' \rightarrow \alpha_2\beta\beta'$ (premature core) $\rightarrow E_{\alpha 2\beta\beta'}$ (active enzyme)

The ω subunit recruits β' to $\alpha_2\beta$ sub-assembly and keeps β' from aggregating (Ghosh, Ishihama, & Chatterji, 2001). (Ebright, 2000)

The structure of the holoenzyme was elucidated at a 2.6 Å resolution with *Thermus thermophilus* RNAP (Vassylyev, et al., 2002). The holoenzyme has the additional σ subunit which is divided into four regions (1-4) based on sequence homology in σ70 family of proteins. These regions are divided into several sub-regions (Vassylyev, et al., 2002). The σ subunit contains determinants for sequence-specific interactions with DNA. The holoenzyme is similar in structure to the core enzyme in that it retains the crab claw structure. However, due to presence of σ, the β’ pincer is extended making it more analogous to a crab claw. The conversion from core to holoenzyme is accompanied by structural changes in all domains of RNAP. (Vassylyev, et al., 2002)

The catalytic site is defined by three invariant aspartates (739, 741, 743) residues in the β’ subunit, which is located in a conserved motif: NADFDGD. These aspartate residues chelate with Mg^{2+} ions at the active site consistent with two-metal-ion mechanism proposed for all nucleic acid polymerases (Vassylyev, et al., 2002).

1.3 Promoters

RNA polymerases have to recognize promoter sequences and bind to them before they can start transcription of the genes downstream to the promoter. Bacterial promoters consist of two defining regions: -10 element and -35 element. The numbering indicates the approximate upstream position (in base pairs) of these regions from the transcription start site. The -10 region has a consensus sequence of 5’-TATAAT-3’. The -35 region has a consensus sequence of 5’-TTGACA-3’. (Harley & Reynolds, 1987)

1.4 Transcription Initiation

The bacterial RNA polymerase carries out a complex set of steps during transcription initiation (Figure 4):

(1) RNAP, along with σ, binds to two conserved hexamers (-35 and -10) in the promoter of the gene that is going to be transcribed. This association forms a closed complex (RP_c). (Vassylyev, et al., 2002)

(2) RNAP then melts a stretch of 14 nucleotides around the transcription start site which results in separation of the DNA strands and formation of an open complex (RP_o).

(3) The RNAP-promoter initial transcribing complex (RP_{itc}) begins synthesizing RNA from the transcription start site.

(4) After the complex is able to synthesize RNA longer than 9-11 nucleotides, the complex leaves the promoter DNA (promoter clearance) and translocates along the DNA synthesizing RNA as a transcription elongation complex (RD_e/TEC). At this time, σ gets released from RNAP for recycling (Darst, 2001). (Ebright, 2000)

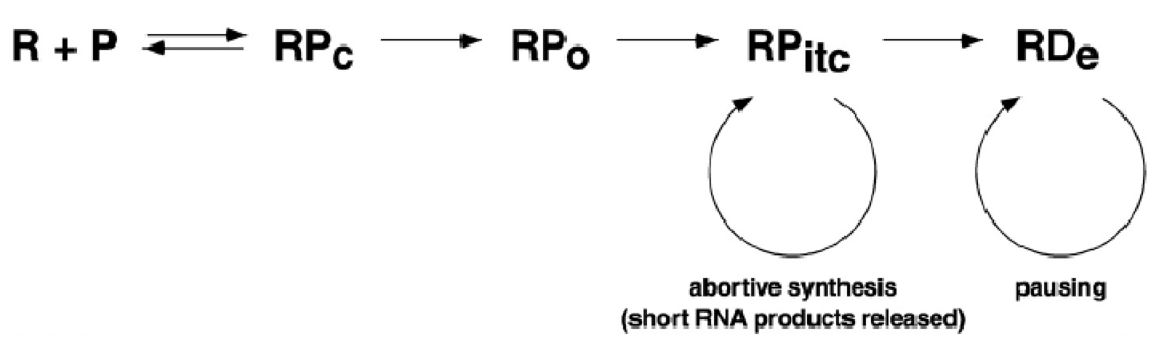


Figure 4. Transcription initiation and elongation

1.5 Nucleotide addition cycle

The nucleotide addition step, either during initiation or elongation, can be subdivided into four separate steps (Erie, Yager, & von Hippel, 1992): First, the RNAP active center translocates relative to the DNA. This is done by a mechanism called scrunching during initiation and by stepping during elongation. Next, the incoming NTP binds which brings about the formation of phosphodiester bond and release of the pyrophosphate. Each nucleotide addition step is accompanied by a conformational change in RNAP active centre: it closes when the incoming NTP is bound and opens up again when the phosphodiester bond forms. This closing and opening of the active center is mediated by a structural element called “Trigger Loop” (Wang, Bushnell, Westover, Kaplan, & Kornberg, 2006).

1.6 Ribosomal RNA (rRNA)

Ribosomal RNA (rRNA) and transfer RNA (tRNA) together make for 95% of the total RNA in a bacterial cell. Transcription of these, thus, not only constitutes a majority of RNAP activity but also is a major metabolic expense. The transcription from rRNA promoters, hence, is highly regulated by cell’s nutritional environment. (Paul, Ross, Gaal, & Gourse, 2004)

E. coli has seven rRNA operons (rrnA-H). For each of the operons, there are two promoters, rrn P1 and rrn P2, for transcription. Both promoters contain the -10 and -35 hexamers which are recognized by σ^{70}. In addition, they consist of a G+C-rich region immediately downstream from -10 hexamer, the discriminator. The P1 promoters are stronger than the P2 promoters at moderate to rapid growth rates in the exponential growth phase. This is due to the presence of UP elements, which are A+T rich regions, upstream to the -35 hexamer. These elements are responsible for an increase in the transcription activity by 300 fold. (Paul, Ross, Gaal, & Gourse, 2004)

1.6.1 Regulation of rRNA promoters

rRNA transcription is regulated in response to cell's environment (Paul, Ross, Gaal, & Gourse, 2004):

1) When cells are diluted into fresh medium, rRNA synthesis increase rapidly.
2) When cells are provided with richer or poorer medium during exponential growth, rapid changes occur in its synthesis.
3) In stationary phase, the rRNA synthesis is inhibited.
4) At steady state growth rate, the rRNA synthesis increases proportionately accounting for growth-rate dependent control.

The rapid changes in the transcription of rRNA synthesis cannot be explained based on protein factors whose synthesis cannot be fast enough. The regulation of rRNA promoters results from small molecule effectors whose concentrations can change as rapidly (Paul, Ross, Gaal, & Gourse, 2004). ppGpp is one such small molecule regulator of rRNA synthesis.

1.7 ppGpp

ppGpp, guanosine tetraphosphate (guanosine-3', 5'-(bis)pyrophosphate) (Figure 5), is an "alarmone" which affects myriad cellular functions (Figure 6) in bacteria and has the ability to induce alterations in metabolic state under stress conditions. ppGpp accumulates in cells during amino acid starvation conditions. This triggers a response in the cell called the "Stringent Response". The overall state of the cell changes from a maintenance-and-growth mode to survival mode. A recent study reported microarray profiling data indicating how ppGpp affects the metabolic state of *E. coli* under amino acid starvation (Traxler, et al., 2008). They found that ppGpp plays a major role in relaying information to genes involved in intermediary metabolism and macromolecule synthesis as ppGpp-null mutants ($ppGpp^0$) showed a lack of such a response. Cells lacking ppGpp were found to be producing more RNA and biomass, but not more protein (Traxler, et al., 2008).

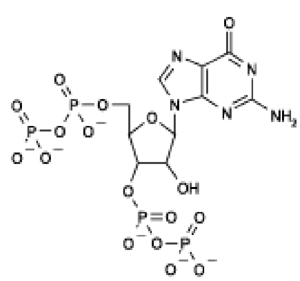


Figure 5. ppGpp

ppGpp ("magic spot") was first discovered by Cashel and Gallant in 1969 (Cashel & Gallant, 1969) when they found that it accumulates during nutrient starvation (Jain, Kumar, & Chatterji, 2006). The protein(s) responsible for the synthesis of ppGpp are RelA/SpoT which transfer a pyrophosphate moiety from ATP to GDP/GTP (Figure 7). This results in formation of ppGpp and pppGpp, together termed as (p)ppGpp (Jain, Kumar, & Chatterji, 2006).

RelA is a synthase, a 3'-kinase, whereas SpoT is a hydrolase which can also show synthesis activity under certain conditions (Chatterji & Ojha, 2001). The two enzymes produce ppGpp at basal levels when the cells are not under stress. This results in some inhibition of rRNA synthesis although ppGpp does not regulate growth-rate dependent rRNA synthesis (Barker, Gaal, Josaitis, & Gourse, 2001). During amino acid starvation, the tRNA in the ribosomal A-site during translation is uncharged. This stalls the ribosomes and RelA is able to recognize them, associate and produce ppGpp (Figure 8) (Magnusson, Farewell, & Nystrom, 2005). SpoT may be responsible for ppGpp accumulation in stress conditions apart from amino acid starvation (Magnusson, Farewell, & Nystrom, 2005).

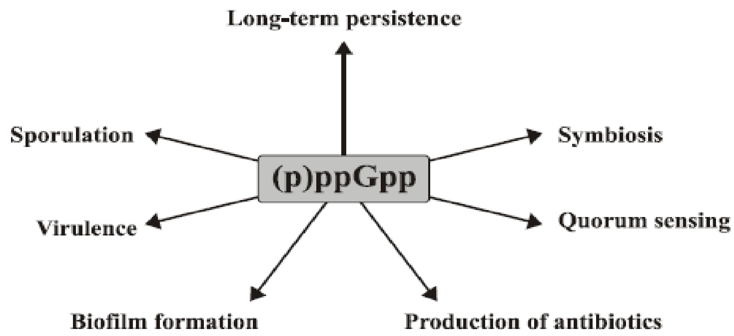


Figure 6. Myriad functions of ppGpp (Jain, Kumar, & Chatterji, 2006)

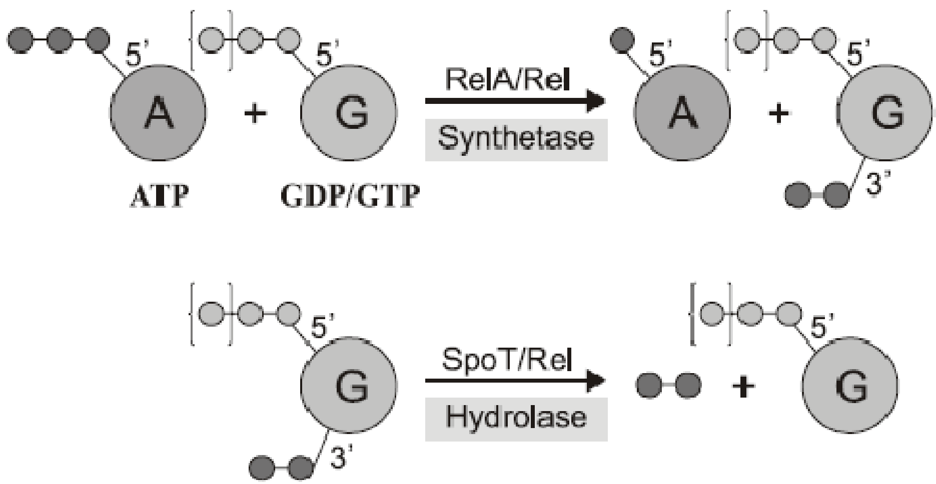


Figure 7. Mechanism of synthesis of (p)ppGpp (Jain, Kumar, & Chatterji, 2006)

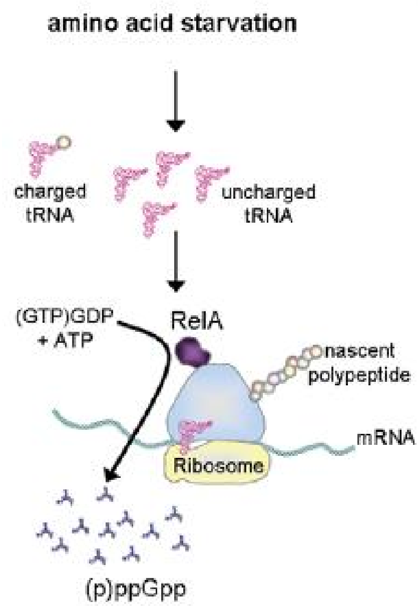


Figure 8. Production of ppGpp (Srivatsan & Wang, 2008)

1.7.1 Stringent response

Bacteria have to survive through adverse environmental conditions varying from paucity of food resources or presence of inhibitory molecules like drugs. They are able to do so by invoking changes at genetic level which change their overall metabolic state. A special class of such a response under nutritional stress is called Stringent Response (Jain, Kumar, & Chatterji, 2006). During nutritional stress bacteria have to up-regulate certain functions like protein degradation and amino acid synthesis (Barker, Gaal, Josaitis, & Gourse, 2001) and down-regulate certain others like protein and nucleic acid synthesis (Figure 9). The effect of positive regulation on amino acid promoters is most likely indirect because the effect can only be observed *in vivo* (Barker, Gaal, Josaitis, & Gourse, 2001)(Barker, Gaal, & Gourse, 2001). It has been shown using the rrnB P1 promoter that ppGpp decreases transcription from this promoter *in vivo* and *in vitro* (Barker, Gaal, Josaitis, & Gourse, 2001). Since the production of ribosomal RNA is largely reduced by ppGpp during stringent response, it has a direct effect on the growth of the organism (Chatterji & Ojha, 2001). This also allows the cells to spend their remaining energy to express stress genes to survive the adversity (Toulokhonov, Shulgina, & Hernandez, 2001). After the adaptation to the stress conditions is over, the ppGpp levels reduce and growth resumes. Failure to reduce ppGpp levels can severely affect cell viability (Toulokhonov, Shulgina, & Hernandez, 2001).

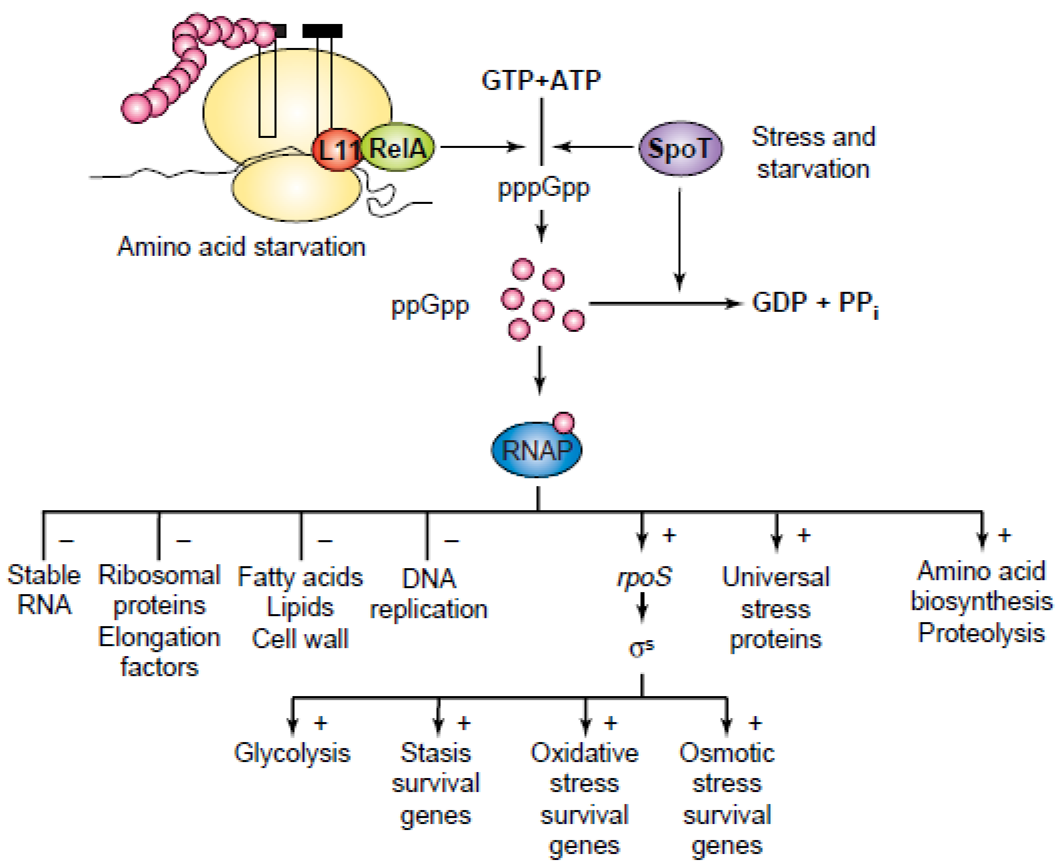


Figure 9. The Stringent Response: Effects of ppGpp on global transcription (Magnusson, Farewell, & Nystrom, 2005)

1.7.2 Proposed mechanisms of ppGpp action

The mechanism of ppGpp action has been a disputed issue. Different explanations have been suggested, any one or any combination of them may be at work at the same time. The negatively regulated promoters have certain characteristic features like the presence of GC-rich discriminator region between the -10 region and the transcription start site which allows only sub-optimal interaction with σ of RNAP, formation of short-lived open complexes etc. (Srivatsan & Wang, 2008). On the other hand, the positively regulated promoters (amino acid biosynthesis) have an A-T rich discriminator region which allows optimal binding with σ of RNAP (Srivatsan & Wang, 2008). The following mechanisms have been proposed for direct negative regulation by ppGpp:

1) **Open complex stability** (Figure 10)**:** rrnB P1 promoter was shown to be sensitive to toxic effects of ppGpp *in vitro*. The mechanism proposed was that ppGpp affects the half-life of the open complexes. It was suggested that ppGpp was affecting all promoters, but only promoters, like rrnB, which have orders of magnitude lesser half-life than unaffected promoters, like amino acid promoters, are able to show the effect. This means that the difference in positive and negative regulation is not due differential interaction with RNAP but due to kinetics of the promoter-RNAP association. It was also shown that ppGpp does not affect initial binding of RNAP to the promoter. (Barker, Gaal, Josaitis, & Gourse, 2001)

2) **Promoter Clearance:** Bacteriophage λP_R promoter, which forms long-lived promoters, was shown to be affected by ppGpp due to failure of the polymerase to escape the promoter for elongation (Potrykus, Wegrzyn, & Hernandez, 2002). They observed here too that ppGpp affects open complex stability, however, to a much lesser degree not accountable for the observed inhibition.

3) **Open Complex formation:** The rate of open complex formation has also been shown to be affected by ppGpp (Magnusson, Farewell, & Nystrom, 2005).

4) **Pausing:** ppGpp has been shown to increase pausing during transcriptional elongation (Jores & Wagner, 2003).

5) **Competition between ppGpp and the NTP substrates:** It has also been suggested that ppGpp competes with the NTP substrates at the active centre of RNAP (Jores & Wagner, 2003)(Artsimovitch, et al., 2004).

6) **Base pairing with cytosines:** ppGpp could base pair with cytosines in the non-template strand of promoter DNA (Artsimovitch, et al., 2004).

Besides the direct effects of ppGpp listed above, various indirect ways in which ppGpp can affect the regulation have been suggested. One study suggests that during growth arrest, there is an increased availability of free RNAP which can result in decreased stability of open complexes. This would also cause the positively regulated promoters by ppGpp to be able to recruit RNAP which they are otherwise poor at. However, this has been contradicted by another study which showed that the ppGpp increases pausing of RNAP which would sequester core RNAP and make them unavailable for positively regulated promoters. Another report indicates that the “stringent promoters”, like rrn promoters, are especially sensitive to reduced levels of RNAP as they require high concentrations of RNAP for maximum activity. It has also been found that amount of σ^{70} bound to RNAP core decreases during stringency, though it is not clear if this is caused by ppGpp. Other sigma factors, like σ^{S}, σ^{32}, σ^{54}, have been shown to be dependent on ppGpp for their production. Thus, during stringency when these sigma factors are needed, they would compete with σ^{70} for the RNAP core (Figure 10). (Magnusson, Farewell, & Nystrom, 2005)

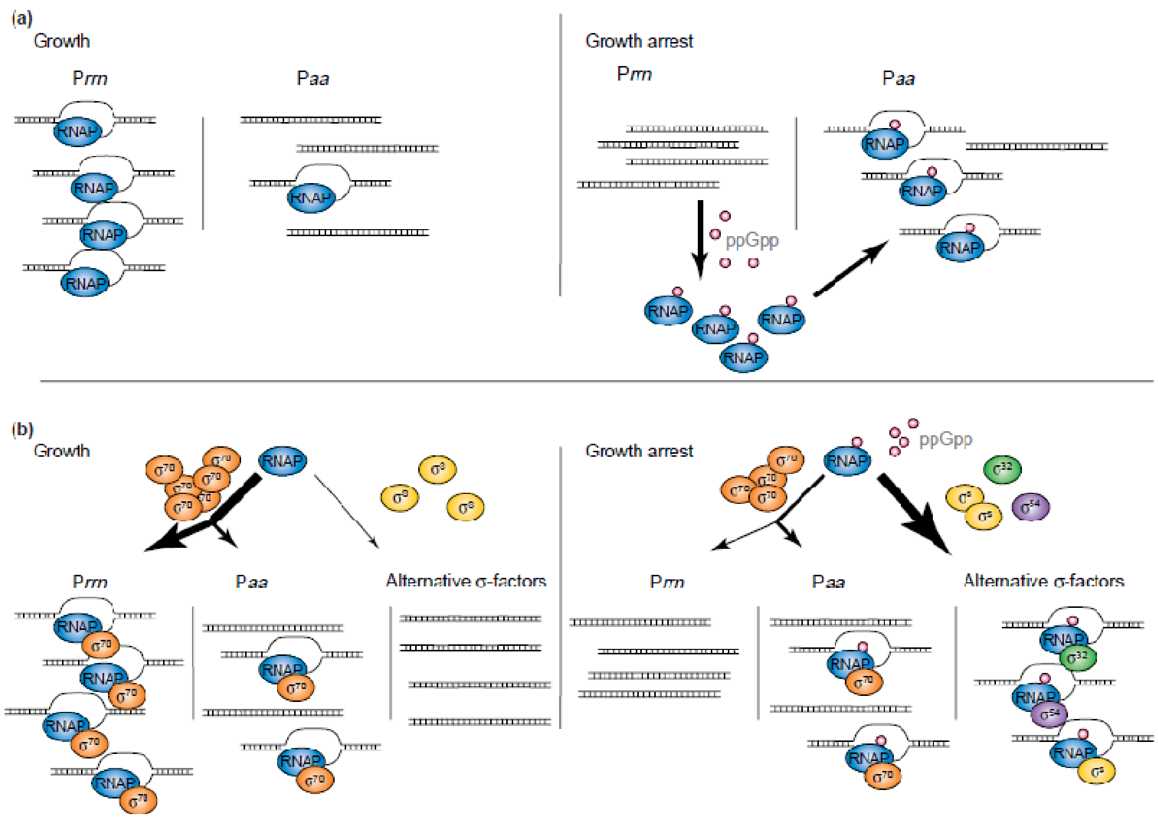


Figure 10. Models for regulation of RNAP levels in by ppGpp. (a) Unstable open complexes (b) Competition from alternate sigma factors (Magnusson, Farewell, & Nystrom, 2005)

1.7.3 ppGpp and RNA polymerase

It has been suggested in several studies that ppGpp directly interacts with RNAP (Toulokhonov, Shulgina, & Hernandez, 2001). Mutants of *E. coli* lacking in relA and spoT gene, unable to synthesize ppGpp (relaxed response, characterized by accumulation of rRNA during starvation (Magnusson, Farewell, & Nystrom, 2005)), cannot grow without amino acid supplements in the media. It has also been shown that these mutants are rescued by suppressor mutations in RNAP subunit genes: 97% in *rpoB* (β subunit) and *rpoC* (β’ subunit) genes and 3% in *rpoD* gene (σ subunit) (Chatterji, Fujita, & Ishihama, 1998).

In one study, ppGpp was shown to bind to the C-terminal region of β subunit with cross-linking studies (Chatterji, Fujita, & Ishihama, 1998). Another study, using photo-cross-linkable derivative of ppGpp, showed that, ppGpp has a binding site near the N-terminus of β’ subunit of RNAP. This region is disordered on the crystal structure of RNAP from *Taq*. The region has been implicated in stable DNA association with the RNAP and is in close proximity to the exiting nascent RNA. A determinant for σ^{70} has also been found to be associated with this region. If this is indeed the binding site of ppGpp, it suggests the mechanism for ppGpp action on RNAP is most likely destabilization of open complexes (Toulokhonov, Shulgina, & Hernandez, 2001).

The N-terminal of β’ and C-terminal of β are spatially close by (Zhang, Campbell, Minakhin, Richter, Severinov, & Darst, 1999), which might suggest that both the above studies may be right and differing in their results only because of the difference in the cross-linking molecule used (Toulokhonov, Shulgina, & Hernandez, 2001).

A 2.7 Å resolution structure of *Thermus thermophilus* RNAP holoenzyme complexed with ppGpp reveals that the ppGpp binding site is near the active center and that it can bind in two different orientations at this site. The site, specifically, is located in the substrate (NTP) channel next to the active center (Artsimovitch, et al., 2004). The residues that ppGpp interacts with according to this model are highly conserved residues in *E. coli* RNAP. One of these residues was an Asparagine at 458 of the β' subunit of *E.coli* RNAP.

However, this finding about the ppGpp binding site has been recently challenged (Vrentas, et al., 2008). The site identified for ppGpp with *Thermus thermophilus* was speculated because: (a) A study had reported that the same site can accommodate a number of negatively charged molecules suggesting that the site could be irrelevant for ppGpp's regulatory role; (b) Another study had reported that *Thermus thermophilus* RNAP does not directly interact with ppGpp (Kasai, Nishizawa, Takahashi, Hosaka, Aoki, & Ochi, 2006) (c) The polymerases were made with over-expression of all subunits but ω. RNAP lacking ω have been shown to be irresponsive to ppGpp. They specifically mutated all the residues identified by the above mentioned crystal structure in *E. coli* RNAP and found that those mutants were not resistant to toxic effects of ppGpp in presence or absence of DksA. They did not alter the effect of ppGpp on promoter complex lifetime. They also reported that ppGpp does not compete with NTP's during transcriptional initiation and ppGpp does not affect *Thermus thermophilus* RNAP *in vitro*. (Vrentas, et al., 2008)

1.8 DksA

The effects of ppGpp have been observed *in vitro*, but the extent of inhibition was much less compared to the effects of ppGpp *in vivo*. DksA, a 151-amino acid protein coded by the dksA gene of *E. coli* (Kung & Craig, 1990), was shown to be absolutely required for regulation of rRNA expression (Figure 11) (Paul, et al., 2004). It was also shown through structural modeling that DksA mediates the effect of ppGpp in transcription inhibition (Perederina, et al., 2004). DksA was shown to be a non-essential gene (Kung & Craig, 1990). However, deletion of *dksA* has pleiotropic effects: gene expression, chaperonin function, cell division, amino acid requirements, quorum sensing, phage sensitivity and virulence (Paul, et al., 2004).

With regards to its interaction with RNAP, it was proposed that DksA might play a role in reducing the lifetime of open complex even in absence of ppGpp and that it helps ppGpp to interact with RNAP (Paul, et al., 2004). It was earlier thought that ppGpp does not directly affect amino acid promoters as the effect of ppGpp was not seen *in vitro* (Barker, Gaal, & Gourse, 2001). After discovery of DksA's role, it was reported that ppGpp can activate amino acid promoters *in vitro* only in presence of DksA by affecting the open complex formation. However, that did not rule out the possibility that some promoters may be controlled indirectly by ppGpp/DksA. Since DksA concentrations inside the cell remain fairly constant in all conditions and ppGpp concentrations vary rapidly, DksA's role is of a coregulator with ppGpp being the one responding to internal signals (Paul, Berkmen, & Gourse, 2005). It was shown in a study that the ω subunit of RNAP is not required for the stringent response *in vivo* (Jain, Kumar, & Chatterji, 2006). This was later explained that RNAP lacking ω subunit was able to show stringent response only in presence of DksA (Vrentas, Gaal, Ross, Ebright, & Gourse, 2005).

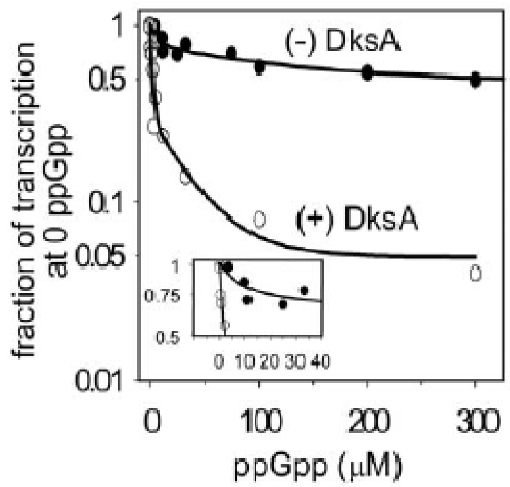


Figure 11. Effect of DksA on rrnB P1 promoter. Concentration of DksA used: 190 nM (Paul, et al., 2004)

Till date the mechanism of DksA or ppGpp action remains obscure and disputed. DksA-null mutants (ΔDksA) and ppGpp-null mutants ($ppGpp^0$) do not show a complete overlap. The two strains show a difference in their amino acid requirements suggesting that may be ppGpp acts alone in some cases and with DksA in others (Magnusson, Farewell, & Nystrom, 2005). The effect of DksA has not been characterized *in vitro*.

The overall structure of DksA resembles that of transcriptional factors GreA which is remarkable considering the fact that they have no sequence similarity (Perederina, et al., 2004). Based on this structural similarity, it is very likely that DksA binds at secondary channel of RNAP just like GreA thereby hindering the access to incoming NTPs although not completely blocking NTP entry and possibly stabilizing ppGpp-RNAP complex (Perederina, et al., 2004)(Paul, et al., 2004).

1.9 ppGpp and Pathogenic bacteria

Pathogenic bacteria like Mycobacterium tuberculosis are able to inflict Tuberculosis on its host because it is able to survive in stressed conditions (Parrish, Dick, & Bishai, 1998). ppGpp has been shown to play a direct role in virulence and persistence of pathogenic bacteria. For example, ppGpp regulates virulence in *Salmonella typhimurium* and *Pseudomonas aeruginosa*; decreases biofilm formation in *E. coli* relA spoT mutants (Godfrey, Bugrysheva, & Cabello, 2002). In a study they immunized mice with ppGpp^0 strain of *Salmonella* and found that the mice were protected from the wild-type *Salmonella* at a dose 10^6-fold above the established LD_{50} values (Potrykus & Cashel, 2008).

Chapter 2: Objective and Strategy

The ppGpp-induced stringent response still has many questions unanswered. From a basic research point of view, it is essential to understand the mechanism of ppGpp action and the myriad processes it can regulate within a cell. The ppGpp and DksA binding sites remains unknown. How they interact with RNAP still is a mystery and as magical as it ever was.

The aim of this study is to determine where ppGpp and DksA bind and how they interact with RNAP. This will help in determining the mechanism of their action and how they regulate the global metabolic state of the cell. The understanding of these mechanisms will also help to develop better small-molecule anti-bacterial compounds targeting the RNAP. The bacteria constantly evolve and become resistant to the anti-bacterial drugs. Thus, developing new antibacterial compounds is essential.

The proposal is to isolate mutants to ppGpp and DksA by performing genetic screens of random mutagenesis libraries of RNAP genes (*rpoB*: β subunit; *rpoC*: β' subunit; *rpoD*: σ subunit). Specifically, this study has looked at *rpoC* gene which encodes for the β' subunit of the RNAP. The mutants will be isolated in a *dksA*+ background, which will yield mutants resistant to toxic effects of both ppGpp and DksA, and in *dksA*- background which will yield mutants resistant to toxic effects of only ppGpp.

2.1 The System

E. coli is the obvious choice for organism for this study. The strain used will be transformed with a plasmid containing the *relA* gene under control of arabinose promoter to over-express ppGpp and mimic the amino acid starvation conditions (Schreiber, Metzger, Aizenman, Roza, Cashel, & Glaser, 1991). The *relA* gene used here is a truncated version which can be expressed constitutively upon induction with arabinose. This expression system is able to produce high levels of ppGpp which results in slowing of growth and inhibition of *rrn* P1 promoter (Schreiber, Metzger, Aizenman, Roza, Cashel, & Glaser, 1991). This approach is akin to the one used in (Tedin & Bremer, 1992) where they isolated a *rpoB* mutant.

The strains carrying the *relA* plasmid will then be transformed with plasmids containing randomly mutagenized *rpoC* gene (Figure 12) (encoding the β' subunit of RNAP) (Zhou, Zhang, & Ebright, 1991)(Mukhopadhyay, Sineva, Knight, Levy, & Ebright, 2004). The mutagenesis will be error-prone-PCR based using Pfu DNA polymerase. The *rpoC* gene will be under the control of the *lac* promoter which can be induced by IPTG (isopropyl-1-thio-8-D-galactopyranoside).

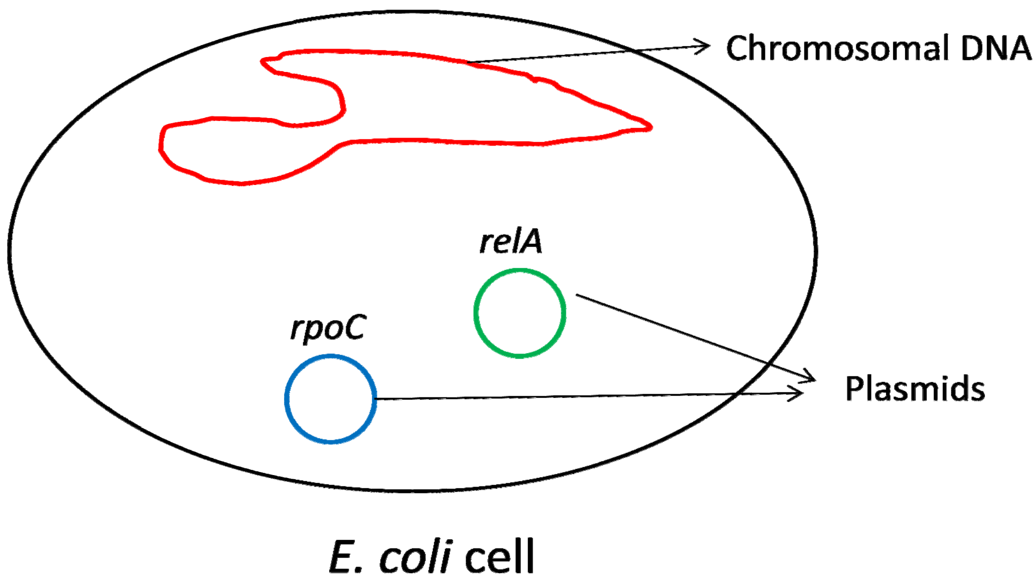


Figure 12. An *E.coli* cells carrying the *relA* and *rpoC* plasmids.

2.2 Primary Screen

The cells carrying the mutagenized *rpoC* gene will be grown on media containing IPTG and appropriate antibiotics with arabinose and without arabinose. At the same time, wild-type *rpoC*-plasmid-containing cells will also be grown in similar media conditions. The wild-type cells show a distinct reduction in their colony size when arabinose is present in media (implying that ppGpp is over-expressed) as compared to when no arabinose is present. The mutants to toxic effects of ppGpp will show a larger colony size in presence of arabinose than wild-type. These mutants will be isolated, sequenced and characterized.

The secondary screen would include linking the mutant phenotype to the *rpoC* plasmid and linking it to the *rpoC* gene. Linkage to *rpoC* plasmid will confirm that a random mutation in the *relA* plasmid is not responsible for the mutant phenotype. Once this is established, selected mutants will be sequenced. The previous linkage test was to establish that the mutation is on the *rpoC* plasmid but it does not confirm that the mutation is on the *rpoC* gene. To confirm this linkage to *rpoC* gene will also be established.

The effects of selected mutants will then be quantified by β-galactosidase assay, an *in-vivo* assay. The β-galactosidase assay will be based on transforming the mutant *rpoC* genes into a strain having the β-galactosidase gene in addition to the *relA* plasmid, growing the cells with and without arabinose, and then performing biochemical assays to measure the β-galactosidase activity. Simultaneously, wild-type *rpoC* gene will be subject to similar protocol to make comparisons.

Mutants which pass all the above screens will then be studied for *in-vitro* effects. The mutants will be purified and assayed for activity using fluorescence-based abortive initiation assay—an assay which characterizes activity of RNAP on transcription initiation. The assay will be performed in presence of ppGpp, DksA and ppGpp+DksA.

Chapter 3: Methods

3.1 Primary Screen

3.1.1 Plasmids and Strains

The strain of *E. coli* used for the primary screen is BW27784 pBADrelA'(1-394) and BW27784 pBADrelA'(1-394) *dksA::tet* (Khlebnikov, Datsenko, Skaug, Wanner, & Keasling, 2001). The former strain is in the DksA+ background and the latter in the DksA- background. These strains are engineered to allow homogeneous induction of arabinose-dependent promoters in a concentration-of-arabinose-dependent manner. These cells contain the pBADrelA'(1-394) (ori-p15A; Cm^r) plasmid which is a derivative of pBAD33 plasmid (Figure 13) (Guzman, Belin, Carson, & Beckwith, 1995). The plasmid carries the truncated, constitutively active *relA* gene (Brown, Gentry, Elliott, & Cashel, 2002) under the control of arabinose-dependent pBAD promoter and chloramphenicol resistance gene (Cm^r).

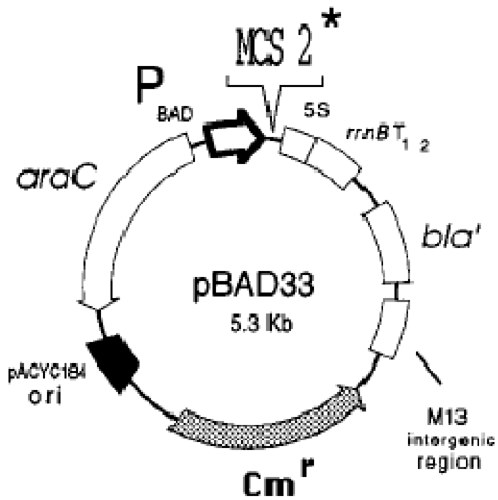


Figure 13. Map of pBAD33 vector. araC: regulator of P_{BAD}; pACYC184 ori: origin of replication; Cm^r: Chloramphenicol resistance gene; MCS: multiple cloning site (Guzman, Belin, Carson, & Beckwith, 1995)

The plasmid carrying the β' subunit encoding *rpoC* gene is the pRL663 plasmid (Wang, Meier, Chan, Feng, Lee, & Landick, 1995). The gene is under control of the lac promoter and can, therefore, be induced using IPTG. The plasmid also carries ampicillin resistance gene. Random mutagenesis was done on the *rpoC* gene using PCR-based methods as used in (Mukhopadhyay, Sineva, Knight, Levy, & Ebright, 2004).

3.1.2 Preparation of Competent cells

Protocol was taken from (Chung, Niemela, & Miller, 1989). Fresh overnight culture of bacteria grown at 37^{o}C was 1:100 diluted into LB liquid broth and incubated at 37^{o}C until the cell density reached an OD_{600} of 0.3-0.4. Cells were spun down at moderate speed and resuspended in 1/10th of original volume of ice-cold TSS buffer (1X LB, 10%

PEG (polyethylene glycol), 5% DMSO, 50 mM Mg^{2+}, pH 6.5; filter-sterilized) and mixed gently. The cells were flash-frozen in dry ice/ethanol and stored at -80°C.

3.1.3 Transformation

Library of pRL663 plasmid containing randomly mutagenized *rpoC* gene were transformed into BW27784 pBADrelA'(1-394). At the same time, wild-type pRL663 was also transformed. For transformation, 50 µL of competent cells were mixed with 1 µL of DNA in a tube and was incubated on ice for 30 minutes. The tube was heat-shocked at 42°C for 40 seconds, incubated in ice for 5 mins, incubated at room temperature for 2 mins. 500 µL of LB Media (1% w/v Bacto-tryptone, 0.5% w/v yeast extract, 1% w/v NaCl, made with MilliQ® water) was added to tube and cells were let to grow for 1 hr at 37°C. IPTG to a final concentration of 1 mM was added and cells were let to grow for another 1.5 hrs at 37°C. 40-60 µL of cells were plated on LB-Agar (1.5%) plates containing 100 µg/ml Ampicillin, 30 µg/ml chloramphenicol, 1 mM IPTG (referred to as LB-Agar-Amp-Cm-IPTG henceforth, unless otherwise stated) with and without Arabinose (0.2% w/v). The plates were incubated for 19 hrs at 37°C. The colonies which were larger than wild-type on arabinose plates were selected and used for further screening.

Same procedure was repeated with BW27784 pBADrelA'(1-394) *dksA::tet* strain.

3.2 Linkage and sequencing

The selected mutants were first isolated by streaking them on fresh LB-Agar-Amp-Cm-IPTG-0.2% arabinose plates along with wild-type colony to verify that a single colony is picked up and that the phenotype is clearly visible.

Linkage to rpoC plasmid

These colonies were then inoculated in 5 ml LB-Amp-Cm liquid cultures O/N for extracting the plasmid from these strains. Plasmid was extracted using Qiagen® miniprep kit. The plasmid prep contains both plasmids: *relA* and *rpoC*. To verify that the *rpoC* plasmid was responsible for the trait and not the relA plasmid, the relA plasmid was digested with restriction enzyme. The enzyme used was Mfe I (New England Biolabs®), which cuts only the *relA* plasmid and not the *rpoC* plasmid. The reaction was carried out for 1 hr at 37^{o}C in 50 µL volume with 5 µL of 10X buffer (provided with the enzyme), 5 µL plasmid DNA, 1 µL enzyme, 39 µL MilliQ® water. The resulting digest was transformed again into BW27784 relA' cells (protocol as described above), grown on LB-Agar-Amp-Cm-IPTG and LB-Agar-Amp-Cm-IPTG-0.2% arabinose and the mutant phenotype verified. (Figure 14)

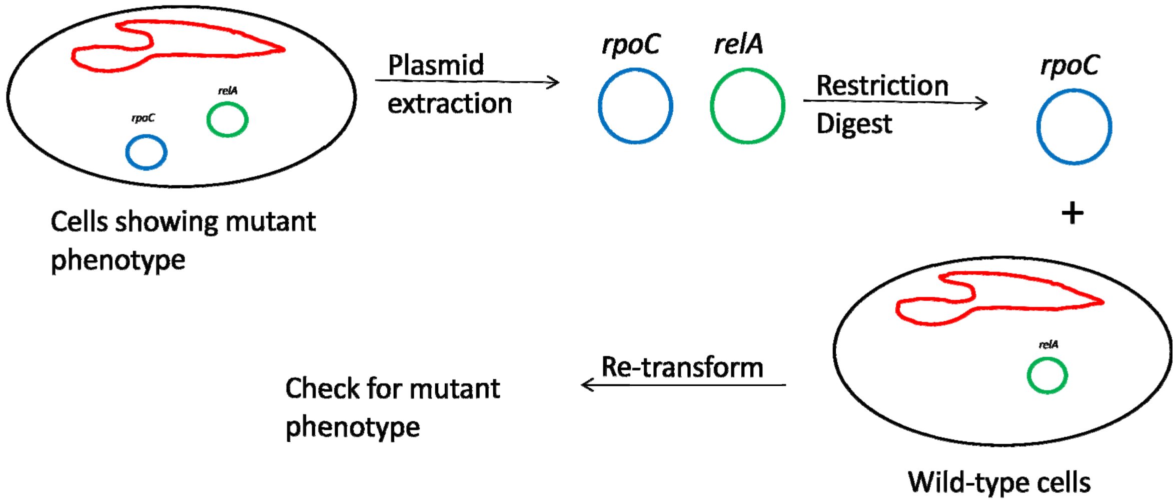


Figure 14. Linkage to *rpoC* plasmid.

Sequencing

All the mutants which passed through these screens were then sequenced. 8 primers were designed to cover the entire *rpoC* gene. The primers and DNA samples were sent to University of Washington or GeneWiz for sequencing. Only mutants which had single amino acid change were considered for further analysis.

Linkage to rpoC gene

The purpose of linkage is to verify that the mutation present on *rpoC* gene is the one causing the mutant phenotype and not a random mutation that may have arisen in the rest of the plasmid. For verifying linkage to *rpoC* gene, the mutant plasmid was digested with Hind III and Xba I to excise out the *rpoC* gene. The wild-type plasmid was also digested similarly to recover the backbone (plasmid without the *rpoC* gene). However, the fragment sizes after the digest were very close by to be separated on agarose gel. So the mutants were further digested with NdeI to further cut the backbone and wild-type plasmid was digested with MscI to cut the *rpoC* gene. The digests were run on an agarose gel (0.8%) by electrophoresis in TAE buffer (40 mM Tris acetate and 1 mM EDTA). For mutants, the *rpoC* gene fragment and for wild-type, the backbone was gel purified using Qiagen® gel purification kit. To ensure that the backbone does not self-ligate, it was treated with Antarctic Phosphatase (NEB®) for 20 mins at 37^{o}C and heat-inactivated for 5 mins at 65^{o}C. The ligation was carried out for 15-20 mins at room temperature in 20 µL volume containing 9 µL of *rpoC* fragment, 3 µL of plasmid backbone (3:1 :: insert:vector), 2 µL of 10X buffer, 1 µL of T4 DNA ligase, 5 µL of water. The ligated product was transformed into BW27784 cells (no relA) and grown on LB-Agar-Amp plates. Colonies were then used to obtain plasmids (as described before), and this plasmid was transformed into BW27784 relA' cells to again check for mutant phenotype (as described above). For mutants which still showed the mutant phenotype it was concluded that the phenotype was linked to the mutation in the *rpoC* gene. (Figure 15)

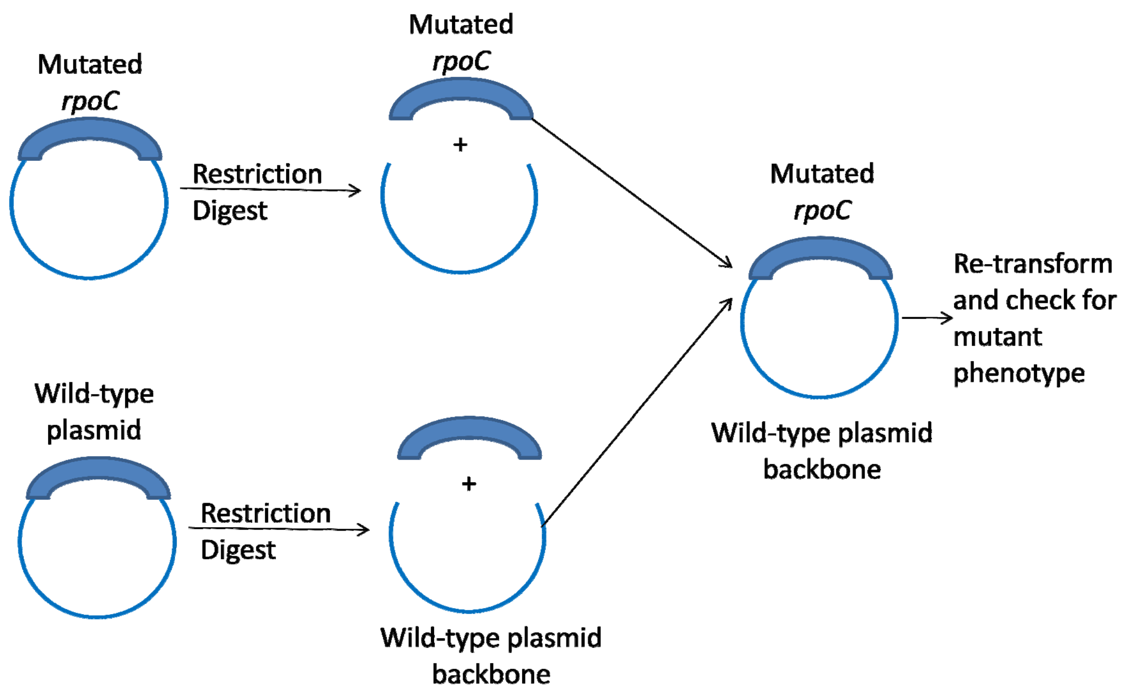


Figure 15. Linkage to *rpoC* gene.

3.3 Testing in DksA- strain

All the mutants that were isolated from the primary screen (in DksA+ background) and which passed the subsequent tests were tested in BW27784 pBADrelA'(1-394) *dksA::tet* strain. This was to differentiate between mutants of ppGpp from mutants of DksA. If the mutant after transformation in this strain retains the mutant phenotype, then it most likely is a ppGpp mutant. On the other hand, if the mutant phenotype is not retained, then it most likely is a DksA mutant. The transformation of mutant plasmid and wild-type plasmid was performed in this strain and the colony sizes compared under conditions of ppGpp overproduction. The methods were analogous to as previously described.

3.4 β-galactosidase assays

This assay (Miller, 1992) was used to quantify the effects of mutants. The strain used was RLG4996 (VH1000 1*rrnB* P1 (-46/+1)-*lacZ*)(Barker & Gourse, 2001) after transforming it with relA plasmid. This strain contains lacZ gene, which codes for the enzyme β-galactosidase, fused with rrnB P1 promoter. The plasmids for mutant *rpoC* genes are transformed into this strain and plated on LB-Agar-Amp-Cm plates also containing X-gal (5-bromo-4-chloro-3-indolyl- beta-D-galactopyranoside). X-gal is acted upon by β-galactosidase and the product appears blue. The colonies which appeared blue were picked up for the assay.

Separate colonies for each mutant and wild-type were grown O/N for 14-18 hours at 37^{o}C in 5 ml LB-Amp-Cm-IPTG liquid cultures without and with arabinose (0.2% and 1%) in triplicates. 100 µL from O/N cultures was transferred into identical media and allowed to grow till the cell density reached OD_{600} of 0.28-0.7 (log phase) and the OD was recorded. The cells were chilled in refrigerator once they reached the correct OD to stop their growth. 100 µL of culture was mixed with 900 µL of Buffer Z (pH 7; 60 mM $Na_2HPO_4.7H_2O$, 40 mM $NaH_2PO_4.H_2O$, 10 mM KCl, 1 mM $MgSO_4.7H_2O$, with 50 mM β-mercaptoethanol added just before use; stored in refrigerator) and 40 µL of chloroform and 20 µL of 0.1 % SDS (sodium dodecyl sulfate) were added to it. The mixture was vortexed at high speed for 10 seconds and the tube was placed at 28^{o}C for 5 minutes.

200 µL of 4 mg/ml ONPG (*ortho*-Nitrophenyl-β-galactoside) was added to the tube, mixed well and start time was recorded. ONPG is another substrate for β-galactosidase which develops a yellow color when it is acted upon. After a fair amount of color developed, 500 µL of 1 M Na_2CO_3 was added to stop the reaction and stop time was recorded. 1 ml of the sample was used to measure absorbance at 420 nm and 550 nm using Buffer Z as blank. The miller units of β-galactosidase were calculated by the following formula:

$$\boldsymbol{Units\ of\ enzyme = \frac{(OD_{420} - 1.75 \times OD_{550})}{time \times volume \times OD_{600}} \times 1000}$$

where time is the time of the reaction, volume is the sample volume taken (100 µL).

The results were averaged for the triplicates. For each mutant, the ratio of activity with arabinose to activity without arabinose was calculated and compared with that of wild-type. For mutants it is expected that the ratio will be close 1 or at least higher than that of wild-type.

Glycerol stocks (in 15% glycerol) of each mutant and wild-type were made with the O/N cultures and stored at -80^{o}C for any future experiments.

3.5 Purification of RNA polymerase

The method used was as described in (Niu, Kim, Tau, Heyduk, & Ebright, 1996).

Growing cells

pRL663 which carries the *rpoC* gene with a His-tag is used for purification. The plasmid was transformed into 397c[*rpoC*^{ts}*397 argG thi lac* (λcI$_{857}$h$_{80}$S$_{t68}$d*lac*$^{+}$)(Christie, et al., 1996)] competent cells and grown at 37^{o}C. For each purification, colonies were grown O/N at 37^{o}C in 20 ml of 4X LB medium containing ampicillin. 2 liters of same media was inoculated with the O/N culture and allowed to grow till the OD$_{600}$ reached 0.4. IPTG was added to a final concentration of 1 mM and the culture was allowed to grow for 3 more hrs. The cells were pelleted down at 4000 rpm for 5 minutes.

Lysis and Ammonium sulfate precipitation

The pellet was re-suspended in 30-40 ml Grinding buffer (0.05 M Tris (pH 7.9), 5% glycerol, 2 mM EDTA, 0.233 M NaCl, 1 mM Dithiothreitol (DTT), 23 µg/ml phenylmethanesulphonylfluoride (PMSF), 0.25% desoxycholate) and lysed using French press. Equal volume of TGED (0.01 M Tris (pH 7.9), 5% glycerol, 0.1 mM EDTA, 1 mM DTT)+0.2M NaCl was added, mixed well and centrifuged for 45 minutes at 8000 rpm at 4^{o}C. The supernatant volume was measured; Polymin P (pH 7.9) was added to a final concentration of 0.6-0.8%, stirred and centrifuged for 20 minutes at 8500 rpm at 4^{o}C. The pellet was re-suspended with 25 ml of TGED+0.5 M NaCl and centrifuged for 15 mins at 8500 rpm at 4^{o}C. This step was repeated once again. The pellet was re-suspended with 30 ml of TGED+1 M NaCl and centrifuged for 15 mins at 8500 rpm at 4^{o}C. The supernatant was collected, precipitated with ammonium sulfate, centrifuged for 25 mins at 13000 rpm at 4^{o}C and the pellet was saved.

Nickel columns

4 ml Ni-NTA columns were made using Qiagen® resin. The column was equilibrated with 5 column volumes (CV) of TGD+0.5 M NaCl (0.01 M Tris (pH7.9), 5% glycerol, 1mM DTT). The pellet from ammonium sulfate precipitation was re-suspended in 10 ml TGD+0.5 M NaCl, centrifuged for 5 mins at 5000 rpm at 4^{o}C and supernatant collected. The supernatant was applied to the column and the flow-through was collected. The column was washed with TGD+0.5 M NaCl for 4-5 CV and flow-through was collected. The column was sequentially washed with TGD+0.5M NaCl+10 mM imidazole, TGD+0.5M NaCl +20 mM imidazole, TGD+0.5M NaCl +40 mM imidazole and TGD+0.5M NaCl +100 mM imidazole and fractions were collected for each of them. The fractions with 40 mM imidazole and 100 mM imidazole were pooled, diluted with TGD and centrifuged for 5 mins at 5000 rpm at 4^{o}C.

FPLC (fast protein liquid chromatography)

The sample was run through anion exchange column (MonoQ) FPLC with a salt gradient of 20 CV between 0.3 M NaCl and 0.5 M NaCl (Buffer: TGED). The chromatogram was analyzed to locate peaks and a SDS-PAGE gel (Polyacrylamide Gel Electrophoresis; see below for protocol) was run to verify the presence of RNA polymerase in the fractions.

Two different peaks show in the chromatogram corresponding to the holo and the core RNAP. The right fragments, separately for core and holoenzyme, were pooled together, dialyzed and concentrated with storage buffer (TGED+0.3 M NaCl, 50% glycerol). Quantification was done using previously quantified samples of RNAP.

SDS-PAGE

4-20% Tris-HCl precast gels were used (Bio-Rad®). Gel-loading buffer-dye: 50 mM Tris-HCl (pH 6.8), 100 mM DTT, 2% SDS, 0.1% Bromophenol Blue, 10% glycerol. Gel running buffer: 5 mM Tris, 50 mM glycine (pH 8.3), 0.02% SDS. 20 µL sample + 5 µL of buffer-dye was loaded onto gels. The molecular weight standard used was Prestained protein standards from Bio-Rad®.

3.6 Purification of DksA

The method from (Paul, et al., 2004) was used for DksA purification. His-tagged DksA was purified using the plasmid pRLG7067 transformed into BL21 (DE3) cells. The cells were grown at 37^{o}C till the OD_{600} reached 0.2, induced with 1 mM IPTG (Gold Biotechnology) for 2 hrs. The cells were pelleted and re-suspended in 40 mM Tris-HCl (pH 8), 300 mM NaCl, 5 mM imidazole, 2 mM β-mercaptoethanol and 0.1 mM PMSF. The cells were sonicated and centrifuged at 15000g for 30 mins. The supernatant was applied to equilibrated Nickel column, washed with 50 volumes of same buffer as above without PMSF and the protein was eluted with same buffer as above with 150 mM imidazole. The fractions were further purified using anion exchange (MonoQ) FPLC via a

15 CV salt gradient from 50 mM NaCl to 1 mM NaCl. SDS-PAGE was done to check the fractions for presence of DksA and the right fractions were pooled, dialyzed and concentrated with storage buffer (25 mM Tris (pH 8), 100 mM NaCl, 50% glycerol and 2 mM β-mercaptoethanol). DksA was quantified using previous samples of DksA of known concentration.

3.7 *In-vitro* Transcription

Fluorescence-based abortive initiation assay was performed as described in (Mukhopadhyay, Sineva, Knight, Levy, & Ebright, 2004). The assay tests for activity of polymerase to perform abortive initiation (see Figure 4). The reaction consists of double stranded DNA consisting of the promoter, the purified RNA polymerase, a fluorescent probe, initiating nucleotides (ApA) and Heparin (Figure 16). The fluorescent probe used was γ-AmNS-UTP (Aminonaphthalenesulfonate labeled UTP) (Molecular Probes). This probe shows fluorescence only when the UTP has been incorporated in the RNA molecule.

Reactions were done in 50 µL volume. The following reactions were performed: without any inhibitors, titration with ppGpp (Trilink®) (50 µM to 2000 µM), titration with DksA (50 nM to 4000 nM), titration with ppGpp when constant concentration of DksA was present and titration with DksA when constant concentration of ppGpp was present. The reaction was started with 49 µL volume containing transcription buffer (50 mM Tris (pH 7.9), 100 mM KCl, 10 mM $MgCl_2$, 1 mM DTT, 10 µg/ml BSA, 5% glycerol), inhibitors/H_2O and 100 nM RNAP holoenzyme, by incubating it for 5 mins at 37°C. After 5 mins, 1 µL of lac(CONS) (-49/30) (Figure 17) was added to the mixture and incubated for 15 mins at 37°C. 0.5 µL of 1 mg/ml heparin and 0.5 µL of 2.5 mM γ-AmNS-UTP were added and the contents were transferred to a submicro fluorometer cuvette (Starna Cells) for fluorescence measurement. 2 µL of 10 mM ApA (initiating nucleotides, primer) was added to the cuvette, mixed well and placed in the Spectrofluorometer (at 37°C). After 1 min, signal was measured with excitation wavelength 360 nm and emission wavelength 500 nm, for 300 secs. Using Felix analysis software®, the slope of the straight-line intensity vs. time curve was calculated. Using SigmaPlot®, graphs were plotted to calculate IC50 values (concentration of drug at half-maximal inhibition) and the values obtained for mutants were compared to that for wild-type.

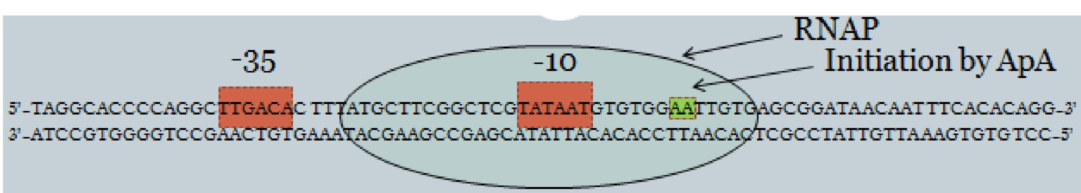


Figure 16. In vitro abortive initiation assay.

5'-TAG GCA CCC CAG GCT TGA CAC TTT ATG CTT CGG CTC GTA TAA TGT GTG GAA TTG TGA GCG GAT AAC AAT TTC ACA CAG G-3'

Figure 17. lac(CONS) (-49/30) (only one strand shown)

Chapter 4: Results

4.1 Mutants isolated and linkage

From the primary screen (Figure 18), a total of more than 120 mutants were picked up. From these mutants, after sequencing and secondary screens, a total of 31 independent single-substitution mutants in the *rpoC* gene were identified (Table 1).

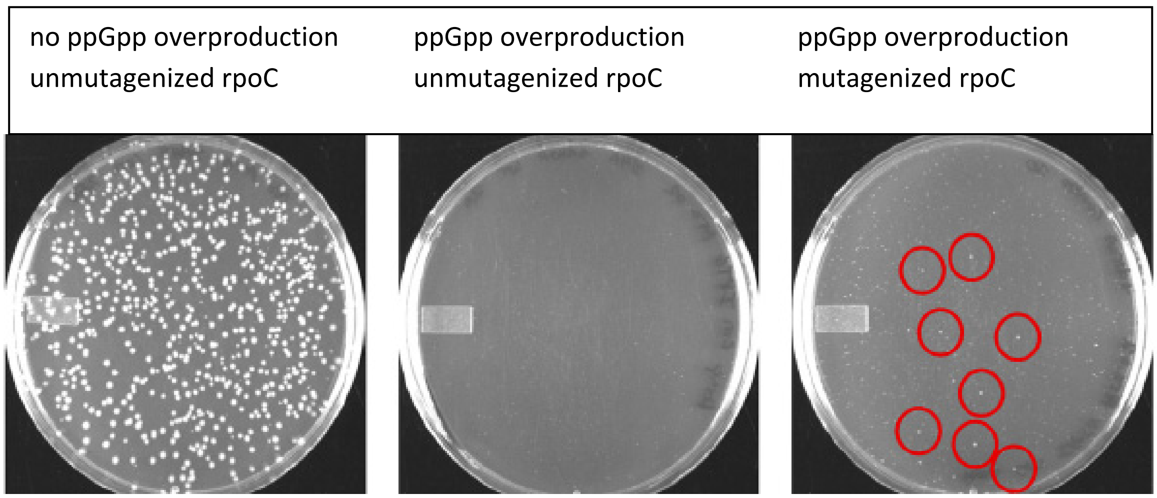


Figure 18. Isolation of mutants from primary screen; red circles show mutant colonies

Table 1. List of independent mutants isolated in *rpoC* gene

Residue #, *E. coli*	Amino acid substitution	# of mutants isolated
339	R>P	1
343	L>R	2
343	L>P	1
350	S>W	1
430	H>R	1

458	N>T	1
460	D>Y	1
460	D>E	1
736	Q>K	1
739	Q>H	1
789	K>I	1
921	Q>K	1
925	E>K	1
930	L>P	1
931	T>G	1
931	T>N	1
931	T>A	1
1133	D>E	2
1133	D>V	1
1134	I>L	1
1134	I>F	1
1136	G>S	2
1136	G>R	1
1136	G>C	2
1324	S>W	1
1382	P>T	1

All these mutants were isolated in DksA+ background. The primary screen in DksA- background did not give any mutants.

These mutants make two distinct clusters on the RNAP molecule (Figure 19). One cluster belongs to the switch region (R339P, L343R, L343P, S350W, S1324W, P1382T) responsible for closing and opening of clamp which disallows and allows, respectively, the entry and exit of DNA (Darst, 2001). The clamp consists of one side of the active center cleft, comprising the N-terminal end of β', the C-terminal of β and ω.

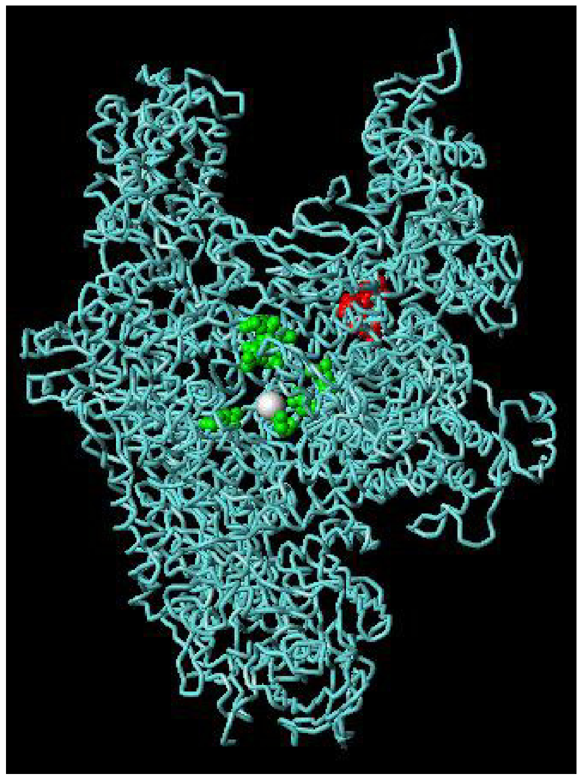


Figure 19. Representation of mutants on the RNAP molecule. The red cluster represents mutants in the switch region; the green cluster is mutants in the secondary channel.

The other cluster of mutants is located in the secondary channel (or close to it). Specifically, the mutants belong to trigger loop (E925K, L930P, T931G, T931N, T931A, D1133E, D1133V, D1134L, D1134F, G1136S, G1136R, and G1136C), which causes the opening and closing of the active center during nucleotide-addition cycle; bridge-helix

(K789I), which spans the active-center cleft and separates the cleft from the secondary channel; and secondary channel wall (Q736K). Some mutants (N458T, D460Y and D460E) also belong to the Mg^{2+} at the active center.

For all these mutants, it was shown that the mutation in the rpoC gene was linked to the mutant phenotype (Figure 20).

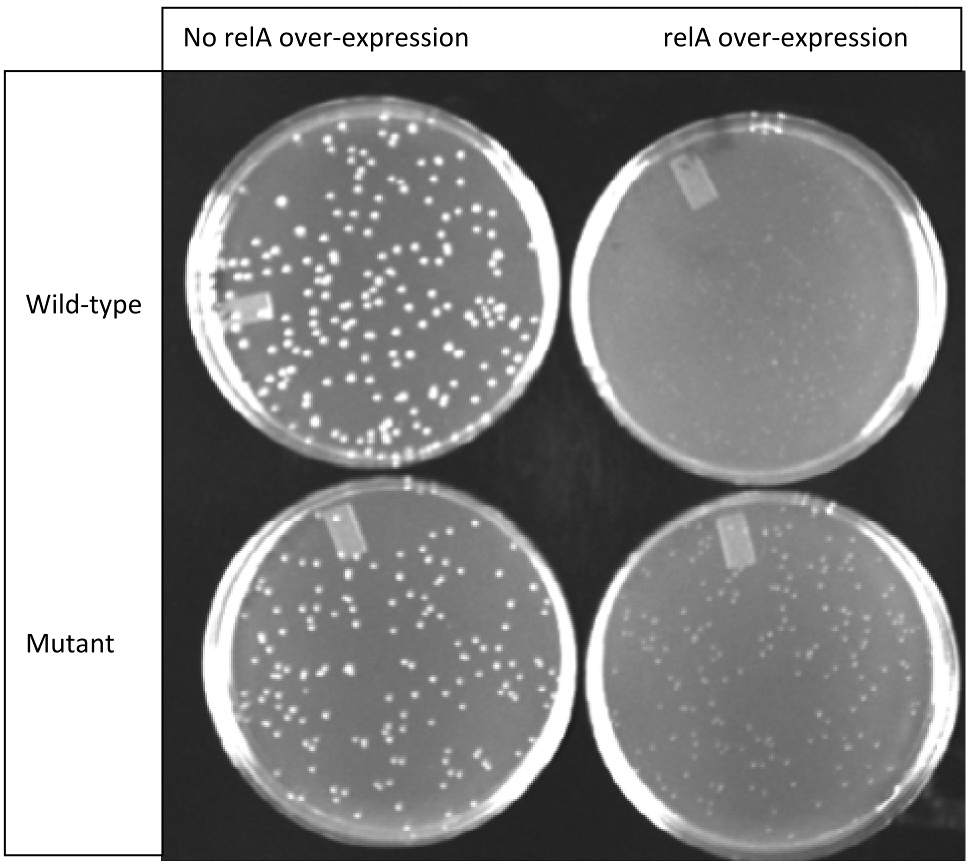


Figure 20. Results from linkage experiment. The mutant has a much larger colony size compared to wild-type when relA is over-expressed.

4.2 β-galactosidase assay results

For some mutants, the β-galactosidase assay (Figure 21) was performed.

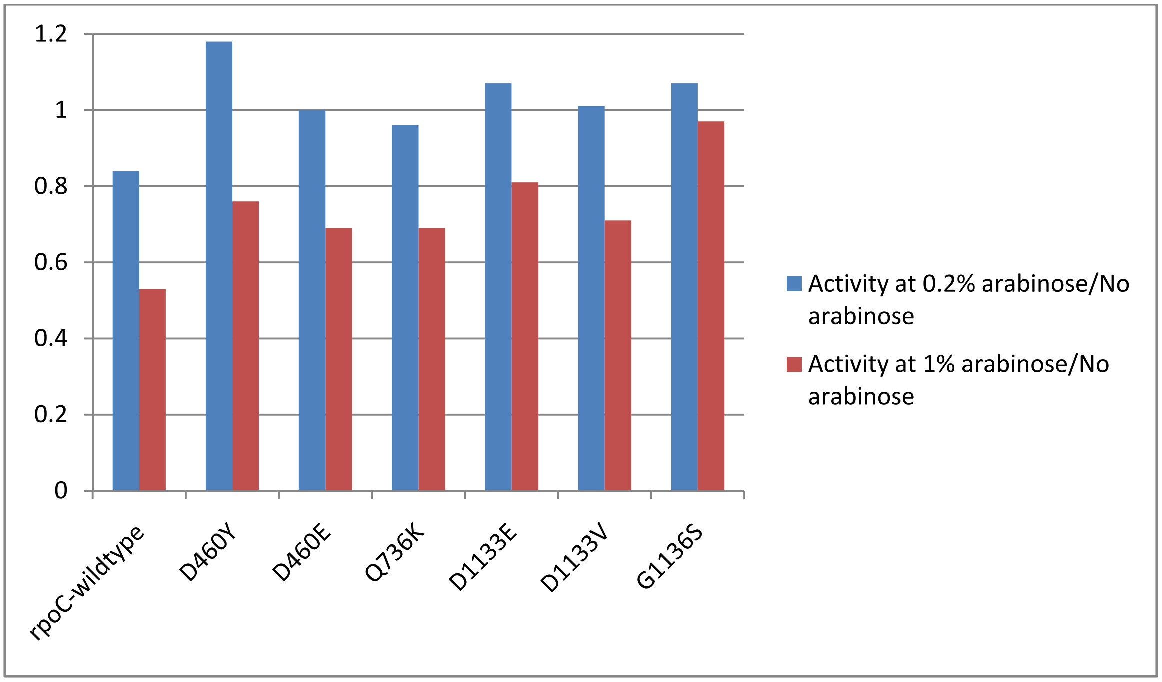


Figure 21. β-galactosidase assay results.

The wild-type shows a ratio of 0.53 in presence of 1% arabinose which implies that when ppGpp is produced the β-galactosidase activity is only ~50% of what it is when ppGpp is not produced. In other words, the rrnB P1 promoter is 50% inhibited when ppGpp is produced in presence of 1% arabinose. These results show that when compared to wild-type, mutants show higher β-galactosidase activity which in turn means that they are resistant to toxic effects of ppGpp.

The assay was also performed for some mutants in strain without DksA (RLG6348), but the results (not shown) were too random to make any logical interpretations and was therefore, not repeated.

4.3 *In-vitro* transcription results

Some mutant RNAP were purified (Figure 22) and assayed for *in-vitro* activity. DksA was purified for *in-vitro* assays (Figure 23).

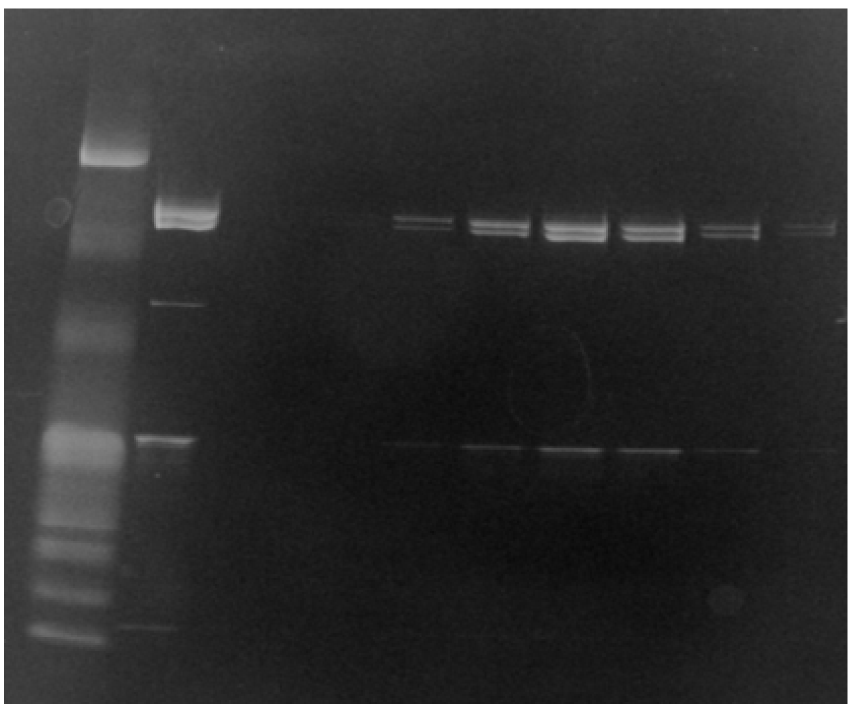


Figure 22. RNA polymerase purification: FPLC fractions. From left to right: protein marker, RNAP standard, FPLC fractions

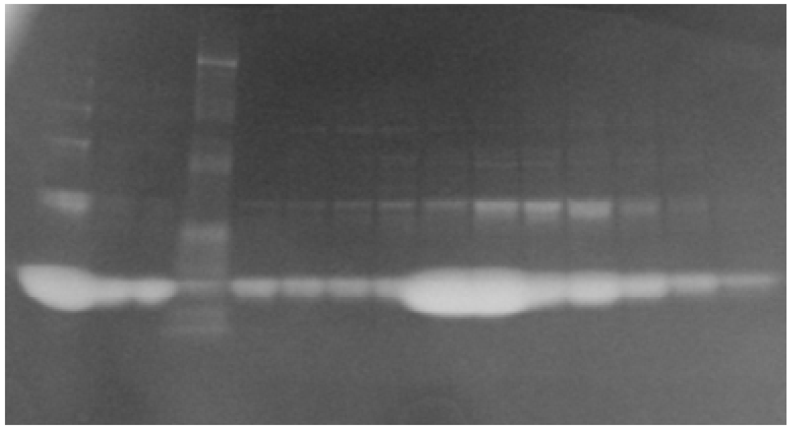


Figure 23. DksA purification.

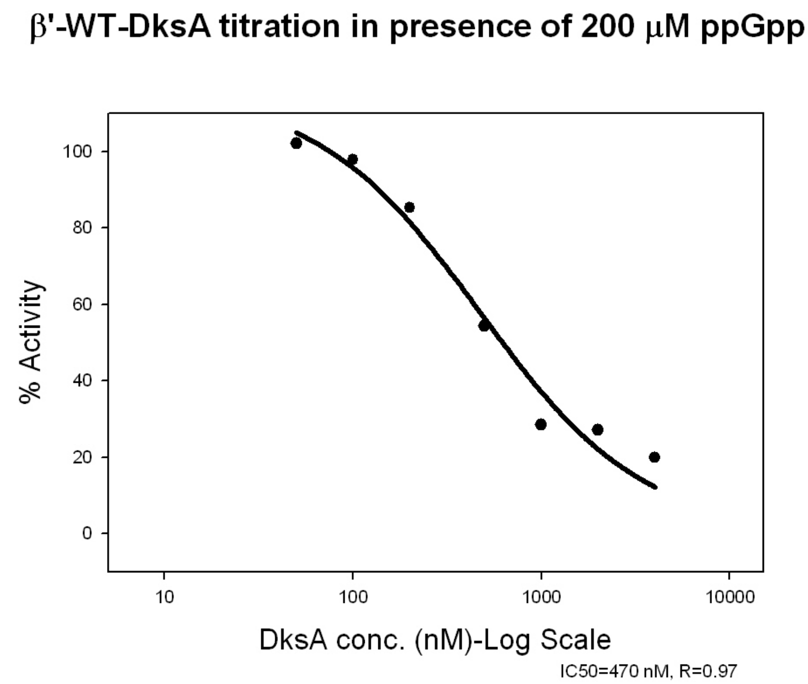
The results for the mutants tested are shown below (Table 2) as fold-change in IC50 values (the concentration of drug at half-maximal inhibition) compared to the wild-type values for the corresponding case.

Table 2. *In-vitro* transcription assay

IC50 values (Fold change as compared to wild-type)				
	ppGpp titration	**ppGpp titration in presence of 1000 nM DksA**	**DksA titration**	**DksA titration in presence of 200 µM ppGpp**
L343P	1	1	1	1
N458T	Polymerase has low activity; showed wild-type response at highest concentrations of ppGpp and DksA			
D460E	Inactive polymerase			
Q736K	1	1	1	1
E925K	1	1	1	1
D1133E	1	1	1	8
I1134L	1	11	35000	46
G1136C	1	1	11	6
G1136R	1	1	75	5

Mutants resistant to toxic effects of ppGpp should show a significantly larger IC50 than wild-type. An example of the graphs drawn to calculate above values is shown below in

Figure 24. A larger IC50 implies that more of the drug is required to inhibit the RNAP activity. Mutants I1134L and G1136R show a significant increase in IC50 values when titrated with DksA and that effect also shows for titration with DksA in presence of ppGpp for both of them. I1134L also shows a significantly higher IC50 when titrated with ppGpp in presence of DksA. D1133E shows a significant fold-change when titrated with DksA in presence of 200 μM ppGpp. G1136C shows a higher IC50 when titrated with DksA and that effect translates when titrated with DksA in presence of ppGpp. All other mutants do not have a significantly higher IC50 than wild-type suggesting that they are not resistant to toxic effects of ppGpp or DksA. (Note: A fold-change of more than 4 was considered significant and a fold change of 1 means wild-type response)



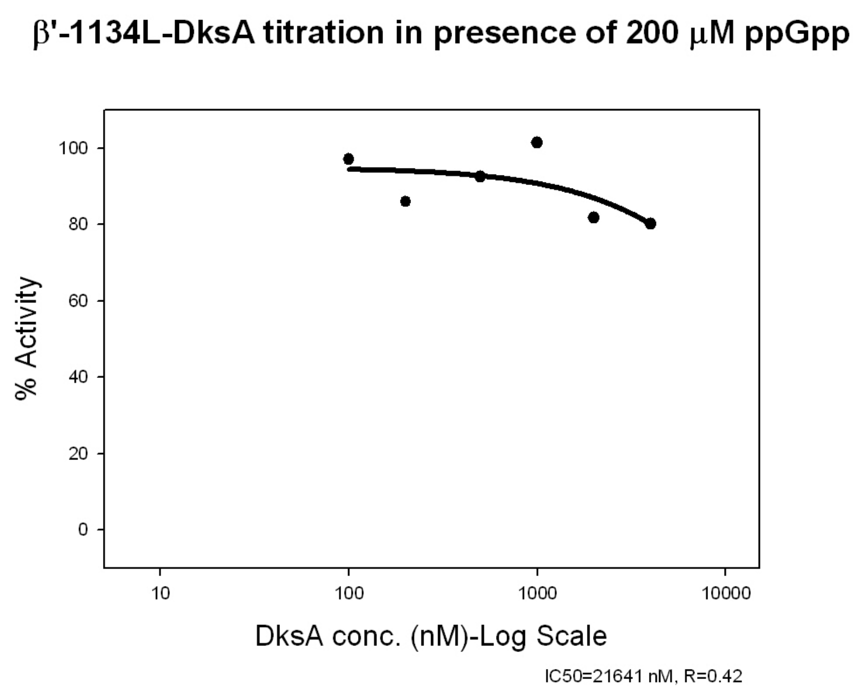


Figure 24. ***In-vitro*** **transcription assay. DksA titration with 200 μM ppGpp for wild-type and 1134L.**

Chapter 5: Discussion

The mutants represent two distinct clusters on the RNAP. The mutants that belong to the secondary channel may be conformation/binding determinants for ppGpp and/or DksA. The secondary channel is the site for NTP entry. The mutants that belong to the switch region are most likely conformational determinants of ppGpp/DksA as the structure and sequence in that region does not seem to be a binding site for either a tetraphosphate or a macromolecule. The switch region is responsible for the opening and closing of the clamp.

Previous studies have reported mutants in *rpoC* gene which mimic the effect of ppGpp in a *ΔrelA ΔspoT* strain of *E.coli*. Several of these mutants map in the clamp domain/switch region (H1445S, A490E). Other mutants (E1264D, R1266C) map near the secondary channel mutants derived from this study and bridge helix (R1087Q). [See (Trinh, Langelier, Archambault, & Coulombe, 2006) for a compilation of these mutants]. These mutants were expected to be most likely playing a role in destabilizing the open complex formation. The mutants got from this study are, thus, consistent with observations made before.

The finding from *in-vitro* transcription assay that I1134L and G1136R are resistant to toxic effects of DksA means these two could be binding determinants for DksA. The putative DksA site, the secondary channel, is very close to these mutants. Several mutants which did not show resistance to ppGpp/DksA may be because they are indirect mutants. Also, there are many proposed mechanisms by which ppGpp could act (see introduction), all of which, like promoter clearance, pausing during elongation, are not assayed with the current experimental setup. In addition, there might be other factors responsible for action of ppGpp which have not been identified yet. The mutants at the active site showed no/little activity and therefore need to be assessed in a different assay than the one currently being used.

β-galactosidase assay results do not form a distinct correlation with *in-vitro* assay results with the current data set. This could be due to the various drawbacks of the β-galactosidase assay namely: limited dynamic range, growth-rate dependence, long duration of stress conditions. The difference in the results may also be plainly due to the fact that β-galactosidase is an *in-vivo* assay, where any accessory molecules, if any, for ppGpp regulation would be present vis-à-vis *in-vitro* assay.

It is also possible that enrichment of the positives from the primary screen (1 in 4-5 mutants assayed is positive on *in-vitro* assays) is such that we require larger set of mutants to get a good representation of the all the determinants of ppGpp/DksA.

It is preposterous to conclude that the mutants isolated from the primary screen, which did not present themselves as mutants on the various assays, are not mutants. The lack of an appropriate assay to account for the various factors affecting the ppGpp/regulation must be the primary issue to be dealt with. It is imperative that the mutants isolated from the primary screen are assayed correctly to select for positives before performing the *in-vitro* assays because these assays are not only time consuming but also, as it has been shown in various studies, may not be a good representation of events taking place inside the cell. Results from similar analysis on other subunits of the enzyme will also play a crucial role in explaining the overall picture. At the end, a holistic picture will only emerge from combination of results from variety of assays, previous knowledge and practical logic.

Bibliography

Artsimovitch, I., Patlan, V., Sekine, S.-i., Vassylyeva, M. N., Hosaka, T., Ochi, K., et al. (2004). Structural Basis for Transcription Regulation by Alarmone ppGpp. *Cell , 117*, 299-310.

Barker, M. M., & Gourse, R. L. (2001). Regulation of rRNA Transcription Correlates with Nucleoside Triphosphate Sensing. *Journal of Bacteriology , 183* (21), 6315-6323.

Barker, M. M., Gaal, T., & Gourse, R. L. (2001). Mechanism of Regulation of Transcription Initiation by ppGpp. II. Models for Positive Control Based on Properties of RNAP Mutants and Competition for RNAP . *Journal of Molecular Biology , 305*, 689-702.

Barker, M. M., Gaal, T., Josaitis, C. A., & Gourse, R. L. (2001). Mechanism of Regulation of Transcription Initiation by ppGpp. I. Effects of Transcription Intiation in vivo and in vitro. *Journal of Molecular Biology , 305*, 673-688.

Brown, L., Gentry, D., Elliott, T., & Cashel, M. (2002). DksA Affects ppGpp Induction of RpoS at a Translational Level. *Journal of Bacteriology , 184* (16), 4455-4465.

Cashel, M., & Gallant, J. (1969). Two Compounds implicated in the Function of the RC Gene of Escherichia coli. *Nature , 221*, 838-841.

Chatterji, D., & Ojha, A. K. (2001). Revisiting the stringent response, ppGpp and starvation signaling. *Current Opinion in Microbiology , 4*, 160-165.

Chatterji, D., Fujita, N., & Ishihama, A. (1998). The mediator for stringent control, ppGpp, binds to the B-subunit of Escherichia coli RNA polymerase . *Genes to Cells , 3*, 279-287.

Christie, G. E., Cale, S. B., Isaksson, L. A., Jin, D. J., Xu, M., Sauer, B., et al. (1996). Escherichia coli rpoC397 Encodes a Temperature-Sensitive C-Terminal Frameshift in the b9 Subunit of RNA Polymerase That Blocks Growth of Bacteriophage P2. *Journal of Bacteriology , 178* (23), 6991-6993.

Chung, C. T., Niemela, S. L., & Miller, R. H. (1989). One-step preparation of competent Escherichia coli: Transformation and storage of bacterial cells in the same solution. *Proceedings of the National Academy of Sciences , 86*, 2172-2175.

Darst, S. A. (2001). Bacterial RNA polymerase. *Current Opinion in Structural Biology , 11*, 155-162.

Ebright, R. H. (2000). RNA polymerase: structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II. *Journal of Molecular Biology , 304* (5), 687-98.

Ebright, R. H., & Busby, S. (1995). The Escherichia coil RNA polymerase 'alpha' subunit: structure and function. *Current Opinion in Genetics and Development* (5), 197-203.

Erie, D. A., Yager, T. D., & von Hippel, P. H. (1992). The Single-Nucleotide Addition Cycle in Transcription: a Biophysical and Biochemical Perspective. *Annual Review of Biophysics and Biomolecular Structure , 21*, 379-415.

Ghosh, P., Ishihama, A., & Chatterji, D. (2001). Escherichia coli RNA polymerase subunit 'omega' and its N-terminal domain bind full-length 'Beta-prime' to facilitate incorporation into the 'alfa2beta' subassembly. *European Journal of Biochemistry , 268* (17), 4621-4627.

Godfrey, H. P., Bugrysheva, J. V., & Cabello, F. C. (2002). The role of the stringent response in the pathogenesis of bacterial infections. *Trends in Microbiology , 10* (8), 349-351.

Guzman, L.-M., Belin, D., Carson, M. J., & Beckwith, J. (1995). Tight Regulation, Modulation, and High-Level Expression by Vectors Containing the Arabinose PBAD Promoter. *Journal of Bacteriology , 177* (14), 4121-4130.

Harley, C. B., & Reynolds, R. P. (1987). Analysis of E. coli promoter sequences. *Nucleic Acids Research , 15* (5), 2343-2361.

Ishihama, A. (1981). Subunit assembly of E. coli RNA polymerase. *Advances in Biophysics , 14*, 1-35.

Jain, V., Kumar, M., & Chatterji, D. (2006). ppGpp: Stringent Response and Survival. *The Journal of Microbiology , 44* (1), 1-10.

Jores, L., & Wagner, R. (2003). Essential Steps in the ppGpp-dependent Regulation of Bacterial Ribosomal RNA Promoters Can Be Explained by Substrate Competition. *The Journal of Biological Chemistry , 278* (19), 16834-16843.

Kasai, K., Nishizawa, T., Takahashi, K., Hosaka, T., Aoki, H., & Ochi, K. (2006). Physiological Analysis of the Stringent Response Elicited in an Extreme Thermophilic Bacterium, Thermus thermophilus. *Journal of Baceteriology , 188* (20), 7111-7122.

Khlebnikov, A., Datsenko, K. A., Skaug, T., Wanner, B. L., & Keasling, J. D. (2001). Homogeneous expression of the PBAD promoter in Escherichia coli by constitutive expression of the low-affinity high-capacity AraE transporter. *Microbiology , 147*, 3241-3247.

Kung, P. J., & Craig, E. A. (1990). Identification and Characterization of a New Escherichia coli Gene That Is a Dosage-Dependent Suppressor of a dnaK Deletion Mutation. *Journal of Bacteriology , 172* (4), 2055-2064.

Magnusson, L. U., Farewell, A., & Nystrom, T. (2005). ppGpp: a global regulator in Escherichia coli . *Trends in Microbiology , 13* (5), 236-242.

Miller, J. H. (1992). *A Short Course in Bacterial Genetics: A Laboratory Manual for Escherichia coli and Related Bacteria.* Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Mukhopadhyay, J., Sineva, E., Knight, J., Levy, R. M., & Ebright, R. H. (2004). Antibacterial Peptide Microcin J25 Inhibits Transcription by Binding within and Obstructing the RNA Polymerase Secondary Channel. *Molecular Cell , 14*, 739-751.

Murakami, K. S., & Darst, S. A. (2003). Bacterial RNA polymerases: the wholo story. *Current Opinion in Structural Biology , 13*, 31-39.

Niu, W., Kim, Y., Tau, G., Heyduk, T., & Ebright, R. H. (1996). Transcription Activation at Class II CAP-Dependent Promoters: Two Interactions between CAP and RNA Polymerase. *Cell , 87*, 1123-1134.

Parrish, N. M., Dick, J. D., & Bishai, W. R. (1998). Mechanisms of latency in Mycobacterium tuberculosis. *Trends in Microbiology , 6*, 107-112.

Paul, B. J., Barker, M. M., Ross, W., Schneider, D. A., Webb, C., Foster, J. W., et al. (2004). DksA: A Critical Component of the Transcription Initiation Machinery that Potentiates the Regulation of rRNA Promoters by ppGpp and the Initiating NTP. *Cell , 118*, 311-322.

Paul, B. J., Berkmen, M. B., & Gourse, R. L. (2005). DksA potentiates direct activation of amino acid promoters by ppGpp. *Proceedings of the National Academy of Sciences , 102* (22), 7823-7828.

Paul, B. J., Ross, W., Gaal, T., & Gourse, R. L. (2004). rRNA transcription in Escherichia Coli. *Annual Review of Genetics , 38*, 749-70.

Perederina, A., Svetlov, V., Vassylyeva, M. N., Tahirov, T. H., Yokoyama, S., Artsimovitch, I., et al. (2004). Regulation through the Secondary Channel—Structural Framework for ppGpp-DksA Synergism during Transcription. *Cell , 118*, 297-309.

Potrykus, K., & Cashel, M. (2008). (p)ppGpp: Still Magical? *Annual Review of Microbiology , 62*, 35-51.

Potrykus, K., Wegrzyn, G., & Hernandez, V. J. (2002). Multiple Mechanisms of Transcription Inhibition by ppGpp at the lamdapR Promoter. *The Journal of Biological Chemistry , 277* (46), 43785-43791.

Schreiber, G., Metzger, S., Aizenman, E., Roza, S., Cashel, M., & Glaser, G. (1991). Overexpression of the relA Gene in Escherichia coli. *The Journal of Biological Chemistry , 266* (6), 3760-3767.

Srivatsan, A., & Wang, J. D. (2008). Control of bacterial transcription, translation and replication by (p)ppGpp. *Current Opinion in Microbiology , 11*, 1-6.

Tedin, K., & Bremer, H. (1992). Toxic Effects of High Levels of ppGpp in Escherichia coli Are Relieved by rpoB Mutations. *The Journal of Biological Chemistry , 267* (4), 2337-2344.

Toulokhonov, I. I., Shulgina, I., & Hernandez, V. J. (2001). Binding of the Transcription Effector ppGpp to Escherichia coli RNA Polymerase Is Allosteric, Modular, and Occurs Near the N Terminus of the b*-Subunit. *The Journal of Biological Chemsitry , 276* (2), 1220-1225.

Traxler, M. F., Summers, S. M., Nguyen, H.-T., Zacharia, V. M., Hightower, G. A., Smoth, J. T., et al. (2008). The global, ppGpp-mediated stringent response to amino acid starvation in Escherichia coli. *Molecular Microbiology , 68* (5), 1128-1148.

Trinh, V., Langelier, M.-F., Archambault, J., & Coulombe, B. (2006). Structural Perspective on Mutations Affecting the Function of Multisubunit RNA Polymerases. *MICROBIOLOGY AND MOLECULAR BIOLOGY REVIEWS , 70* (1), 12-36.

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

Vrentas, C. E., Gaal, T., Berkmen, M. B., Rutherford, S. T., Haugen, S. P., Ross, W., et al. (2008). Still Looking for the Magic Spot: The Crystallographically Defined Binding Site for ppGpp on RNA Polymerase Is Unlikely to Be Responsible for rRNA Transcription Regulation. *Journal of Molecular Biology , 377*, 551-564.

Vrentas, C. E., Gaal, T., Ross, W., Ebright, R. H., & Gourse, R. L. (2005). Response of RNA polymerase to ppGpp: requirement of the 'omega' subunit and relief of this requirement by DksA. *Genes & Development , 19*, 2378-2387.

Wang, D., Bushnell, D. A., Westover, K. D., Kaplan, C. D., & Kornberg, R. D. (2006). Structural basis of transcription: role of the trigger loop in substrate specifity and catalysis. *Cell , 127* (5), 941-954.

Wang, D., Meier, T. I., Chan, C. L., Feng, G., Lee, D. N., & Landick, R. (1995). Discontinuous Movements of DNA and RNA in RNA Polymerase Accompany Formation of a Paused Transcription Complex. *Cell , 81*, 341-350.

Young, B. A., Gruber, T. M., & Gross, C. A. (2002). Views of Transcription Initiation. *Cell , 109*, 417-420.

Zhang, G., Campbell, E. A., Minakhin, L., Richter, C., Severinov, K., & Darst, S. A. (1999). Crystal Structure of Thermus aquaticus Core RNA Polymerase at 3.3 A° resolution. *Cell , 98*, 811-824.

Zhou, Y., Zhang, X., & Ebright, R. H. (1991). Random mutagenesis of gene-sized DNA molecules by use of PCR with Taq DNA polymerase. *Nucleic Acids Research , 19* (21), 6052.

Curriculum Vita

Nimish Gupta

July 2002-July 2006	Indian Institute of Technology Madras, India
	Bachelor of Technology in Biotechnology
August 2006-Dec. 2008	Rutgers-The state university of New Jersey, New Brunswick
	Master of Science in Chemical & Biochemical Engineering
August 2007-Dec. 2007	Teaching Assistant. General Biology Laboratory
Jan 2008-May 2008	Teaching Assistant, Genetics
August 2008-Dec. 2008	Teaching Assistant, Genetics