

Structural and Functional Analysis of the NusB-S10 Complex Shared Between Transcription and Translation

Dissertation

for the award of the degree

“Doctor of Philosophy” (Ph.D.)

Division of Mathematics and Natural Sciences
of the Georg-August-Universität Göttingen

submitted by

Xiao Luo

from HeiLongJiang, China

Göttingen, 2010

Prof. Dr. Markus Wahl (Reviewer)

AG Strukturbiochemie, Freie Universität Berlin

Prof. Dr. Ralf Ficner (Reviewer)

AG Molekulare Strukturbiologie, Institut für Mikrobiologie und Genetik, GZMB

Dr. Dirk Fasshauer

AG Strukturelle Biochemie, Max-Planck-Institut für biophysikalische Chemie

Date of the oral examination: 21.05.2010

Affidavit

I hereby declare that my doctoral thesis “Structural and functional analysis of the NusB-S10 complex shared between transcription and translation” has been written independently and with no other sources and aids than quoted. I have made clear what was done by project collaborators.

Parts of this thesis were published in:

Burmann B. M., Schweimer K., Luo X., Wahl M. C., Stitt B. L., Gottesman M. E. and Rösch P. (2010). A NusE:NusG complex links transcription and translation. *Science*, **328**: 501-504.

Burmann B. M., Luo X., Rösch P., Wahl M. C. and Gottesman M. E. (2010). Fine tuning of the *E. coli* NusB:NusE complex affinity to BoxA RNA is required for processive antitermination. *Nucleic Acids Research*, **38**: 314-326.

Luo X., Hsiao H. H., Bubunenko M., Weber G., Court D. L., Gottesman M. E., Urlaub H. and Wahl M. C. (2008). Structural and functional analysis of the *E. coli* NusB-S10 transcription antitermination complex. *Molecular Cell*, **32**: 791-802.

Xiao Luo

Göttingen, 2010

Table of Contents

1 Summary	1
2 Introduction	3
2.1 Overview of the prokaryotic transcription machinery	3
2.1.1 Prokaryotic transcription initiation	4
2.1.2 Prokaryotic transcription elongation	4
2.1.3 Prokaryotic transcription termination	4
2.2 Overview of the prokaryotic translation machinery	5
2.2.1 Prokaryotic translation initiation	6
2.2.2 Prokaryotic translation elongation	7
2.2.3 Prokaryotic translation termination and ribosome recycling	8
2.3 The coupling between prokaryotic transcription and translation machineries	8
2.3.1 Transcription antitermination systems	8
2.3.2 Proteins shared by the transcription and translation machineries	11
2.3.2.1 S10 (NusE)	11
2.3.2.2 L4	13
2.3.2.3 S1	13
2.3.2.4 NusG	14
2.3.2.5 NusB	14
2.4 Aims of this study	15
3 Materials and methods	17
3.1 Materials	17
3.1.1 Chemicals	17
3.1.2 Media	18
3.1.3 Antibiotics	18
3.1.4 Nucleotides	18
3.1.5 Radionucleotides	18
3.1.6 Antibodies	18
3.1.7 Enzymes and inhibitors	19
3.1.8 DNA oligonucleotides	19

3.1.9 RNA oligonucleotides	19
3.1.10 Vectors	20
3.1.11 Plasmids	20
3.1.12 Bacterial strains	20
3.1.13 Commercial kits	21
3.1.14 Crystallization screens	21
3.1.15 Equipments	21
3.1.16 Consumption materials	22
3.2 Methods	23
3.2.1 Molecular cloning	23
3.2.1.1 PCR amplification	23
3.2.1.2 Agarose gel electrophoresis and DNA fragment isolation	24
3.2.1.3 Enzyme digestion and ligation	24
3.2.1.4 Preparation of competent cells for electroporation transformation	24
3.2.1.5 Preparation of competent cells for chemical transformation	25
3.2.1.6 Electroporation transformation	25
3.2.1.7 Chemical transformation	25
3.2.1.8 Mini-preparation of plasmid and DNA sequencing	25
3.2.1.9 PCR-based site directed mutagenesis	26
3.2.2 Protein Production	26
3.2.2.1 Co-expression of protein complexes	26
3.2.2.2 Cell lysis	27
3.2.2.3 Co-purification of protein complexes	27
3.2.2.4 Determination of protein concentrations	27
3.2.2.5 SDS-polyacrylamide gel electrophoresis	28
3.2.2.6 Gel staining	28
3.2.3 Protein crystallography	28
3.2.3.1 Pre-crystallization test	29
3.2.3.2 Protein crystallization	29
3.2.3.3 Data collection and processing	30
3.2.3.4 Phasing, model building and refinement	30
3.2.3.5 Structure analysis	30
3.2.4 Biochemical assays	31
3.2.4.1 GST pull-down assay	31

3.2.4.2 5'-End labeling of RNA-oligonucleotides	31
3.2.4.3 Double filter-binding assay	31
3.2.4.4 UV-induced crosslinking assay	32
3.2.4.5 Deduction of protein-RNA crosslinking sites	32
3.2.4.6 Ribosome preparation	33
3.2.4.7 Western blot	34
3.2.4.8 Analytical size exclusion chromatography	35
4 Results	37
4.1 Transcriptional and translational functions are attributed to distinct regions of S10	37
4.1.1 The long ribosome-binding loop of S10 is dispensable for transcriptional functions	37
4.1.2 The loop-deleted S10 variant does not bind to ribosomes	37
4.2 Structural analysis of the NusB-S10 complex	41
4.2.1 Crystal structure of a transcriptionally active NusB-S10 Complex	41
4.2.2 NusB and S10 retain their overall folds upon complex formation but interact <i>via</i> local induced fit	43
4.2.3 Binding of S10 to NusB is mutually exclusive with its incorporation into the ribosome and with NusB dimerization	45
4.2.4 Molecular basis of the conserved proline motif on S10	47
4.2.5 NusB does not influence the <i>cis/trans</i> equilibrium at Pro39 of S10	49
4.2.6 Molecular basis of the <i>nusB5</i> and <i>nusE100</i> phenotypes	51
4.2.7 The <i>nusE71</i> mutation defines an additional interaction surface on S10	52
4.3 BoxA RNA binding by the NusB-S10 complex	54
4.3.1 The NusB-S10 complex exhibits an intermolecular, mosaic and contiguous BoxA RNA-binding surface	54
4.3.2 <i>nusB101</i> represents a gain-of-function mutation with increased RNA affinity	56
4.3.3 The structure of NusB ^{Asp118Asn} -S10 ^{Δloop} closely resembles the structure of NusB-S10 ^{Δloop}	59
4.4 Roles of S10 and NusB in transcription and translation	62
4.4.1 S10 supports transcription antitermination in the absence of NusB	62
4.4.2 NusB delivers S10 into other molecular environment	63
4.4.3 Does NusB escort S10 into ribosomes?	65
4.4.4 NusG couples transcription and translation <i>via</i> S10	66

5 Discussion	69
5.1 S10 ^{Δloop} is a tool to dissect translational and transcriptional functions of S10	69
5.2 S10 is adapted to different functional contexts without global structural remodeling	70
5.3 Mutually exclusive binding of S10 to the 30S subunit or NusB may provide for feedback control of ribosome biogenesis	70
5.4 S10 and NusB form a functional module for recognition of BoxA	71
5.5 S10 is the active antitermination factor of the NusB-S10 complex and NusB serves as an adaptor in the transcription process	71
5.6 Hypothesis: NusB may deliver S10 into ribosomes	72
6 References	75
7 Appendixes	85
7.1 Principles of protein X-ray crystallography	85
7.1.1 Crystal growth	85
7.1.2 Data collection and processing	85
7.1.3 Solutions of the phase problem	87
7.1.4 Fitting, refinement and validation of crystal structure	88
7.2 Abbreviations	90
7.3 Acknowledgements	94
7.4 Curriculum Vitae	95

List of Figures

Figure 2.1 The prokaryotic transcription cycle	3
Figure 2.2 The prokaryotic translation cycle	6
Figure 2.3 Map of the regulatory region of phage λ lytic phase	9
Figure 2.4 Transcription anitermination or termination models	10
Figure 2.5 Structures of S10 and NusB	12
Figure 4.1 Analysis of the S10 ^{Δloop} mutant	38
Figure 4.2 Gel analysis of <i>nusE</i> <> <i>kan</i> recombinants and ribosome binding of S10 ^{Δloop} ...	40
Figure 4.3 Structure of the NusB-S10 ^{Δloop} complex	43
Figure 4.4 Aspects of the NusB-S10 ^{Δloop} interaction	46
Figure 4.5 Aspects of proline motif of S10	48
Figure 4.6 Structure of the NusB-S10 ^{Δloop, Pro39Ala} complex	50
Figure 4.7 Structure of the NusB-S10 ^{Δloop, Ala86Asp} complex	53
Figure 4.8 Mapping of crosslinked peptides on the surface of the NusB-S10 ^{Δloop} complex .	56
Figure 4.9 Protein-RNA and protein-protein crosslinking analysis	58
Figure 4.10 Structure of the NusB ^{Asp118Asn} -S10 ^{Δloop} complex	61
Figure 4.11 Aspects of the S10 ^{Δloop} in isolation	65
Figure 4.12 Ribosome binding of S10	66
Figure 4.13 Size exclusion chromatography analyses	67
Figure 5.1 A feedback control circuit by S10	71

List of Tables

Table 3.1 DNA oligonucleotides	19
Table 3.2 RNA oligonucleotides	20
Table 3.3 Vectors	20
Table 3.4 Plasmids	20
Table 3.5 High-throughput crystallization experiments	29
Table 3.6 Data collection strategy	30
Table 4.1 <i>nusE</i> ⁺ and <i>nusE</i> ^{Δloop} are dominant to <i>nusE71</i>	38
Table 4.2 Crystallographic data	44
Table 4.3 Structural comparisons	45
Table 4.4 Transcription activities tests by overproduction of S10 or S10 proline mutants ..	48
Table 4.5 Overexpression of NusB5 overcomes the <i>nusB5</i> defect	52
Table 4.6 Electrospray-ionization tandem mass spectrometry identification of protein- RNA crosslinks	54
Table 4.7 Overproduction of S10 or S10 ^{Δloop} allows λ to grow on a Δ <i>nusB</i> strain (<i>nusB::Cam</i>)	62

1 Summary

Transcription and translation are two highly coupled processes during prokaryotic gene expression where ribosomes initiate translation on mRNAs already during transcription, in contrast to eukaryotes where two principle processes occur in two different cellular compartments.

One of mechanisms by which transcription and translation in prokaryotes communicate directly with one other is sharing proteins, which have dual activity. As one such example, S10 protein was initially defined as a ribosomal (r-) protein before an additional role in transcription was discovered. S10 is a component of the 30S ribosomal subunit and participates together with NusB protein in processive transcription antitermination. NusB is implicated in translation through studies of its mutations that slow down the translation elongation rate. However, the exact role of NusB in translation remains unknown and the molecular mechanisms by which S10 and NusB can act as transcription or translation factors are still a mystery.

Here, regions of S10 dispensable for transcription antitermination were delineated through complementation assays and recombineering. The crystal structure of a transcriptionally active NusB-S10 complex was determined. In the complex, S10 adopts the same fold as in the 30S subunit and is blocked from simultaneous association with the ribosome. Mass spectrometric mapping of UV-induced crosslinks revealed that the NusB-S10 complex presents an intermolecular, composite, and contiguous binding surface for RNAs containing BoxA antitermination signals. Furthermore, S10 overproduction complemented a *nusB* null phenotype. These data demonstrate that S10 and NusB together form a BoxA binding module, that NusB facilitates entry of S10 into the transcription machinery, and that S10 represents a central hub in processive antitermination. Last, the evidence that NusB plays a role of a loading factor in delivering S10 into transcription antitermination complex and into other molecular environment *in vitro* (crystals) allowed me to deduce a hypothesis that in translation NusB may still function as a loading factor that delivers S10 into ribosomes.

2 Introduction

2.1 Overview of the prokaryotic transcription machinery

Transcription (RNA synthesis) is the first step of gene expression where RNA polymerase (RNAP) reads DNA and produces a complementary, antiparallel copy of DNA sequence as an RNA product, which is then translated by ribosomes to yield proteins (Squires and Zaporjets, 2000). The prokaryotic RNAP molecule is a complex enzyme composed of two α subunits, one β subunit, one β' subunit, one ω subunit and one σ subunit. The σ subunit can be separated from other subunits to give rise to a core RNAP. Three main types of RNA are obtained from transcription: messenger RNA (mRNA) that will be translated into amino acids for protein biosynthesis; transfer RNA (tRNA) that transfers amino acids to ribosomes; ribosomal RNA (rRNA) that involves in the ribosomes assembly and catalysis. The prokaryotic transcription is divided into several major phases that are illustrated in the Figure 2.1.

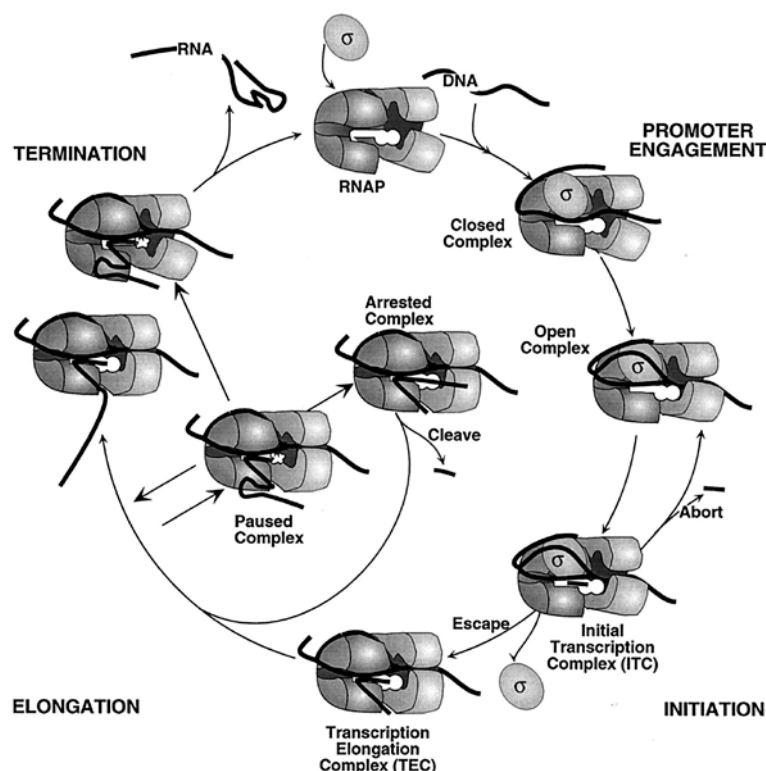


Figure 2.1 The prokaryotic transcription cycle (Mooney et al., 1998)

For simplicity, not all intermediate steps are shown. The four major phases of the transcription (promoter engagement, initiation, elongation and termination) are discussed in the main text.

2.1.1 Prokaryotic transcription initiation

The initial stage of the prokaryotic transcription begins with the binding of RNAP to the promoter in DNA, which localizes approximately 10 and 35 base pairs upstream from the start site of transcription. The binding of RNAP to the promoter is aided by σ subunit. The transcription is then initiated with melting of about 15 base pairs of DNA around the initiation site and scrunching of DNA of the growing bubble into RNAP (Roberts et al., 2008). The σ subunit is released from RNAP after addition of about first 10 nucleotides (nt) by accumulated stress from DNA scrunching, which also drives breakage of the interaction between RNAP and promoter DNA, as well as between RNAP and other initiation factors for promoter clearance (Kapanidis et al., 2006). The initiation process is also affected by many other initiation factors, including both positively acting factors like AraC, CAP and Fis, and negatively acting factors like repressors (Squires and Zaporozhets, 2000).

2.1.2 Prokaryotic transcription elongation

After the initiation, RNAP moves along the DNA template strand (non-coding strand) to make the elongation of the growing RNA chain. Transcription elongation starts with binding of a template-complementary nucleotide triphosphate (NTP) into the growing bubble, followed by the reaction between the RNA chain 3'-OH and the NTP α -PO₄ group (Roberts et al., 2008; Vassylyev et al., 2007). This chemical reaction, catalyzed by a pair of bound Mg²⁺ ions, results in the addition of one nucleotide monophosphate (NMP) to the RNA and release of a pyrophosphate and subsequently, the next template base is placed in the growing bubble (Roberts et al., 2008; Vassylyev et al., 2007). This process produces an RNA molecule that is an exact copy of the coding strand of DNA with the exceptions that thymines (T) are replaced with uracils (U) and that the nucleotides are made of ribose sugars. During the elongation process, at certain template sites RNAP pauses frequently to modulate the elongation rate, reflecting a finely detailed evolution of transcription rate to match the particular fate of the transcript (Roberts et al., 2008). Pausing and antipausing are fine-tuned by specific factors which lead RNAP to vary its elongation rate from 40-45 nucleotides/s on most mRNAs to 80-90 nucleotides/s on rRNAs (Squires and Zaporozhets, 2000).

2.1.3 Prokaryotic transcription termination

Transcription proceeds until RNAP encounters a termination signal, where RNA synthesis stops, the growing RNA chain is released from RNAP, and RNAP dissociates from

the DNA template. Prokaryotes use two different ways to terminate transcription: intrinsic termination (Rho-independent termination) and Rho-dependent termination.

An intrinsic termination has a termination signal in DNA consisting of a GC-rich dyad symmetry element and an oligo T sequence (T stretch), thus the transcribed RNA contains a stable hairpin followed by an element of seven to nine U residues (U stretch) at the 3' terminus (d'Aubenton Carafa et al., 1990; Gusarov and Nudler, 1999). The transcription elongation complex stops at the end of U stretch (usually at U7 and U8 positions), and then it is converted to an irreversibly trapped configuration (Gusarov and Nudler, 1999). The hairpin disrupts most of the A-U base pairs in the DNA-RNA hybrid and also disrupts the interaction between single stranded RNA and RNAP, thus destabilizing the trapped elongation under physiological salt conditions (Gusarov and Nudler, 1999).

Rho-dependent termination depends on the Rho factor, a protein having both ATPase and helicase activities (Ciampi, 2006). Rho forms a hexameric ring by six protomers joined together through the N- and C-terminal domains of each protomer, where the Rho hexamer is split open to accommodate single-stranded RNA (Skordalakes and Berger, 2003). Each of six N-terminal domains of the Rho is working as the primary RNA binding site to mediate the tethering of the Rho to the Rho utilization site (*rut*), an element of 70-80 nt exhibiting a high-C and low-G content (Ciampi, 2006). The C-terminal domain in each protomer of the Rho contains several key motifs: P loop, a part of a Walker-type ATP binding protein required for ATP binding and hydrolysis; R loop and Q loop, together forming Rho's secondary RNA binding site (Skordalakes and Berger, 2003). Rho-dependent termination starts with the loading of the Rho to the *rut* site through the primary RNA binding sites in the N-terminal domains of the Rho (Ciampi, 2006; Skordalakes and Berger, 2003). To allow association between mRNA and the secondary RNA binding sites (R loop and Q loop) in the interior of the hexamer, the Rho ring opens during *rut* site binding (Ciampi, 2006; Skordalakes and Berger, 2003). Rho-mRNA binding activates the Rho's ATPase functionality, providing the energy for Rho translocation along the mRNA, and finally the transcript is released by Rho's helicase activity (Ciampi, 2006; Skordalakes and Berger, 2003).

2.2 Overview of the prokaryotic translation machinery

Translation is the first stage of protein biosynthesis by ribosomes through decoding mRNA generated in transcription to produce a specific polypeptide according to genetic codes. Ribosomes (70S) in prokaryotes consist of a small 30S ribosomal subunit and a large

50S ribosomal subunit. Assembly of the 30S ribosomal subunit needs proteins S1 through S21 along with the 16S rRNA, and the assembly of the 50S ribosomal subunit requires proteins L1 through L36 along with the 23S and 5S rRNA (Squires and Zaporozhets, 2000). The fully assembled prokaryotic ribosomes have three tRNA binding sites, defined as the aminoacyl (A), peptidyl (P) and exit (E) sites. The prokaryotic translation initiation, elongation, termination and ribosome recycling are illustrated in the Figure 2.2.

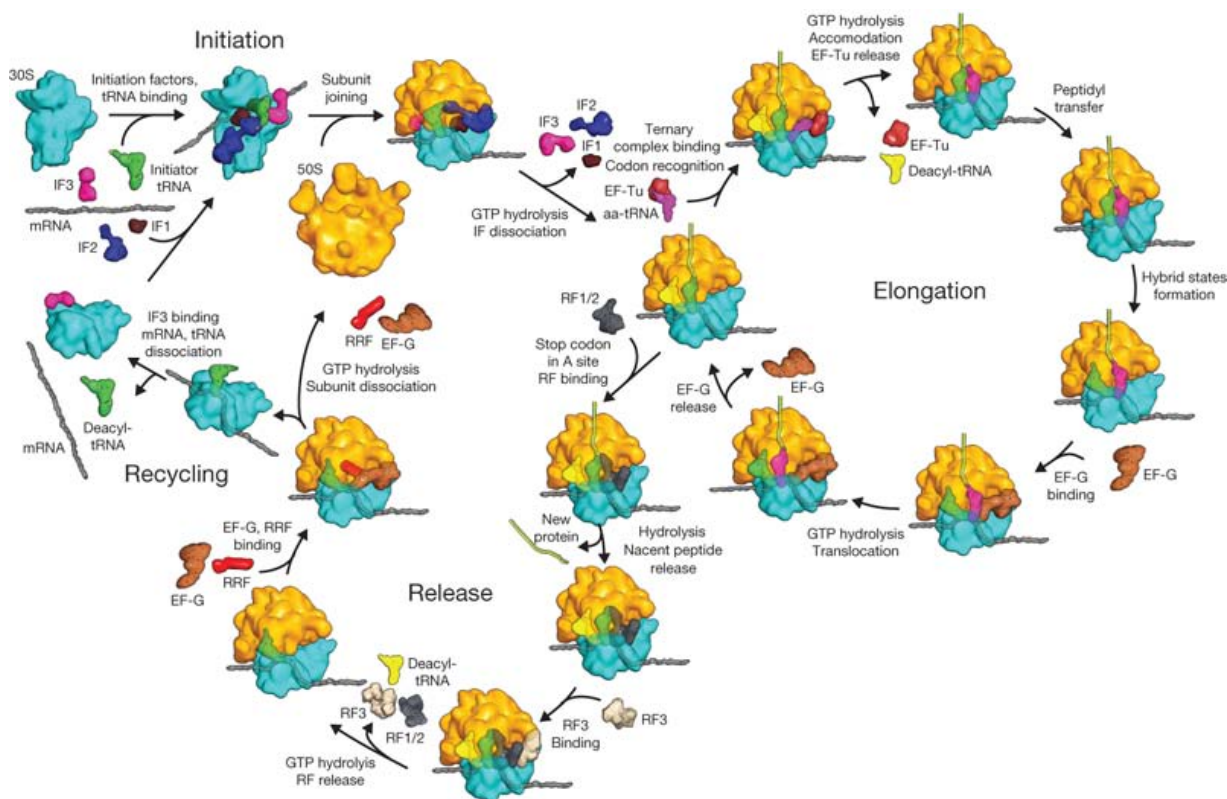


Figure 2.2 The prokaryotic translation cycle (Schmeing and Ramakrishnan, 2009)

For simplicity, not all intermediate steps are shown. The four major phases of the translation (initiation, elongation, termination and ribosome recycling) are discussed in the main text.

2.2.1 Prokaryotic translation initiation

The initiation of prokaryotic translation begins with the binding of the 30S ribosomal subunit to the initiation factor 3 (IF3), which leads the dissociation of ribosomes into subunits and couples translation initiation and ribosome recycling (Gualerzi and Pon, 1990). Initiation factor 1 (IF1) stimulates the activity of the IF3, specifically interacts with bases of the A site of the 30S subunit, and thus indirectly guides the initiator tRNA (fMet-tRNA^{fMet}, a tRNA carrying a formylmethionine) to the ribosomal P site by blocking the A site (Laursen et al.,

2005). The mRNA associates with the 30S subunit through complementary base pairing between its Shine-Dalgarno (SD) sequence and the anti-SD sequence of the 16S RNA, and hence the initiation codon (AUG) is adjusted in the P site of the 30S subunit (Yusupova et al., 2001). The initiator tRNA is accurately positioned at the P site of 30S subunit by the promotion of initiation factor 2 (IF2), where the binding of initiator tRNA to the P site is further stabilized by the IF3 (Laursen et al., 2005). In the presence of GTP, the GTPase activity of the IF2 is activated upon association of the 50S subunit to the 30S initiation complex that gives rise to the 70S initiation complex, during which process the IF1 and IF3 are ejected, GTP is hydrolyzed to GDP and phosphate, and the IF2 is released (Brock et al., 1998). At this stage the initiator tRNA is ready to form the first peptide bond with the second coded aminoacyl-tRNA (Brock et al., 1998).

2.2.2 Prokaryotic translation elongation

An initiator tRNA in the P site and an empty A site of ribosome coming from the end of the initiation process serves to initiate translation elongation. The second coded aminoacyl-tRNA is brought to the A site as a ternary complex with elongation factor Tu (EF-Tu) and GTP (Ramakrishnan, 2002). The recognition of the anticodon of the second aminoacyl-tRNA with the mRNA codon causes conformational changes in the ribosome which stabilizes tRNA binding and stimulates GTP hydrolysis by EF-Tu (Ramakrishnan, 2002). The resulting EF-Tu:GDP complex exhibits a low binding affinity for the aminoacyl-tRNA, which is then released from the A site of ribosome (Dell et al., 1990). A peptide bond between the initiator tRNA from the P site and the aminoacyl-tRNA accommodated in the A site is formed through the peptidyl transferase reaction that takes place in the 50S subunit, in which 23S rRNA is viewed as the catalytic element (Nissen et al., 2000; Noller et al., 1992). During the peptide bond formation, the initiator tRNA from the P site is deacylated and the peptidyl-tRNA is transferred to the A site (Ramakrishnan, 2002). Next, the deacylated tRNA is translocated to the E site from the P site and the peptidyl-tRNA to the P site from the A site, where translocation is catalyzed by the elongation factor G (EF-G) (Rodnina et al., 1997). As a result, the ribosome is ready for the next cycle of elongation, with deacylated tRNA and peptidyl-tRNA in the P site, and an empty A site to accommodate the next cognate ternary complex (Ramakrishnan, 2002).

2.2.3 Prokaryotic translation termination and ribosome recycling

The translation elongation cycle is repeated until a stop codon is encountered on mRNA in the A site of ribosome (Ramakrishnan, 2002). Stop codons are recognized by “class I” release factors (RF): RF1 recognizing UAA and UAG; RF2 recognizing UAA and UGA (Kisselev and Buckingham, 2000). The RF1 and RF2 trigger the hydrolysis of the ester bond in peptidyl-tRNA and release the newly synthesized peptide chain from the ribosome (Kisselev and Buckingham, 2000). The “class II” release factor, RF3, binds to the RF1-RF2 complex and induces the release of the RF1 and RF2 from the ribosome at the end of termination process (Ramakrishnan, 2002).

After the release of the peptide chain, the complex formed at the end of the termination process, comprising the 70S ribosome, mRNA and deacylated tRNA in the P site, is disassembled by the ribosome recycling factor (RRF) and EF-G to prepare the ribosome for a new round of protein synthesis (Janosi et al., 1996). During this ribosome recycling process, RRF and EF-G trigger the dissociation of the ribosome into 30S and 50S subunits on the basis of GTP hydrolysis (Karimi et al., 1999). Subsequently, the IF3 removes the deacylated tRNA from the 30S subunit, and all translational components are free for the next round of translation (Ramakrishnan, 2002).

2.3 The coupling between prokaryotic transcription and translation machineries

Transcription and translation are two highly coupled processes during prokaryotic gene expression where ribosomes initiate translation on mRNAs already during transcription (Laursen et al., 2005). As one strategy to communicate directly with one each other in transcription and translation machineries, prokaryotes make use of “moonlighting” proteins that can be shared in more than one cellular context (Jeffery, 1999). This dual activity of proteins was first noted in phage λ transcription antitermination system (Friedman et al., 1981).

2.3.1 Transcription antitermination systems

During lytic phase, early gene transcription of the phage λ genome initiates at pL and pR promoters proceeding in opposite directions (Figure 2.3), in which transcription on the left transcribes the *N* gene and stops on the tL₁ termination site, and about 50 % of transcripts on

the right stops on the tR_1 termination site with the remainder continuing and terminating at the combined $tR_{2,3,4}$ terminators in the *nin* region (Friedman and Court, 1995). In the presence of N-dependent processive transcription antitermination system, transcription can overcome these termination barriers and allow phage λ to switch from early to delayed early gene expression (Friedman and Court, 1995).

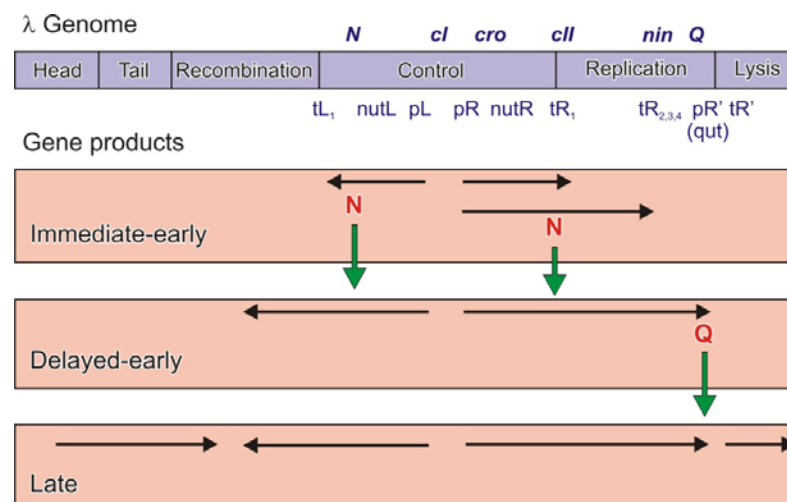


Figure 2.3 Map of the regulatory region of phage λ lytic phase

A representative collection of genes are shown at the top of the figure. The following lists the activities of the products of gene not discussed in details in the text: *cl* and *cro*, repressors; *cII*, transcription activator; N, phage λ protein involved in early gene expression; Q, phage λ protein involved in late gene expression; pR' and tR' , promoter and terminator for late gene expression. Blue, λ genome; Pink, gene products; Green arrow, phage λ gene expression switch; Black arrow, directions of phage λ gene expression. Figure is modified according to (Friedman and Court, 1995).

The processive transcription antitermination system of phage λ relies on the phage-encoded protein N, an RNA control sequence (N-utilization site, Nut; comprising two linear elements, BoxA and a “spacer”, followed by a stem loop, BoxB) and four host N-utilization substances (NusA, NusB, NusE and NusG) (Figure 2.4 (Left); (Friedman and Court, 1995; Friedman and Gottesman, 1983)). Phage λ N protein belongs to a family of proteins containing an arginine rich motif of about 6-10 amino acids which directly interacts with BoxB RNA (Legault et al., 1998). The NusA-binding region (amino acids 34-47) of N protein suppresses NusA’s enhancement of termination, and the C-terminal region of that makes contacts with RNAP (Mogridge et al., 1998a). NusA consists of three functional domains: the N-terminal domain that binds β and β' of RNAP (Mah et al., 1999); the RNA-binding domain that comprises an S1 motif and two KH motifs (Mah et al., 2000); the

regulatory domain encompassing the acidic repeats AR1 for N binding and AR2 (Bonin et al., 2004a). NusG is a conserved regulatory protein comprising two largely independent N- and C-terminal domains (NTD and CTD; (Steiner et al., 2002)), where the NTD interacts with RNAP, and the CTD interacts with the Rho and other Nus factors (Mooney et al., 2009). NusE is identical to r-protein S10 (Friedman et al., 1981). NusE forms a stable complex with NusB (Mason et al., 1992) that has enhanced affinity for BoxA-containing RNAs compared to NusB alone (Luttgen et al., 2002; Mogridge et al., 1998b; Nodwell and Greenblatt, 1993). N, Nus RNA and the Nus factors form a ribonucleoprotein complex on the surface of RNAP, in which RNA and protein factors engage in numerous, predominantly weak and cooperative contacts (Mogridge et al., 1995). The N-Nus factor complex accompanies RNAP during elongation *via* RNA looping (Whalen and Das, 1990) and promotes processive transcription elongation through downstream intrinsic and factor-dependent termination sites (Weisberg and Gottesman, 1999).

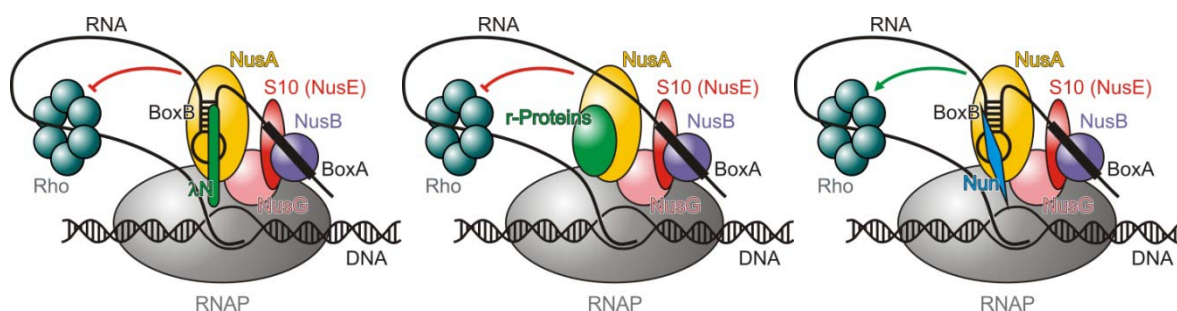


Figure 2.4 Transcription antitermination or termination models

(Left) Model of phage λ N-dependent transcription antitermination. (Middle) Model of *E. coli* ribosomal RNA transcription antitermination. (Right) Model of phage HK022 transcription termination. The Nus factors, NusA (yellow), NusB (blue), NusE (red) and NusG (pink), are involved in all three models. In *E. coli* ribosomal RNA transcription antitermination, BoxB-like element is dispensable and one of r-proteins (green) participates. λ N, green; Nun, cyan; RNAP, grey; Rho, maroon; DNA and RNA, black. Rho factor is blocked (represented by a red curve) in phage λ N-dependent transcription antitermination and in *E. coli* ribosomal RNA transcription antitermination. Rho factor is provoked (represented by a green arrow) to terminate phage HK022 transcription.

Other bacteria like *Escherichia coli* (*E. coli* or *eco*) utilize a similar mode of processive antitermination during their ribosomal RNA gene (*rrn*) transcription (Figure 2.4 (Middle)), in that the same factors, NusA, NusB, NusE and NusG, are involved (Li et al., 1984; Quan et al., 2005). BoxA RNA is strictly conserved in all seven *rrn* operons of *E. coli*, whereas the BoxB-like element is dispensable for *rrn* antitermination (Berg et al., 1989). However, there

is no known analogue of N itself, so either the analogue has not been found or one of r-proteins participates in *rrn* antitermination (Roberts et al., 2008). In addition to Nus factors, other r-proteins, including S2, S4, L4 and L13, participate in this latter process for transcription antitermination (Torres et al., 2004; Torres et al., 2001).

A relative of λ , phage HK022, expresses the Nun protein which is a transcription factor related to the λ N protein (Friedman and Court, 1995). Nun protein acts at the λ Nut BoxB site and, after enlisting the four Nus factors, provokes transcription termination (Figure 2.4 (Right); (Robert et al., 1987)). Moreover, the Nun protein competes with N at λ Nut sites and represses phage λ N transcription in order to avoid superinfection with phage λ of a bacterial cell that is already HK022 infected, thus securing phage HK022 survival (Robert et al., 1987).

2.3.2 Proteins shared by the transcription and translation machineries

Prokaryotic transcription and translation leading to gene expression communicate directly with one another by sharing proteins (Squires and Zaporjets, 2000). More than 60 proteins are required to fine-tune the transcriptional and translational processes, providing ample candidates for proteins to be shared between two activities (Squires and Zaporjets, 2000). For example, three r-proteins, S10, L4 and S1, have been clearly discovered to participate in both transcription and translation, and two transcription antitermination factors, NusG and NusB, are implicated in translation through genetic mutation studies but their exact roles in translation are still not well understood.

2.3.2.1 S10 (NusE)

S10 was initially defined as an r-protein before an additional role in transcription was discovered (Friedman et al., 1981). It is an important architectural element in the 30S ribosomal subunit (Figure 2.5A), as revealed by reconstitution (Mizushima and Nomura, 1970) and crystal structure analyses (Schlunzen et al., 2000; Wimberly et al., 2000). S10 is one of last six r-proteins involved in the final step of 30S ribosomal subunit assembly (Squires and Zaporjets, 2000). In the 30S ribosomal subunit, S10 exhibits a globular domain that is located at the surface of the particle and an extended ribosome-binding loop that deeply penetrates the subunit and interacts with several other r-proteins and the 16S rRNA (Figure 2.5A and 2.5B; (Schlunzen et al., 2000; Wimberly et al., 2000)). It was suggested

that the fold of S10 by itself is unstable (Das et al., 2008; Gopal et al., 2001); thus, the other r-proteins and 16S rRNA may act to stabilize S10 in the ribosome.

The participation of S10 in transcription represents a first example of an r-protein involved in two cellular machineries (Friedman et al., 1981). The role of S10 in transcription antitermination is highly cooperative to NusB *in vivo*, where NusB and S10 form a complex for BoxA RNA binding (Figure 2.4), suggesting that S10 is involved in the formation of functional transcription antitermination complex (Nodwell and Greenblatt, 1993). The addition of S10 increases the efficiency of terminator read-through in an *in vitro* rRNA transcription antitermination system (Squires et al., 1993; Squires and Zaporozjets, 2000). In addition, it was also found that S10 directly contacts RNAP (Mason and Greenblatt, 1991) and it has been shown to bind phage λ N *in vitro* (Mogridge et al., 1995, 1998b).

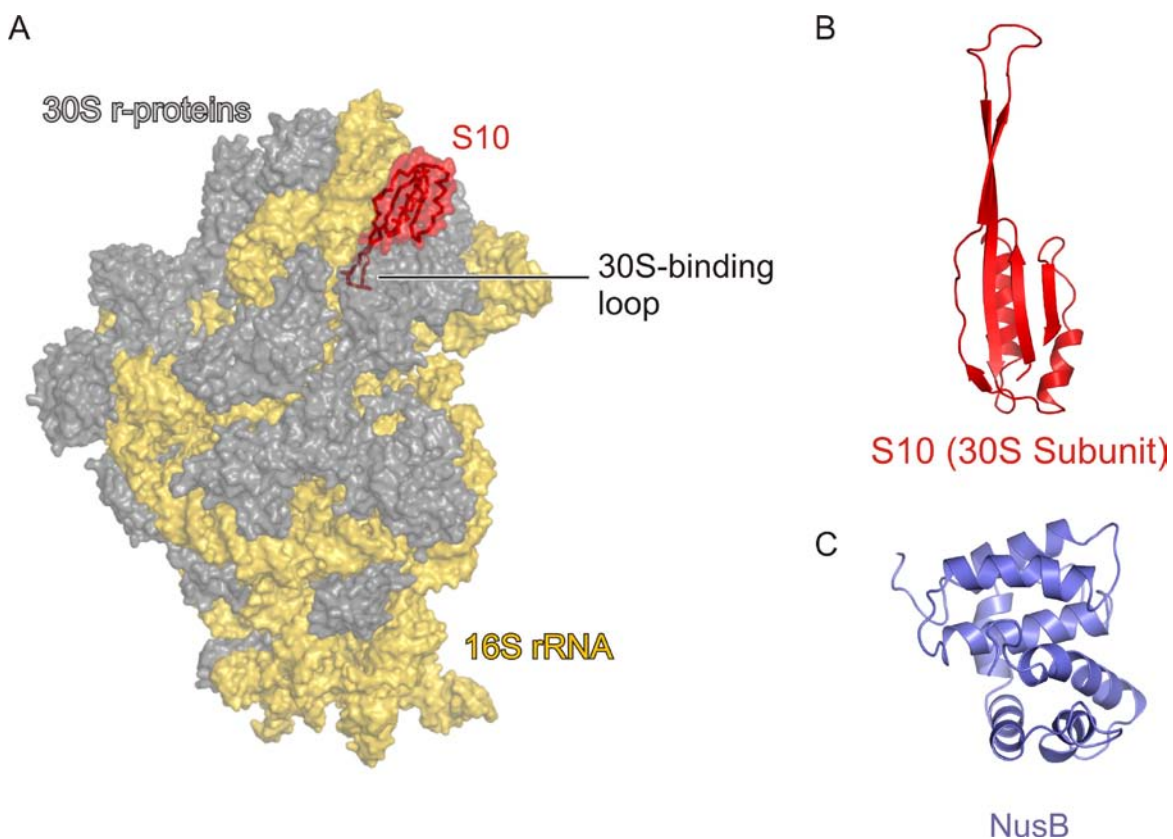


Figure 2.5 Structures of S10 and NusB

(A) The global view of *E. coli* S10 in 30S ribosomal subunit (PDB ID 2AVY; (Schuwirth et al., 2005)). S10, red; 30S r-proteins, grey; 16S rRNA, gold.

(B) Ribbon plot of the crystal structure of S10 from *E. coli* 30S ribosomal subunit (PDB ID 2AVY; (Schuwirth et al., 2005)).

(C) Ribbon plot of the NMR structure of *E. coli* NusB (PDB ID 1EY1 (Altieri et al., 2000))

2.3.2.2 L4

L4 participates in the early assembly of the 50S ribosomal subunit in which L4 fixes the tertiary structure of the 23S rRNA (Nierhaus, 1991; Worbs et al., 2000). L4 has a globular domain that sits on the surface of the 50S subunit and an extended loop penetrates the core of the 50S subunit (Ban et al., 2000; Zengel et al., 2003). L4 is involved with the peptidyl transferase RNA region and may participate in the catalysis of peptide bond formation (Worbs et al., 2000).

Some r-proteins have a function as regulators, autogenously inhibiting expression of their own operons when they are produced in excess of available binding sites on nascent rRNA during ribosomes assembly (Zengel et al., 2003; Zengel and Lindahl, 1994). These r-protein operons feedback translational regulations by a particular r-protein encoded in the operon (Squires and Zaporjets, 2000). L4 is one component of the S10 operon which contains genes for eleven r-proteins (Lindahl and Zengel, 1986). Like other r-protein operons, the S10 operon is autogenously regulated by one of its products, L4, which inhibits translation by preventing initiation of translation of the most proximal gene of the S10 operon (Zengel and Lindahl, 1992). Regulation of the S10 operon by L4 occurs not only at the translation level, but also at the transcription level. L4 inhibits the transcription of the S10 operon, with the cooperation of NusA, by leading transcription termination at a particular site in the S10 operon leader region, where NusA stabilizes the RNAP pause, and L4 further reinforces this pause and converts RNAP into a termination activity (Squires and Zaporjets, 2000; Zengel and Lindahl, 1992).

2.3.2.3 S1

S1 is one of last proteins involved in the 30S ribosomal subunit assembly. It recognizes the nascent mRNA structures and opens up these mRNA structures to initiate translation by ribosomes (Squires and Zaporjets, 2000). S1 is the largest r-protein (mass, 70 kDa) with an NTD for ribosome binding and with a CTD formed by six copies of approximately ~70 amino acids for mRNA binding (Gribskov, 1992).

S1 is involved in the cellular transcription on the basis of the finding that it competitively inhibits the binding of NusB and S10 to the BoxA RNA during the *rrn* antitermination (Mogridge and Greenblatt, 1998). The affinity of the *rrn* BoxA RNA for S1 is 200-fold-higher than that of for NusB-S10 complex that suggests that S1 might be an inhibitor of transcription antitermination (Mogridge and Greenblatt, 1998). The competitive binding ability of S1 for phage λ Nut BoxA RNA was identified in the same manner, but it did not

inhibit the λ N-dependent transcription antitermination *in vitro* in reactions containing other antitermination factors (Mogridge and Greenblatt, 1998; Squires and Zaporojets, 2000). The transcription antitermination roles of S1 in both ribosomal RNA and phage λ N transcription systems have to be further studied.

2.3.2.4 NusG

NusG is a transcription elongation factor that was originally discovered as a key component in the phage λ N-dependent antitermination complex (Figure 2.4), where NusG interacts directly with RNAP (Li et al., 1992; Squires and Zaporojets, 2000). NusG inhibits transcription pausing and increases the rate of elongation (Burova et al., 1995). NusG has also been shown to directly bind the Rho factor (Pasman and von Hippel, 2000) and stimulate Rho-dependent termination (Sullivan and Gottesman, 1992). Thus, NusG establishes a bridge between RNAP and Rho to help recruit the Rho into the termination complex in a way independent of its effect on elongation (Li et al., 1993).

A role for NusG in translation has been identified by the finding that the peptide elongation rate *in vivo* is reduced in the *nusG*-depleted cells by measuring rate of synthesis of a *lacZ* construct (Zellars and Squires, 1999). Thus, NusG was viewed to serve as a linker to couple the rate of transcription and the rate of translation (Zellars and Squires, 1999). Moreover, all known members of the NusG family at the C-terminus of proteins carry a KOW motif, a 27-amino acids sequence with a glycine at position 11, which is highly conserved in r-protein families RL24, RL26 and RL27 (Squires and Zaporojets, 2000). The phenomenon that a transcription factor has been linked by a sequence motif to r-protein families suggests that NusG potentially participates in translation (Squires and Zaporojets, 2000).

2.3.2.5 NusB

The *nusB* gene was discovered by Shiba et al when they were seeking the extragenic suppressors of *secY24* mutation that causes a defect in secretion (Shiba et al., 1986). The NMR structure of *E. coli* NusB protein shows that NusB is composed of five helices and adopts an all helical fold (Figure 2.5C; (Altieri et al., 2000)). NusB is implicated in translation elongation on the basis of characterizations of several mutations in *nusB* (Court et al., 1995). One of these *nusB* mutations, *nusB::IS10*, suppresses the *secY24* defect, leads to a cold-sensitive growth defect in *E. coli* cells and slows down the peptide chain elongation rate by 30% (Court et al., 1995; Shiba et al., 1986; Taura et al., 1992). As most mutations that

suppress a *secY* mutation have been identified in genes whose products are related to protein synthesis, NusB was speculated to play a role in translation (Squires and Zaporjets, 2000).

Much evidence indicates NusB is an important transcription antitermination factor during λ N-mediated antitermination (Figure 2.4). NusB forms a stable complex with S10 even in the absence of other Nus factors (Mason et al., 1992) to sustain unstable S10 a proper fold in transcription antitermination (Das et al., 2008; Gopal et al., 2001). NusB alone interacts specifically with BoxA RNA, and the enhanced binding affinity is achieved by the addition of S10 (Nodwell and Greenblatt, 1993). Since BoxA is strictly conserved in all seven *rrn* operons of *E. coli*, where the BoxB-like element is dispensable for *rrn* antitermination (Berg et al., 1989), association of NusB, S10 and BoxA is considered as a key nucleation event during processive antitermination (Greive et al., 2005).

2.4 Aims of this study

There is much evidence showing S10 participates in transcription and translation, but presently, it is still unclear how S10 is reprogrammed as a transcription factor. In particular, it is unknown how S10 interacts with NusB, whether the conformation of S10 in transcription is different from that in the 30S subunit (Gopal et al., 2001), whether the protein can remain part of the ribosome while participating in antitermination (Das et al., 1985) and why the NusB-S10 complex exhibits enhanced affinity for BoxA RNA.

The effects of *nusB* mutations on translation are certainly indirect evidence that NusB participates in translation. However, the exact role of NusB in translation remains unknown and the dual activity of NusB in transcription and translation is still not understood.

Mutations in *nusB* and *nusE* have served as important genetic tools to study processive antitermination. Some mutations were found to affect antitermination activities. For instances, the *nusB5* mutation leads to a defect in N-dependent antitermination that blocks λ growth (Friedman et al., 1976); the *nusB101* mutation suppresses the N antitermination defects of *nusA1* and *nusE71* mutations at high temperatures (Ward et al., 1983); the *nusE71* mutation blocks N antitermination λ growth at high temperatures (Friedman et al., 1981); the *nusE100* mutation restricts Nun termination but not N antitermination (Robledo et al., 1991). However, the biochemical basis for the dysfunction or suppressor activity of any of the mutant proteins was not defined.

The aims of this study are to answer these questions and to define the dual roles of S10 and NusB during transcription and translation processes.

3 Materials and Methods

3.1 Materials

3.1.1 Chemicals

Acetic acid	Merck, Darmstadt
Acrylamide solution	Roth, Karlsruhe
Agarose	Invitrogen, USA
Ammonium persulfate (APS)	Merck, Darmstadt
Bovine serum albumin (BSA), acetyliert	Sigma, Deisenhofen
Bradford solution	Biorad, München
Bromophenol blue	Merck, Darmstadt
Coomassie brillant blau R250	Serva, Heidelberg
Dimethyl sulfoxide (DMSO)	Sigma, Deisenhofen
Dithiothreitol (DTT)	Roth, Karlsruhe
DNA ladder (1 kb)	Invitrogen, USA
Ethanol	Merck, Darmstadt
Ethylendiamine tetra-acetic acid (EDTA)	Roth, Karlsruhe
Ethidium bromide solution	Roth, Karlsruhe
Glutathione (reduced)	Sigma, Deisenhofen
Glycin	Merck, Darmstadt
Glycerol	Merck, Darmstadt
HEPES	Calbiochem, USA
Imidazole	Merck, Darmstadt
Isopropyl- β -D-thiogalactoside (IPTG)	Sigma, Deisenhofen
Lysozyme	Boehringer, Mannheim
Milk powder, instant	Cenovis GmbH, Radolfzell
Magnesium chloride	Merck, Darmstadt
Methanol	Merck, Darmstadt
Polyethylenglycol 3350 (PEG3350)	Sigma, Deisenhofen
Ponceau S	Serva, Heidelberg
Potassium chloride	Merck, Darmstadt
Precision protein standard marker	Biorad, München

Roti-Phenol/Chloroform	Roth, Karlsruhe
Silver nitrate	Merck, Darmstadt
Sodium chloride	Merck, Darmstadt
Sodium dodecyl sulfate (SDS)	Merck, Darmstadt
Sodium thiosulfate	Merck, Darmstadt
N, N, N', N'-Tetramethylethyldiamin (TEMED)	Sigma, Deisenhofen
tRNA, <i>E. coli</i>	Boehringer, Mannheim
Tris-(hydroxymethylen) aminomethan	Roth, Karlsruhe
Triton X-100	Sigma, Deisenhofen
Tween 20	Sigma, Deisenhofen
Xylene cyanol FF	Fluka, Schweiz

Standard chemicals, organic substances and solvents (purification grade p.a.), which are not listed here, were ordered from one of the following companies: Merck (Darmstadt), Roth (Karlsruhe), Sigma (Taufkirchen), Serva (Heidelberg) or Fluka (Switzerland).

3.1.2 Media

Auto-inducing medium	Own production
LB-medium	Q-Biogene, USA
Luria-Bertani-broth (LB)-Agar	Q-Biogene, USA

3.1.3 Antibiotics

Ampicillin	Sigma, Deisenhofen
Chloramphenicol	Boehringer, Mannheim
Kanamycin sulphate	Sigma, Deisenhofen

3.1.4 Nucleotides

Deoxynucleoside-5'-Triphosphate (dNTPs, 100 mM)	Amersham, Freiburg
---	--------------------

3.1.5 Radionucleotides

[$\gamma^{32}\text{P}$]-ATP (6000 Ci/mmol, 10 Ci/l)	Amersham, Freiburg
---	--------------------

3.1.6 Antibodies

Rabbit anti-GST antibody	Invitrogen, USA
Goat anti-rabbit IgG	Dianova, Hamburg

3.1.7 Enzymes and inhibitors

DNase I	Roche, Mannheim
<i>Pfu</i> DNA polymerase (2,5 U/ μ l)	Stratagene, Heidelberg
PreScission protease	Own production
Proteinase inhibitor cocktail complete™, EDTA-free	Roche, Mannheim
Restriction endonucleases	New England Biolabs, France
RNAsin (40 U/ μ l)	Promega, USA
T4 DNA ligase (400 U/ μ l)	New England Biolabs, France
T4 polynucleotide kinase (20 U/ μ l)	New England Biolabs, France
Taq DNA polymerase (5000 U/ μ l)	Promega, USA
TEV-protease	Own production

3.1.8 DNA oligonucleotides

Synthetic DNA oligonucleotides (Table 3.1) were purchased from MWG/Operon (Ebersberg, Germany).

Table 3.1 DNA oligonucleotides

Protein	Description	Oligo	Sequence (5'→3')
<i>ecoNusB</i>	pETM11; E2K; Full-length	Forward	ACGTACCCATGGAACCTGCTGCTCGTCGCCGCGC
		Reverse	ACGTACGGTACCTCACTTTTTGTTAGGGCGAATCACAG
	pBAD; Full-length	Forward	TATCCGTCTCCCATGAAACCTGCTGCTCGTCGCC
		Reverse	AGCCTCGAGTCACTTTTTGTTAGGGCGAATCACAGG
<i>ecoNusE</i>	pGEX-6p-1; Full-length	Forward	CGCGGATCCATGCAGAACCAAGAATCCGTATCC
		Reverse	CGCGAATTCTTAACCCAGGCTGATCTGCACGTC
	pBAD; Full-length	Forward	TATCCGTCTCCCATGCAGAACCAAGAATCCGTATCCGCTG
		Reverse	AGCCTCGAGTTAACCCAGGCTGATCTGCACGTC
<i>ecoNusEΔ</i>	pGEX-6p-1; AA46-67S	Forward	CCGATCCCCTGCGGACACGCAGCCGTACTCACTTGCGTCTGG
		Reverse	CCAGACGCAAGTGAGTACGGCTGCGTGTCGGCAGCGGGATCGG
	pBAD; AA46-67S	Forward	TATCCGTCTCCCATGCAGAACCAAGAATCCGTATCCGCTG
		Reverse	AGCCTCGAGTTAACCCAGGCTGATCTGCACGTC

3.1.9 RNA oligonucleotides

Synthetic RNA oligonucleotides (Table 3.2) were purchased from Dharmacon (Lafayette, USA).

Table 3.2 RNA oligonucleotides

Oligo	Description	Sequence (5'→3')
<i>rrn</i> BoxA RNA	<i>E. coli</i> ; 19mer	CACUGCUCUUUAACAAUUA
NutR BoxA RNA	Phage λ ; 19mer	CACCGCUCUUACACAAUUA
NutR BoxA RNA	Phage λ ; 12mer; 5- ^{Br} U labeled	CGC ^{Br} UCUUACACAAUUA

3.1.10 Vectors

pBAD vector was used to express proteins for *in vivo* complementation and recombineering analysis (Table 3.3). pETM11 and pGEX-6p-1 vectors were used to express proteins for crystallization and biochemical assays (Table 3.3).

Table 3.3 Vectors

Vector	Description	Source
pBAD	Expression vector; His-Tag; araBAD promoter; Amp ^r	Invitrogen
pETM11	Expression vector; His-tag; T7 promoter; Kan ^r	Novagen
pGEX-6P-1	Expression vector; GST-tag; tac promoter; Amp ^r	GE Healthcare

3.1.11 Plasmids (Table 3.4)

The plasmids generated by site directed mutagenesis were not listed.

Table 3.4 Plasmids

Plasmid	Description
pBAD- <i>eco</i> NusB	Cleavage sites: <i>NcoI</i> → <i>XhoI</i> ; Full-length
pBAD- <i>eco</i> NusE	Cleavage sites: <i>NcoI</i> → <i>XhoI</i> ; Full-length
pBAD- <i>eco</i> NusE Δ	Cleavage sites: <i>NcoI</i> → <i>XhoI</i> ; AA 46-67 were replaced with a serine
pETM11- <i>eco</i> NusB	Cleavage sites: <i>NcoI</i> → <i>Acc65I</i> ; Full-length
pGEX-6P-1- <i>eco</i> NusE	Cleavage sites: <i>BamHI</i> → <i>EcoRI</i> ; Full-length
pGEX-6P-1- <i>eco</i> NusE Δ	Cleavage sites: <i>BamHI</i> → <i>EcoRI</i> ; AA 46-67 were replaced with a Serine

3.1.12 Bacterial strains

<i>E. coli</i> BL21(DE3)	Novagen, Darmstadt
<i>E. coli</i> DH5 α	Invitrogen, USA
<i>E. coli</i> XL-1 blue	Stratagene, Heidelberg
<i>E. coli</i> 9739	Max Gottesman

E. coli 9976 (*nusB::Cam*)

Max Gottesman

3.1.13 Commercial kits

ECL Western blot detection kit	Amersham, Freiburg
Pre-crystallization test kit	Hampton Research, USA
QIAprep spin miniprep kit	Qiagen, Hilden
QIAquick gel extraction kit	Qiagen, Hilden
QIAquick PCR purification kit	Qiagen, Hilden
Stratagene QuikChange™ kit	Stratagene Amsterdam

3.1.14 Crystallization screens

Amonium sulfate screen	Qiagen, Hilden
Anions and cations suites	Qiagen, Hilden
Classics and classics lite	Qiagen, Hilden
Crystal screen I and II	Hampton Research, USA
Index I and II	Hampton Research, USA
JCSG screen	Qiagen, Hilden
Mb class I and II	Qiagen, Hilden
MPD suite	Qiagen, Hilden
Nucleix suite	Qiagen, Hilden
PACT screen	Qiagen, Hilden
PEG I and II	Qiagen, Hilden
pH clear I and II	Qiagen, Hilden
Protein complex screen	Qiagen, Hilden
SM I, II and III	Qiagen, Hilden
Salt Rx screen	Hampton Research, USA

3.1.15 Equipments

Anode X-ray generator (in-house source)	Rigaku, Tokyo
Åkta explorer/prime/purifier and columns	Amersham, Freiburg
Biofuge (pico/fresco)	Heraeus, Hanau
Cartesian NanoDrop robot	Zinsser Analytik, Frankfurt
Electrophoresis appartaus	BiorRad, München
Gel documentation system	Biorad, München

Heating block	Hybaid Biometra, UK
Head over tail rotor 7637-01	Cole-Parmer, USA
Image plate detector (in-house source)	MAR Research, Norderstedt
Incubator BK-600	Heraeus, Hanau
Incubation shaker Multitron	Infors, Switzerland
Multi-well filtration manifold	Biorad, München
pH Meter	MettlerToledo, Switzerland
Phosphorimager Typhoon 8600	Molecular Dynamics
Scintillation counter LS	Beckman/Packard, USA
SMART system	Pharmacia Biotech
Sonifier	Heinemann Labortechnik
Sorvall rotor	Kendro, USA
SpeedVac concentrator 5301	Eppendorf, Hamburg
Spectrophotometer Ultropsec 300 pro	Amersham, Freiburg
SW60 rotor	Beckman, USA
Synchrotron beamline 14-2	BESSY, Berlin
Synchrotron beamline PXI/II	SLS, Villigen
Trans-Blot electrophoresis transfer cell	Bio-Rad, München
Ultracentrifuge	Sorvall/Beckman, USA
UV lamp 254 nm	Bachofer, Reutlingen
Thermal cycler	Hybaid Omni Gene, UK
Vortex	Janke & Kunkel, Staufen i. Br.
X-ray film developer X-Omat 2000	Kodak, USA

3.1.16 Consumption materials

Amicon centriplus concentrator	Millipore, France
Chemiluminescence film	Amersham, Freiburg
Cuvettes for monolight 3010	Pharmingen, USA
Collodium bags	Sartorius GmbH, Göttingen
Dialysis cassettes	Pierce, USA
Electroporation cuvettes	Bio-Rad, München
Falcon tubes (5, 15, 50 ml)	Greiner, Kremsmünster
Glass beads (425-600 microns)	Sigma, Deisenhofen
Glutathione sepharose 4B	Amersham, Freiburg

Ni-NTA agarose	Quiagen, Hilden
Nylon membrane hybond-(N+)	Qiagen, Hilden
Pipettes	Eppendorf, Hamburg
Probe Quant™ G-25 micro columns	Amersham, Freiburg
Protran nitrocellulose membrane	Schleicher & Schuell, Dassel
Reaction tubes (0.5; 1.5; 2 ml)	Eppendorf, Hamburg
Sterile filter (0.2; 0.45 µm)	Millipore, France
Talon metal affinity resin	Clontech, Heidelberg
Vivaspin concentrators	Vivascience, Sartorius
X-ray film BioMax MR	Kodak, USA

3.2 Methods

3.2.1 Molecular cloning

3.2.1.1 PCR amplification

Polymerase chain reaction (PCR) was used for target amplification from *E. coli* genomic DNA and plasmid construction. Both forward and reverse primers (Table 3.2) were designed to introduce compatible restriction enzyme sites and 3-6 additional bases were added before these sites to allow efficient digestion by restriction enzymes. The annealing temperature was chosen on the basis of the melting temperatures of the primers. A typical PCR reaction and cycling programme are shown below:

PCR reaction mixture (50 µl)

1 µl	DNA sample (100 ng/µl)
5 µl	10x <i>Pfu</i> buffer
5 µl	DMSO
1 µl	5' primer (20 pmol/µl)
1 µl	3' primer (20 pmol/µl)
4 µl	dNTP (10 mM each)
2 µl	<i>Pfu</i> polymerase
31 µl	H ₂ O

PCR cycling programme

94 °C	2'	} 30 repetitions
94 °C	15"	
60 °C	30"	
72 °C	1'	
72 °C	5'	
4 °C	hold temperature	

3.2.1.2 Agarose gel electrophoresis and DNA fragment isolation

Agarose gel electrophoresis was performed for the analysis of PCR products. PCR samples were mixed with 5x DNA loading buffer and loaded to a 1.5 % agarose gel. A 1-kb DNA ladder at the concentration of 0.05 mg/ml was loaded in one lane as a marker. Gel was then run in 1x TBE buffer at 50-100 V and stained in 0.5 µg/ml ethidium bromide. DNA was visualized under UV light. QIAquick gel extraction kit was used for DNA fragments isolation from agarose gel. The band of interest on the agarose gel was cut out and treated according to the manufacturer's protocol.

10x TBE buffer (pH 8.3)

1 M	Tris base
0.83 M	Boric acid
10 mM	EDTA

5x DNA loading buffer

30 % (v/v)	Glycerol
0.25 % (w/v)	Bromophenol blue
0.25 % (w/v)	Xylene cyanol FF

3.2.1.3 Enzyme digestion and ligation

For the ligation reaction both vector DNA and the insert DNA were digested with appropriate restriction enzymes, and then purified with QIAquick PCR Purification Kit (Qiagen) according to the manufacturer's protocol. 3:1 to 5:1 molar ratio of insert to vector was performed in the reaction to achieve the optimal ligation efficiency. The reaction mixture was incubated at 16°C for 3 hours, and then incubated at 65°C for 15 minutes to inactivate the enzymes. The sample was spun down briefly before the transformation. A typical ligation reaction is shown below:

Ligation reaction mixture (20 µl)

2 µl	10x Buffer for T4 DNA ligase
2 µl	Linearised vector DNA
8 µl	Insert DNA
7.5 µl	H ₂ O
0.5 µl	T4 DNA ligase (400 U/µl)

3.2.1.4 Preparation of competent cells for electroporation transformation

E. coli competent cells from manufacturers were grown in 1 L of LB medium until the OD₆₀₀ of 0.4-0.6 was reached. The cell culture was centrifuged at 4000 g for 15 min at 4°C. The pellet was washed twice with 500 ml of ice-cooled, sterilized water, once with 200 ml of ice-cooled, sterilized 10 % glycerol and once with 50 ml of ice-cooled, sterilized 10 % glycerol. The resulting pellet was resuspended in 4 ml of 10% glycerol, divided into 50 µl aliquots and then flash-frozen in liquid nitrogen and stored at -80 °C.

3.2.1.5 Preparation of competent cells for chemical transformation

E. coli competent cells from manufacturers were grown in 250 ml of LB medium until the OD₆₀₀ of 0.4-0.6 was reached. The culture was centrifuged at 4000 g for 15 min at 4 °C. The cell pellet was resuspended in 100 ml of ice-cooled TFB I buffer and incubated at 4 °C for 10 min. The resuspended cells were centrifuged at 4000 g for 5 min at 4°C. The pellet was resuspended in 10 ml of ice-cooled TFB II buffer and incubated on ice for at least 15 min. 100 µl aliquots were flash-frozen in liquid nitrogen and stored at -80 °C.

<u>TFB I</u>		<u>TFB II</u>	
30 mM	potassium acetate	10 mM	MOPS
100 mM	rubidium chloride	75 mM	calcium chloride
10 mM	calcium chloride	10 mM	rubidium chloride
50 mM	manganese chloride	15% v/v	glycerol
15% v/v	glycerol	Adjust pH to 6.5 with KOH and sterilize.	
Adjust to pH 5.8 with acetic acid and sterilize.			

3.2.1.6 Electroporation transformation

The plasmid was mixed with 50 µl of competent cells already thawed on ice. The mixture was transferred to a pre-chilled electroporation cuvette and subjected to a voltage of 1.8 kV (for cuvettes with 0.1 mm width) by the *E. coli* Pulser. After the resuspension in 950 µl of LB medium, the cells were grown at 37 °C for 1 h without any antibiotics. Subsequently, the cells were pelleted, resuspended in a small volume of fresh LB medium, streaked out on an agar plate containing the selective antibiotics and incubated at 37 °C overnight.

3.2.1.7 Chemical transformation

The plasmid was mixed with 100 µl of competent cells already thawed on ice. The mixture was incubated on ice for 20-30 min, heat shocked at 42°C for 90 sec and then cooled down on ice for 2 min. After the resuspension in 950 µl of LB medium, the cells were grown at 37 °C for 1 h without any antibiotics. Subsequently, the cells were pelleted, resuspended in a small volume of fresh LB medium, streaked out on an agar plate containing the selective antibiotics and incubated at 37 °C overnight.

3.2.1.8 Mini-preparation of plasmid and DNA sequencing

A single colony was picked up from an agar plate for the mini-preparation of plasmid DNA by the QIAprep spin miniprep kit according to the manufacturer's protocol. DNA

sequencing was performed by the Seqlab, Göttingen. The appropriate amounts of DNA and sequencing primers were supplied as suggested by the Seqlab.

3.2.1.9 PCR-based site directed mutagenesis

Site directed mutagenesis was performed by the Stratagene QuikChange™ kit according to the manufacturer's protocol. All the primers were designed by the online QuikChange primers design tool provided by Stratagene. A typical mutagenesis PCR reaction and cycling programme are shown below:

<u>PCR reaction mixture (50 µl)</u>		<u>PCR cycling programme</u>	
2 µl	DNA sample (10 ng/µl)	95 °C	30"
5 µl	10x <i>Pfu</i> buffer	95 °C	30"
1.5 µl	5' primer (10 pmol/µl)	55 °C	1'
1.5 µl	3' primer (10 pmol/µl)	68 °C	6'30"
4 µl	dNTP (10 mM each)	4 °C	hold temperature
1 µl	<i>PfuTurbo</i> DNA polymerase (2.5 U/µl)		
35 µl	H ₂ O		

} 16 repetitions

3.2.2 Protein Production

3.2.2.1 Co-expression of protein complexes

Equal amounts (20 ng) of two plasmids containing the genes of interest were co-transformed into an *E. coli* strain BL21(DE3). The cells were grown in auto-inducing medium (Studier, 2005) in the presence of appropriate antibiotics to an OD₆₀₀ of 0.5 at 37 °C, and then incubated for an additional 16 hours at 20 °C. After harvesting at 4 °C, the cell pellets were washed with binding buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl) and stored at -80 °C.

<u>Auto-inducing medium (1 L)</u>		<u>Per 950 ml ZY</u>	
ZY	950 ml	N-Z-Amine AS	10 g
50× 5052	20 ml	Yeast Extract	5 g
50× M	20 ml	Autoclave at 121°C for 15 min	
2 M MgSO ₄	1 ml		
1000× Metals	200 µl		

3.2.2.2 Cell lysis

Frozen cells were thawed in binding buffer and disrupted by a sonifier. Proper cooling was accomplished with a NaCl ice-bath and a number of short pulses (duration 5-10 s) with pauses (duration 10-30 s) to sustain a low temperature. Subsequently, the lysate was centrifuged at 4 °C and 15000 rpm for 30 min to pellet the cell debris. The supernatant was supplied to purification.

3.2.2.3 Co-purification of protein complexes

For purification, the cleared lysate was incubated with glutathione-sepharose equilibrated with binding buffer to trap the complex *via* the N-terminal GST-tag of the S10 or S10^{Δloop} protein. Proteins were eluted in a single step with binding buffer containing 15 mM reduced glutathione and then treated with PreScission protease overnight at 4 °C in order to remove the GST-tag. After PreScission cleavage, the protein complex was trapped *via* the N-terminal His₆-tag of the NusB protein on Ni²⁺-NTA-agarose equilibrated with binding buffer containing 20 mM imidazole, washed with 50 mM imidazole and eluted with 500 mM imidazole. During dialysis against binding buffer plus 2 mM DTT, proteins were treated with TEV protease overnight at 4 °C in order to remove the His₆-tag. After TEV cleavage and dialysis, the sample was passed again over Ni²⁺-NTA-agarose. The flow-through was concentrated by ultrafiltration and further purified by gel filtration on a Superdex-75 26/60 column equilibrated with crystallization buffer (10 mM Tris-HCl, pH 7.5, 50 mM NaCl, 2 mM DTT). Purified protein complex was concentrated by ultrafiltration to 16 mg/ml and stored at -80 °C after flash-freezing in liquid nitrogen.

3.2.2.4 Determination of protein concentrations

Protein solution was concentrated using an Amicon centriplus concentrator with an appropriate molecular weight cutoff (around 3 times less than the molecular weight of the respective protein). The protein concentration was determined with a Bradford assay: 1 μl of concentrated protein was mixed in 1 ml of 5× diluted Bradford solution; the absorbance at 595 nm was measured in a spectrophotometer; the BSA protein was employed to make a standard curve in the same manner; by comparison with a BSA standard curve, the concentration of the protein solution was determined.

3.2.2.5 SDS-polyacrylamide gel electrophoresis

The denaturing SDS polyacrylamide gel electrophoresis (SDS-PAGE) was performed according to Laemmli method (Laemmli, 1970). In this study, acrylamide gels of 12 % and 15 % (37.5:1 acrylamide:bis-acrylamide, 1mm thickness) were used depending on the protein-mixture that had to be separated. Before sample loading on the gel, proteins were mixed with Laemmli buffer and incubated 5 min at 95 °C to ensure complete denaturation. After loading the samples on the gel in a gel chamber filled with protein running buffer, the proteins were focused in the stacking gel at 15-25 mA and subsequently separated in the resolving gel at 30-45 mA.

<u>Laemmli buffer</u>		<u>Protein running buffer</u>		<u>5 % of stacking gel (10 ml)</u>	
75 mM	Tris, pH 6.8	25 mM	Tris, pH 8.8	H ₂ O	6.9 ml
1.25 mM	EDTA	192 mM	Glycine	1M Tris, pH 6.8	1.25 ml
2.5 % (w/v)	SDS	0.1 % (w/v)	SDS	30 % Acrylamide	1.67 ml
20 % (w/v)	Glycerol			10 % SDS	0.1 ml
0.1 % (w/v)	Bromphenolbue			10 % APS	0.1 ml
50 mM	DTT			TEMED	0.01 ml
<u>12 % of resolving gel (30 ml)</u>		<u>15 % of resolving gel (30 ml)</u>			
H ₂ O	6.3 ml	H ₂ O	3.3 ml		
1M Tris, pH 7.8	11.25 ml	1M Tris, pH 7.8	11.25 ml		
30 % Acrylamide	12 ml	30 % Acrylamide	15 ml		
10 % SDS	0.3 ml	10 % SDS	0.3 ml		
10 % APS	0.15 ml	10 % APS	0.15 ml		
TEMED	0.03 ml	TEMED	0.03 ml		

3.2.2.6 Gel staining

Proteins on SDS-PAGE gels were visualized either by staining with Coomassie brilliant blue R250 and destaining (Sambrook and Fritsch, 1989) or by silver-staining (Blum and Beier, 1987). Whereas Coomassie-staining reveals a band formed by up to 1 µg of protein and silver-staining can detect up to 5 ng of protein in a single band.

3.2.3 Protein crystallography

In this section applied methods for protein crystallization, data collection and processing, phasing, model building and refinement, and structure analysis are described. Basic principles of protein X-ray crystallography are provided in the appendixes (Section 7.1).

3.2.3.1 Pre-crystallization test

To determine the optimal concentration for crystallization, a pre-crystallization test was carried out by using the pre-crystallization kit (Hampton) according to manufacturer's instruction. Alternatively, the protein was crystallized in the Hampton classics screen and the number of drops, where the protein was precipitated, was counted. The target concentration was determined when approximately 1/3rd of all conditions showed precipitation 1 h later after drop setting.

3.2.3.2 Protein crystallization

Initial screening was performed in a 96-well format crystallization plate. The protein was spun down prior to crystallization at 13 krpm for 5 min. Drop volumes of 200 nl with a 1:1 ratio of protein and reservoir solution were set up by the Cartesian NanoDrop robot at 20 °C *via* sitting drop vapor diffusion method. The screens listed in the Section 3.1.14 were usually tested in initial screening. An overview of the crystallization experiments performed with different protein complexes is provided in the Table 3.5. The initial conditions that yielded crystals were subsequently scaled up to microliter range and optimized by screening the effects of precipitant and pH. Droplets were set up manually at 20 °C *via* sitting drop vapor diffusion by mixing 1 µl of sample with 1 µl of reservoir solution in a 24-well format crystallization plate. Crystals could be cooled at cryogenic temperatures after transfer into certain cryo-protectants (Table 3.5). Cryo-protectants were determined by checking a titration curve of cryo-protectant mixed with reservoir for the scattering behavior upon exposure to an X-ray beam.

Table 3.5 High-throughput crystallization experiments

Protein name	Concentration (mg/ml)	Number of screened conditions	Optimized buffer condition	Cryo-protectant
NusB-S10 ^{Δloop}	16	960	0.1 M CHES, pH 8.8, 18 % PEG 8000	10 % propylene glycol
NusB ^{Asp118Asn} -S10 ^{Δloop}	16	672	0.2 M K ₃ C ₆ H ₅ O ₇ , 20 % PEG 3350	40 % glycerol
NusB-S10 ^{Δloop, Ala86Asp}	16	768	0.1 M Tris, pH 7.1, 0.2 M (NH ₄)SO ₄ , 25 % PEG 3350	10 % propylene glycol
NusB-S10 ^{Δloop, Pro39Ala}	16	672	0.17 M Mg(CH ₃ COO) ₂ , 14 % PEG 3350	30 % glycerol
S10 ^{Δloop}	16	1248	0.16 M (NH ₄)SO ₄ , 0.08 M NaCH ₃ COO, pH 4.4, 22 % PEG 4000, 20 % glycerol	None

3.2.3.3 Data collection and processing

The diffraction quality of crystals was tested on an in-house source equipped with a MAR image plate detector coupled to a RU-200 rotating anode X-ray generator producing CuK α radiation with a wavelength of 1.5148 Å. The complete diffraction dataset for NusB-S10 ^{Δ loop, Ala86Asp} protein complex was collected on beamline PXI in SLS using a Pilatus detector (Broennimann et al., 2006). The complete diffraction datasets for the rest protein complexes were collected on beamline PXII in SLS using a MarCCD 225 mm detector. Data collection strategies are shown in the Table 3.6. The data were processed with the XDS package (Kabsch, 1993). Crystallographic data can be found in the Table 4.2.

Table 3.6 Data collection strategy

Protein	Data collection strategy				
	Distance (mm)	λ (Å)	$\Delta \phi$ (°)	Exposure time (s)	Frames
NusB-S10 ^{Δloop}	100	0.9840	0.5	0.5	720
NusB ^{Asp118Asn} -S10 ^{Δloop}	250	0.9788	1	1	180
NusB-S10 ^{Δloop, Ala86Asp}	400	0.9763	0.2	0.2	500
NusB-S10 ^{Δloop, Pro39Ala}	180	0.9999	0.5	1	190
S10 ^{Δloop}	230	0.9200	0.5	1	360

3.2.3.4 Phasing, model building and refinement

The usage of the *NcoI* site for NusB gene cloning into pETM-11 vector gave rise to a Lys2Glu point mutation. The mutated protein was initially used for crystallographic analysis. The phase of NusB^{Lys2Glu}-S10 ^{Δ loop} was calculated by molecular replacement through Molrep (CCP4, 1994) using the coordinates of *Thermotoga maritima* NusB (PDB ID 1TZV; (Bonin et al., 2004b)) and of *Thermus thermophilus* S10 taken from the structure of the *T. thermophilus* 30S ribosomal subunit (chain J of PDB ID 1J5E; (Wimberly et al., 2000)). The coordinates of NusB^{Lys2Glu}-S10 ^{Δ loop} were employed to solve the other protein complexes' structures by molecular replacement. The models were built using COOT (Emsley and Cowtan, 2004) and refined by standard methodology using Refmac5 (Murshudov et al., 1997) including TLS refinement (Winn et al., 2001).

3.2.3.5 Structure analysis

The geometric quality assessment on the refined models was done with PROCHECK (Laskowski et al., 1993). Illustrations of the structures were prepared by Pymol (<http://pymol.sourceforge.net/>).

3.2.4 Biochemical assays

3.2.4.1 GST pull-down assay

Frozen cells were lysed as above (Section 3.2.2.2). The cleared lysates were incubated with glutathione-sepharose beads equilibrated with binding buffer. Trapped proteins were washed with binding buffer and eluted with 15 mM reduced glutathione. Aliquots of the samples were analyzed by SDS-PAGE.

3.2.4.2 5'-End labeling of RNA-oligonucleotides

19mer BoxA RNA oligonucleotides containing *rrn* BoxA or Nut BoxA were 5'-end-labeled with [γ - 32 P]-ATP (6000 Ci/mmol) using T4 polynucleotide kinase (PNK). 10 pmol of RNA-oligonucleotide was incubated in a volume of 10 μ l with 1 μ l of T4 PNK and 1 μ l of 10 \times PNK buffer in the presence of 2 μ l of [γ - 32 P]-ATP. The mixture was incubated at 37 $^{\circ}$ C for 30 min. The 5'-end labeled product was diluted with 40 μ l of H₂O and purified *via* MicroSpin G25 columns according to the manufacturer's protocol. The labeled product was stored at -20 $^{\circ}$ C. A typical [γ - 32 P]-ATP labeled reaction is shown below:

[γ - 32 P]-ATP labeled reaction mixture (10 μ l)

1 μ l	RNA oligo (10 pmol/ μ l)
2 μ l	[γ - 32 P]-ATP
1 μ l	10 \times PNK buffer
1 μ l	T4 PNK enzyme
5 μ l	H ₂ O

3.2.4.3 Double filter-binding assay

[γ - 32 P]-ATP labeled RNA oligonucleotide was diluted by a factor of 50. Varying concentrations of protein complex (0, 0.04, 0.08, 0.1, 0.2, 0.4, 0.6, 0.8, 1.0, 2.0, 4.0, 8.0 μ M) were incubated with 1 μ l of diluted [32 P]-labeled RNA oligonucleotide for 30 min at 4 $^{\circ}$ C in 10 μ l reaction volumes. The upper nitrocellulose membrane and the lower nylon membrane served to trap protein-RNA complexes and unbound RNA, respectively. The membranes were pre-washed with MilliQ water and soaked for one hour in crystallization buffer at 4 $^{\circ}$ C (Wong and Lohman, 1993). A multi-well filtration manifold was used to spot samples onto the membranes according to the manufacturer's instruction. After membranes had been washed with 200 μ l of crystallization buffer and air-dried, radioactivity retained on the membranes was visualized by a Typhoon 8600 phosphoimager.

3.2.4.4 UV-induced crosslinking assay

[γ - 32 P]-ATP labeled RNA oligonucleotide was diluted by a factor of 50. Varying concentrations (0, 0.15, 0.31, 0.62, 1.25 and 2.5 μ M) of NusB-S10 $^{\Delta$ loop or NusB101-S10 $^{\Delta$ loop (NusB^{Asp118Asn}-S10 $^{\Delta$ loop) were mixed with 1 μ l of diluted [32 P]-labeled RNA oligonucleotide in 10 μ l reaction volumes and exposed to 254 nm ultraviolet light for 5 min at 4 $^{\circ}$ C (Lingel et al., 2003). Reactions were analyzed by 15 % SDS-PAGE. Gels were dried and developed on a phosphoimager.

Under saturating conditions, a maximum of ca. 7 % of the total radioactivity was shifted on gels. For quantification, 0.31 and 0.62 μ M of NusB-S10 $^{\Delta$ loop or NusB101-S10 $^{\Delta$ loop were crosslinked as above. Crosslinked samples from three independent experiments were analyzed on the same SDS-PAGE gel. For loading control, each sample was divided and averaged. Radiolabeled bands were quantified by densitometry using Image Quant software (GE Healthcare). Crosslink yields for the components of the wt NusB-based complex were normalized to 1 and the yields for the corresponding components of the NusB^{Asp118Asn}-based complex were represented relative to the wt sample.

3.2.4.5 Deduction of protein-RNA crosslinking sites

Crosslinks identified are listed in the Table 4.6. NusB peptide B1 (96-SDVPYKVAINEAIELAK-112) was found crosslinked to a CU (or UC) dinucleotide. The only such sequences are found at positions 3-6 of λ or *rrn* BoxA elements. Thus, peptide B1 must be in contact with this region in either RNA. Consistently, peptide B1' that is elongated by an arginine at the C-terminus compared to peptide B1 shows identical crosslinking behavior as B1.

NusB peptide B3 (122-FVNGVLDK-129) was found crosslinked to a UU dinucleotide employing either λ or *rrn* BoxA RNA. There are two regions encompassing UU dinucleotides. Therefore, the peptide B3 is in close proximity to the UUU sequence at positions 6-8 of *rrn* BoxA and to the UU element at positions 6/7 of λ BoxA. This conclusion is based on the observation that the same peptide in isolated NusB was found crosslinked to a triple-U sequence of the *rrn* BoxA oligomer. The latter crosslink rules out the UU di-nucleotide 3'-terminal of the core BoxA as a crosslinking site for peptide B3.

NusB peptide B2 (113-SFGAEDSHK FVNGVLDK-129) encompasses the linker between peptides B1 and B3 plus the entire B3 peptide. In complex with S10 and λ BoxA, peptide B2,

but not the shorter peptide B3, crosslinks to a UAC (or permuted) trinucleotide. There is only one such sequence at positions 7-9 of the λ BoxA RNA.

Peptide E1 of S10 (10-LKAFDHR-16) was found crosslinked to a UA (or AU) element. The observation that this peptide also crosslinks to a UA/AU in a shortened λ BoxA core oligomer, which harbors only one such dinucleotide, unequivocally identifies positions 8/9 of *rrn* BoxA and positions 7/8 of λ BoxA as the contact sites of this peptide and rules out contacts to UA/AU-elements 3' of the BoxA cores. Peptides E2 (49-FTVLISPHVNK-58) and E3 (63-DQYEIR-68) are entirely and partially contained in the ribosome-binding loop of S10, respectively. They both crosslink to an AAU (or permuted) trinucleotide. For peptide E3, this crosslink was observed with either type of RNA, ruling out the possibility that peptide E3 crosslinked to positions 8-10 of *rrn* BoxA. Instead, this peptide must be in close proximity to the AAU elements at position 12 and beyond. Peptide E2 also crosslinks to this latter region in the RNAs, since in the case of *rrn* BoxA an additional unequivocal crosslink to an AAUU (or permuted) oligo was observed.

3.2.4.6 Ribosome preparation

Ribosomes were prepared by sedimentation from whole-cell extracts as described previously (Worbs et al., 2002). Cells were transformed with the plasmids and grown at 32 °C in 200 ml LB medium with 100 μ g/ml ampicillin. At $OD_{450} = 0.1$, IPTG was added to 0.5 mM final concentration. At $OD_{450} = 1.5$, cells were harvested by centrifugation. The cell pellets were washed once with 1 ml of buffer A (20 mM HEPES-KOH, pH 7.5, 6 mM $MgCl_2$, 30 mM NH_4Cl , 6 mM β -mercaptoethanol), resuspended in 2 ml of buffer A and split into two 1-ml aliquots.

After addition of lysozyme (100 μ l of a 15 mg/ml solution) and incubation on ice for 3 min, cell lysis was completed by freeze-thawing. Lysates were clarified by spinning at 23,000 rpm for 30 min in an S100-AT4 rotor. The duplicate supernatants from each culture were pooled and centrifuged for 4 hours at 43,000 rpm in an S100-AT4 rotor. The pellets were resuspended overnight in 200 μ l of buffer A and centrifuged at 7900 x g in a Fresco 17 centrifuge for 10 min at 4 °C. The resulting supernatants contain 'crude' ribosomes. 150 μ l of the crude ribosome preparations were mixed with 1.8 ml of buffer B (20 mM HEPES-KOH, pH 7.5, 30 mM $MgCl_2$, 1 M NH_4Cl , 6 mM β -mercaptoethanol), incubated for 1 hour at 4 °C and then centrifuged for 4 hours at 53,000 rpm in an S100-AT4 rotor. The pellets were rinsed once with 200 μ l of buffer A, resuspended overnight in 100 μ l of buffer A and centrifuged at

7900 x g in a Fresco 17 centrifuge for 10 min at 4 °C. The resulting supernatants contain ‘salt-washed’ ribosomes.

For analysis of total-cell extracts, 1-ml aliquots of the 200-ml cultures were removed prior to IPTG addition and grown in parallel (uninduced extract, ‘ext -’). Also, 1-ml aliquots were removed from the IPTG-induced culture immediately before harvesting (induced extract, ‘ext +’). Both sets of samples were centrifuged and the pellets were resuspended in 200 µl of Laemmli sample buffer, incubated for 2 min at 95 °C and stored at -20 °C.

3.2.4.7 Western blot

Proteins from ribosomes (0.1 A₂₆₀ equivalents) were separated by 12 % SDS PAGE and electro-blotted on a nitrocellulose transfer membrane by a Trans-Blot Electrophoretic Transfer Cell according to the manufacturer’s instruction. A pre-stained MW standard was applied onto one lane as an indication of a successful transfer. The gel, membrane and filter paper were soaked in the Blot buffer and sandwiched by a cassette. The transfer was performed in the Blot buffer for 2 h at 70 V at 4 °C. To decrease non-specific binding of antibodies, the membrane was blocked in the block buffer for overnight at 4 °C. For probing GST-tagged S10 or S10^{Δloop}, the membrane was first incubated with a rabbit anti-GST antibody (primary Ab solution) for 1 h at room temperature, washed 3 times for 15 min/ each with washing-1 solution at room temperature and subsequently by a goat anti-rabbit IgG (secondary Ab solution) for 1 h at room temperature, and then washed 3 times for 15 min/ each with washing-2 solution at room temperature. The signal was detected by using ECL Western blotting detection reagents and exposing a high performance chemiluminescence film. The film was developed on a KONICA developer.

<u>Slab 4 (5 L)</u>		<u>10× TBS (1 L)</u>		<u>Blot buffer (3 L)</u>		<u>Block buffer</u>
30 g	Tris	24.2 g	Tris	1.5 L	Slab 4	1× TBS
142.6 g	Glycine	87.66 g	NaCl	0.6 L	Methanol	1 % Tween
5 g	SDS	Add H ₂ O up to 1 L		0.9 L	H ₂ O	5 % Milk
Add H ₂ O up to 5 L		Adjust to pH 7.6 with HCl				
<u>Primary Ab solution</u>		<u>Washing-1</u>		<u>Secondary Ab solution</u>		<u>Washing-2</u>
1× TBS		1× TBS		1× TBS		1× TBS
1 % Tween		1 % Tween		1 % Tween		1 % Tween
1 % Milk		1 % Milk		1 % Milk		
Rabbit anti-GST antibody				Goat anti-rabbit IgG		

3.2.4.8 Analytical size exclusion chromatography

NusB-S10^{Δloop} and NusG proteins were mixed in approximately equimolar ratios, applied on a Superdex-75 PC 3.2 column (GE Healthcare), and chromatographed in 10 mM Tris-HCl, pH 7.5, 50 mM NaCl and 2 mM DTT using a SMART protein purification system (GE Healthcare). For a typical run, 30 μl of sample were loaded on the column at a flow rate of 40 μl/min. 40-μl fractions were collected and analyzed by 15 % SDS-PAGE gel.

4 Results

4.1 Transcriptional and translational functions are attributed to distinct regions of S10

4.1.1 The long ribosome-binding loop of S10 is dispensable for transcriptional functions

To investigate the structural requirements of S10 (NusE) as a transcription factor, I attempted to delineate molecular regions that are dispensable for processive transcription antitermination. In the 30S ribosomal subunit, S10 exhibits a globular domain that is located at the surface of the particle and an extended ribosome-binding loop that deeply penetrates the subunit and interacts with several other r-proteins and the 16S rRNA (Figure 2.5A and 2.5B; (Schluzen et al., 2000; Wimberly et al., 2000)). I speculated that the ribosome-binding loop may be dispensable for transcription antitermination. To test this idea, a truncated S10 variant was generated, in which this loop (residues 46-67) was replaced by a serine (S10^{Δloop}). To test whether the truncation affected the interaction with NusB, full-length S10 or S10^{Δloop} were co-expressed with NusB in *E. coli* and purified *via* a GST-tag on the S10 molecules. Both wild type (wt) and truncated S10 remained stably associated with NusB during purification (Figure 4.1A, lanes 1-6). During antitermination, the NusB-S10 complex interacts with the BoxA element of the mRNA, a function that should be preserved in the NusB-S10^{Δloop} complex. Indeed, the affinities of the full-length and loop-deleted complexes for BoxA-containing RNAs were comparable in a filter-binding assay (Figure 4.1B). Most importantly, the antitermination activity of the loop-deleted S10 variant was tested directly by Max Gottesman's group. They found that S10^{Δloop} complemented λ growth at 42 °C in an *E. coli* strain bearing a chromosomal *nusE71* defect (Table 4.1) that normally blocks N-antitermination and λ growth at high temperatures (Friedman et al., 1981). Therefore, the transcription antitermination activity of S10 is unaffected by deletion of its ribosome-binding loop.

4.1.2 The loop-deleted S10 variant does not bind to ribosomes

It was known that the *nusE* gene is essential for cell growth (Bubunenko et al., 2007). The question is whether *nusE*^{Δloop} gene is also essential for cell growth. To answer this question,

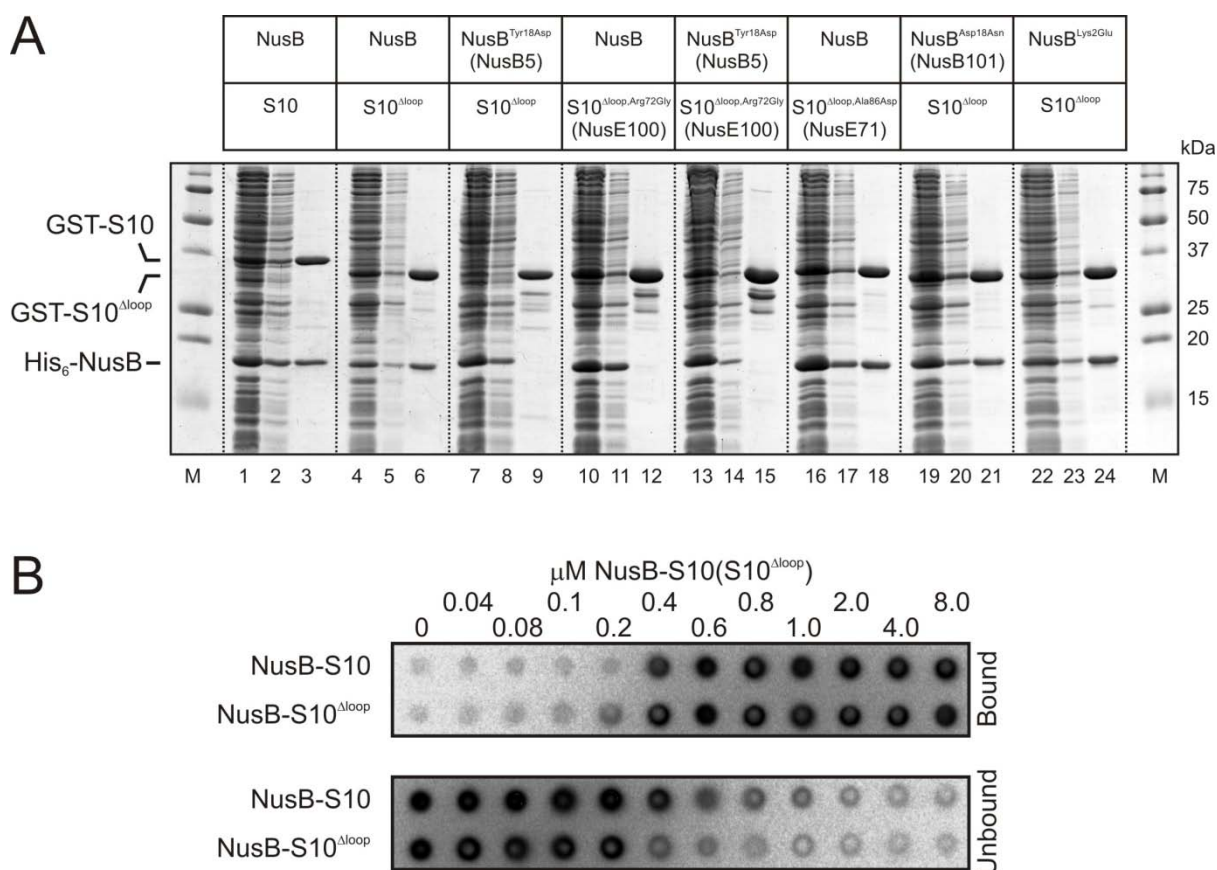


Figure 4.1 Analysis of the S10^{Δloop} mutant

(A) Copurification of GST-S10 or GST-S10^{Δloop} and mutants with His₆-NusB and mutants. Groups of three lanes show the soluble extract from co-overexpression experiments (first lane), the wash (second lane) and the elution (third lane) from glutathione beads. Co-expressed proteins are indicated above the group of lanes. M, molecular mass marker; sizes of marker bands in kDa are indicated on the right.

(B) Double filter-binding assays of a NusB-S10 complex to an *rrn* BoxA-containing 19mer RNA. (Upper panel) nitrocellulose layer representing bound RNA. (Lower panel) nylon filter representing unbound RNA. The upper lanes correspond to the full-length complex, the lower lanes to the NusB-S10^{Δloop} complex. Numbers indicate protein concentrations in μM.

Table 4.1 *nusE*⁺ and *nusE*^{Δloop} are dominant to *nusE71*

Chromosomal <i>nusE</i>	pBAD Plasmid	Arabinose	λ EOP
+	-	-	1
71	-	-	<10 ⁻⁵
71	<i>nusB</i> ⁺	-	<10 ⁻³
71	<i>nusE</i> ⁺	-	1.0
71	<i>nusE</i> ^{Δloop}	-	1.0
71	<i>nusB</i> ⁺	+	<10 ⁻³
71	<i>nusE</i> ⁺	+	1.0
71	<i>nusE</i> ^{Δloop}	+	1.0

nusE71 is non-permissive for λ growth at 42 °C. Strains are W3102 derivatives that carry *nusE*⁺ or the *nusE71* mutation in the chromosome and the indicated plasmid. *λimm434* was titered on LB or LB plus ampicillin (50 μg/ml) at 42 °C and Efficiencies of Plating (EOP) were determined. Where indicated, 0.1 % arabinose was added to the plate.

Donald Court's group used recombineering technique to test if *nusE*^{Δloop} gene is able to suppress deletion of the chromosomal *nusE* gene to support cell growth. They found that in cells containing a plasmid with *nusE* under arabinose control, the chromosomal *nusE* could be replaced with a *kan* gene (*kan* open reading frame fused with *nusE* open reading frame), conferring kanamycin resistance, in an arabinose-dependent manner (Figure 4.2A). The appearance of a single *nusE*<>*kan* fragment indicates that S10 expressed from a plasmid is functional and is able to complement the nonviable chromosomal *nusE* knockout (Bubunenko et al., 2007). In contrast, cells containing plasmid-borne *nusE*^{Δloop} yielded only rare *nusE*<>*kan* recombinants irrespective of arabinose induction. The 40 such recombinants tested all carried an additional *nusE*⁺ gene as a tandem duplicate in the chromosome. The appearance of two fragments representing *nusE*<>*kan* and *nusE* indicates that S10^{Δloop} copy expressed from a plasmid is not functional and is unable to complement a chromosomal knockout. In this case, recombinants are rare and have a knockout copy and a wt copy of *nusE*, which reflects the special diploid nature of these strains (Bubunenko et al., 2007). Thus, *nusE*^{Δloop} does not encode all vital functions of *nusE*.

I speculated that the long ribosome-binding loop of S10 is essential for cell growth most likely due to its interaction role in ribosomes, in which S10^{Δloop} may fail to bind to ribosomes and therefore fail to support translation. In order to test this idea, I directly monitored binding of glutathione S-transferase (GST)-S10 and GST-S10^{Δloop} to ribosomes. GST-S10 and GST-S10^{Δloop} were overexpressed in an *E. coli* BL21(DE3) strain, and then the fully assembled ribosomes were prepared. Crude and salt washed ribosomes were purified by ultracentrifugation of a whole cell lysate. The procedure of salt washing ribosomes in the crude ribosome pellet is used to distinguish between true ribosomal proteins and proteins associated with ribosome as contaminants or accessory translational factors (Zengel et al., 2003). Associations of GST-S10 and GST-S10^{Δloop} with the crude and salt washed ribosomes were evaluated by Western blot by using an anti-GST antibody. Since the N-terminus of S10 is accessible on the surface of the 30S ribosomal subunit (Wimberly et al., 2000), an N-terminal GST fusion should not interfere with stable ribosome incorporation of the protein. Indeed, GST-S10 was incorporated readily and in a salt-stable manner into ribosomes (Figure

4.2B, lanes 1-4). In contrast, while residual amounts of GST-S10^{Δloop} were seen associated with crude ribosome preparations, the truncated fusion protein was completely lacking from salt-washed ribosomes (Figure 4.2B, lane 5-8). Thus, S10 behaves differently from some other r-proteins, such as L4, in which analogous ribosome-penetrating loops are not required for stable ribosome association (Zengel et al., 2003).

These data show that transcriptional and translational functions can be attributed to distinct regions of S10. Namely, only is the globular part of S10 necessary for transcription function, while its loop is exclusively required for S10 function in translation.

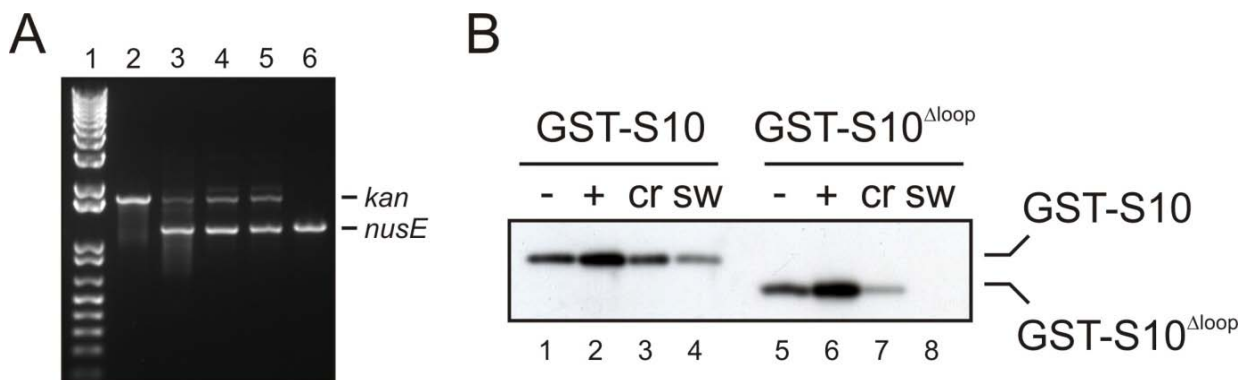


Figure 4.2 Gel analysis of *nusE*<->*kan* recombinants and ribosome binding of S10^{Δloop}

(A) Gel analysis of *nusE*<->*kan* recombinants. Kanamycin resistant cells from a single colony were analyzed by PCR for configuration of the targeted chromosomal *nusE* region. Lane 1, DNA markers (Invitrogen). Lanes 2 and 3, PCR products from recombinant cells that contained pBAD*nusE*. Lanes 4 and 5, PCR products from recombinant cells that contained pBAD*nusE*^{Δloop} initially selected either with (lanes 2 and 4) or without (lanes 3 and 5) 0.2 % arabinose. Lane 6, PCR product control of wt *nusE* from the bacterial chromosome. Note that a haploid *nusE*<->*kan* knockout can be made only when pBAD*nusE* is induced by arabinose, *i.e.* when wt *nusE* is expressed from the plasmid (lane 2).

(B) Western blot probing the binding of GST-S10 and GST-S10^{Δloop} to ribosomes. Equal amounts of cells before (-; lanes 1 and 5) and after (+; lanes 2 and 6) induction with IPTG as well as equal amounts (0.1 A₂₆₀ equivalents) of crude (cr; lanes 3 and 7) and salt-washed (sw; lanes 4 and 8) ribosomes from the *E. coli* BL21(DE3) strains expressing GST-S10 (lanes 1-4) or GST-S10^{Δloop} (lanes 5-8) were analyzed on a 12 % SDS gel, transferred to a nitrocellulose membrane and analyzed by Western blotting.

4.2 Structural analysis of the NusB-S10 complex

4.2.1 Crystal structure of a transcriptionally active NusB-S10 Complex

The gene encoding NusB was PCR-amplified from *E. coli* chromosomal DNA and ligated into the *NcoI* and *Acc65I* restriction enzyme sites of the pETM11 vector. The usage of the *NcoI* site for cloning into pETM11 gave rise to a Lys2Glu point mutation. The mutated protein was used for crystallographic analysis (the NusB^{Lys2Glu}-S10^{Δloop} complex is referred as wt NusB-S10^{Δloop} complex). For functional studies, Glu2 was converted back to a lysine by site directed mutagenesis. NusB proteins with Glu or Lys at position 2 behaved identically in biochemical, RNA crosslinking and *in vivo* studies (see for example GST pull-down assays in Figure 4.1A, lanes 22-24).

I exploited the results from the functional dissection of S10 in order to devise a high-resolution crystal structure of a transcriptionally active NusB-S10 complex. Crystals obtained from the complex of the full-length proteins did not diffract well. The ribosome-binding loop of S10 might be flexible off the ribosome and disturb the crystalline order. Therefore, the S10^{Δloop} was employed, instead of full-length S10, to co-express and co-purify with NusB. The NusB-S10^{Δloop} complex gave rise to crystals that diffracted to 1.3 Å resolution and allowed structure solution by molecular replacement. The structure was refined to R_{work} and R_{free} factors of 17.3 and 20.4 %, respectively (Table 4.2).

In the structure of the complex (Figure 4.3A), NusB adopts an all-helical fold with two perpendicular three-helix bundles. S10^{Δloop} exhibits a four-stranded antiparallel β -sheet backed by two α -helices on one side. Helix α 1 and an irregular strand, β 2, of S10^{Δloop} bridge the two helical bundles of NusB (contact regions I and II in Figure 4.3A). The region on NusB contacted by S10^{Δloop} coincides with NusB residues that show NMR chemical shift changes upon addition of full-length S10 (Das et al., 2008). These observations further corroborate the equivalence of the wt and loop-deleted S10 in transcription.

NusB and S10^{Δloop} approach each other *via* complementary electrostatic surfaces (Figure 4.3B), burying ca. 1700 Å² of combined surface area upon complex formation. The two proteins engage in mixed hydrophobic and hydrophilic interactions (Figure 4.3C–4.3F). For example, an intramolecular Asp19-Arg72 ion pair of S10^{Δloop} forms hydrogen bonds to Tyr18 of NusB, thereby positioning Tyr18 between Pro39 and Pro41 of a proline motif (Pro39-Ile40-Pro41-Leu42-Pro43) on strand β 2 of S10 (Figure 4.3C). The remainder of the proline

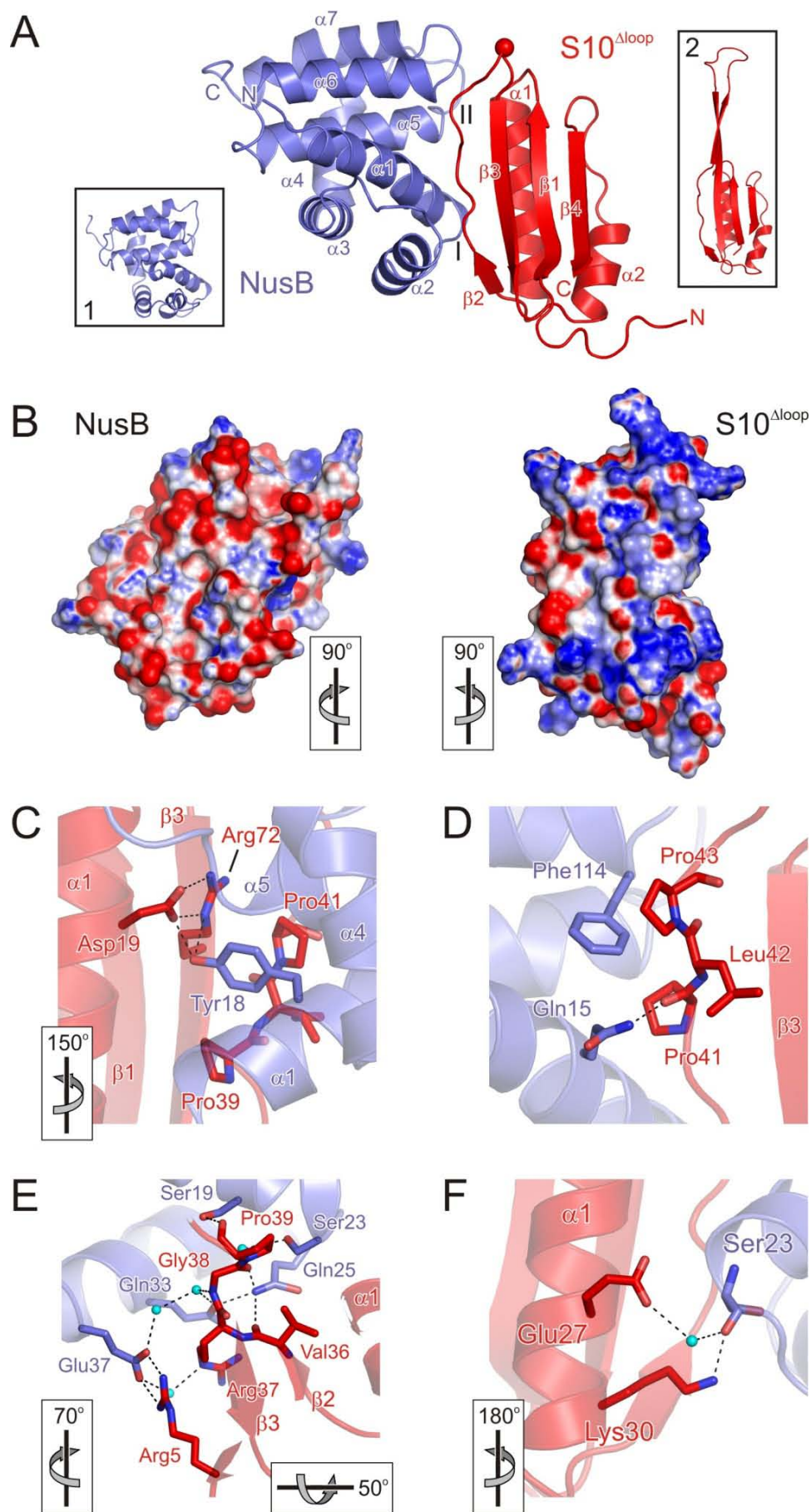


Figure 4.3 Structure of the NusB-S10^{Δloop} complex

(A) Ribbon plot of the *E. coli* NusB-S10^{Δloop} complex. NusB, blue, S10^{Δloop}, red. Secondary structure elements and termini are labeled. The red sphere marks the site at which the ribosome-binding loop of S10 has been replaced by a single serine. (I and II) Interaction regions on the flank of the first three helix bundle (I) and on a tip of the second three helix bundle (II) of NusB. (Inset 1) NMR structure of *eco*NusB (PDB ID 1EY1; (Altieri et al., 2000)) after global superpositioning on the NusB molecule of the present complex. (Inset 2) Structure of S10 from the *E. coli* 30S subunit (PDB ID 2AVY; (Schuwirth et al., 2005)) after global superpositioning on the S10^{Δloop} molecule of the present complex.

(B) Electrostatic surface potentials mapped on the surfaces of NusB (left) and S10^{Δloop} (right) showing a view on the interfaces of both molecules. Blue, positive charge, red, negative charge. Protomers were rotated 90° relative to panel (A) as indicated.

(C-F) Details of the NusB-S10^{Δloop} interaction. Interacting residues and secondary structure elements are labeled. Residues of interest are colored by atom type: carbon, the respective molecules; oxygen, red; nitrogen, blue. Cyan spheres indicate water molecules. Dashed lines are hydrogen bonds or salt bridges. Views relative to (A) are indicated.

motif with two intervening apolar side chains engages in snug van-der-Waals contacts with NusB-Phe114, sandwiching it between S10-Pro41 and S10-Pro43 (Figure 4.3D). Pro39 is molded into a *cis* conformation that allows it to participate in intra- and intermolecular hydrogen bonding networks (Figure 4.3E).

4.2.2 NusB and S10 retain their overall folds upon complex formation but interact via local induced fit

The global structures of isolated NusB (Altieri et al., 2000; Bonin et al., 2004b; Das et al., 2008; Gopal et al., 2000) and of NusB in complex with S10^{Δloop} are very similar (Figure 4.3A, inset 1; Table 4.3). S10^{Δloop} in complex with NusB likewise resembles the structure of S10 in the 30S subunit (Schuwirth et al., 2005) (Figure 4.3A, inset 2; Table 4.3). It was suggested that the fold of S10 by itself is unstable (Das et al., 2008; Gopal et al., 2001); thus, NusB apparently acts to stabilize S10 in the same overall conformation it takes in the ribosome. Clearly, the data exclude the possibility that S10 is extensively remodeled by NusB as a mechanism for partitioning of S10 between the translation and transcription machineries as suggested by Gopal et al (Gopal et al., 2001).

While the global structures of both proteins are conserved, they are apparently adjusted by local induced fit upon complex formation. A pronounced difference to the structure of NusB determined in isolation (Altieri et al., 2000; Bonin et al., 2004b; Das et al., 2008) is seen in

Table 4.2 Crystallographic data

Data collection	NusB-S10 ^{Δloop}	NusB ^{Asp118Asn} -S10 ^{Δloop}	NusB-S10 ^{Δloop, Ala86Asp}	NusB-S10 ^{Δloop, Pro39Ala}	S10 ^{Δloop}
Wavelength (Å)	0.9051	0.9788	0.9763	0.9999	0.9200
Temperature (K)	100	100	100	100	100
Space Group	P2 ₁ 2 ₁ 2 ₁	I4 ₁ 22	I4 ₁ 22	C222 ₁	P6 ₁
Unit Cell Parameters (Å, °)	a = 40.7, b = 49.0, c = 122.8	a = 112.64, b = 112.64, c = 263.25	a = 112.2, b = 112.2, c = 266.2	a = 38.57, b = 76.42, c = 151.76	a = 63.22, b = 63.22, c = 61.24
Resolution (Å)	30.0 - 1.3 (1.4 - 1.3) ^a	30.0 - 2.5 (2.6 - 2.5) ^a	50.0 - 2.6 (2.7 - 2.6) ^a	30.0 - 1.7 (1.8 - 1.7) ^a	50.0 - 1.9 (2.0 - 1.9) ^a
Reflections					
Unique	56411 (11095)	29761 (3263)	26358 (2780)	23587 (2808)	16062 (952)
Completeness (%)	100 (100)	100 (100)	100 (100)	93.6 (72.2)	99.5 (92.5)
Redundancy	15.3 (14.6)	7.22 (7.42)	7.1 (6.5)	3.7 (3.3)	4.1 (2.4)
I/σ(I)	17.5 (5.3)	18.1 (4.1)	19.8 (3.9)	18.2 (5.4)	16.9 (3.9)
R _{sym} (I) ^b	7.4 (64.8)	8.6 (72.5)	6.8 (46.2)	6.6 (35.0)	7.1 (21.0)
Refinement					
Resolution (Å)	20.0 - 1.3 (1.33 - 1.30)	30.0 - 2.5 (2.56 - 2.50)	30.0 - 2.6 (2.67 - 2.60)	30.0 - 1.7 (1.74 - 1.70)	15.0 - 1.9 (1.95 - 1.90)
Reflections					
Number	56394 (4109)	29753 (2178)	26337 (1921)	23522 (1121)	10944 (745)
Completeness (%)	100 (100)	100 (100)	100 (100)	100 (100)	99.4 (93.7)
Test Set (%)	5.0	5.0	5.0	5.0	4.8
R _{work} ^c	17.3 (27.8)	20.4 (23.1)	21.8 (29.5)	20.1 (24.0)	25.3 (24.3)
R _{free} ^c	20.4 (29.7)	25.6 (29.1)	28.0 (35.8)	23.4 (25.7)	29.3 (26.7)
Contents of A.U.^d					
Protein Molecules, Refined Atoms	1/1, 1779	3/3, 5313	3/3, 5308	1/1, 1766	1, 636
Water	316	155	197	185	62
Ligand /Atoms	3 CHES/39	1 K ⁺ /1	-	-	-
Mean B-Factors (Å²)					
Wilson	24.6	52.7	54.3	25.6	29.9
Protein	21.5	60.5	48.7	17.9	31.3
Water	41.0	22.8	46.0	25.7	45.4
Ligand	38.2	29.5	-	-	-
Ramachandran Plot^e					
Favored	99.1	97.12	96.8	97.7	100
Allowed	0.9	2.43	2.3	1.8	0
Outliers	0	0.45	0.9	0.5	0
RMSD^f from Target Geometry					
Bond Lengths (Å)	0.013	0.01	0.006	0.01	0.01
Bond Angles (°)	1.51	1.22	1.09	1.38	1.45
RMSD B-Factors (Å²)					
Main Chain Bonds	0.86	0.41	0.28	0.76	0.87
Main Chain Angles	1.31	0.82	0.51	1.44	1.68
Side Chain Bonds	2.47	1.43	0.45	2.49	2.47
Side Chain Angles	3.58	2.48	0.80	4.24	4.36
PDB ID	3D3B	3IMQ	3D3C	-	-

^a Data for the highest resolution shell in parentheses

^b $R_{\text{sym}}(I) = \frac{\sum_{\text{hkl}} \sum_i |I_i(\text{hkl}) - \langle I(\text{hkl}) \rangle|}{\sum_{\text{hkl}} \sum_i I_i(\text{hkl})}$; for n independent reflections and i observations of a given reflection; $\langle I(\text{hkl}) \rangle$ – average intensity of the i observations

^c $R = \frac{\sum_{\text{hkl}} ||F_{\text{obs}}| - |F_{\text{calc}}||}{\sum_{\text{hkl}} |F_{\text{obs}}|}$; $R_{\text{work}} - \text{hkl} \notin T$; $R_{\text{free}} - \text{hkl} \in T$; T – test set

^d A.U. – asymmetric unit

^e Calculated with MolProbity (<http://molprobity.biochem.duke.edu/>) (Davis et al., 2004)

^f RMSD – root-mean-square deviation

the loop connecting helices $\alpha 4$ and $\alpha 5$, which rearranges to allow an ionic interaction between NusB-Glu75 and S10-Arg16 (Figure 4.4A). In agreement with this observation, strong NMR chemical shift changes were previously observed in this loop of NusB upon addition of full-length S10 (Das et al., 2008). In ribosome-bound S10 (Schuwirth et al., 2005; Selmer et al., 2006; Wimberly et al., 2000) several residues that contact the 16S rRNA, including Pro39 and Arg72, have been refined with different conformations compared to the present structure of S10 in complex with NusB. While these data suggest that S10 also adjusts locally to accommodate different binding partners, the limited resolution of the ribosome structures precludes a more detailed comparison.

Table 4.3 Structural comparisons

	<i>eco</i> NusB	<i>mtu</i> NusB	<i>tma</i> NusB	<i>aae</i> NusB	<i>eco</i> S10
PDB ID	1EY1	1EYV	1TZV	2JR0	2AVY
Reference	(Altieri et al., 2000)	(Gopal et al., 2000)	(Bonin et al., 2004b)	(Das et al., 2008)	(Schuwirth et al., 2005)
Sequence Identity (%)	100	29.9	35.1	25.2	100
Matching Cα Atoms	106	111	125	126	76
RMSD (Å)	2.47	1.36	1.45	2.35	1.09

Comparison of NusB protein structures to the NusB molecule of the present NusB-S10^{Alloop} crystal structure and of S10 from the 30S subunit to the S10^{Alloop} molecule of the present NusB-S10^{Alloop} crystal structure. *eco*NusB, NMR structure of *E. coli* NusB; *mtu*NusB, crystal structure of *M. tuberculosis* NusB; *tma*NusB, crystal structure of *Thermotoga maritima* NusB; *aae*NusB, NMR structure of *Aquifex aeolicus* NusB; *eco*S10, crystal structure of *E. coli* S10 in complex with the 30S ribosomal subunit.

4.2.3 Binding of S10 to NusB is mutually exclusive with its incorporation into the ribosome and with NusB dimerization

S10 residues His15, Arg37, Pro39, Ile40, Pro41, Pro43, Thr44, His70 and Arg72, which directly interact with NusB, also directly contact 16S rRNA in the 30S subunit (within 3.5 Å

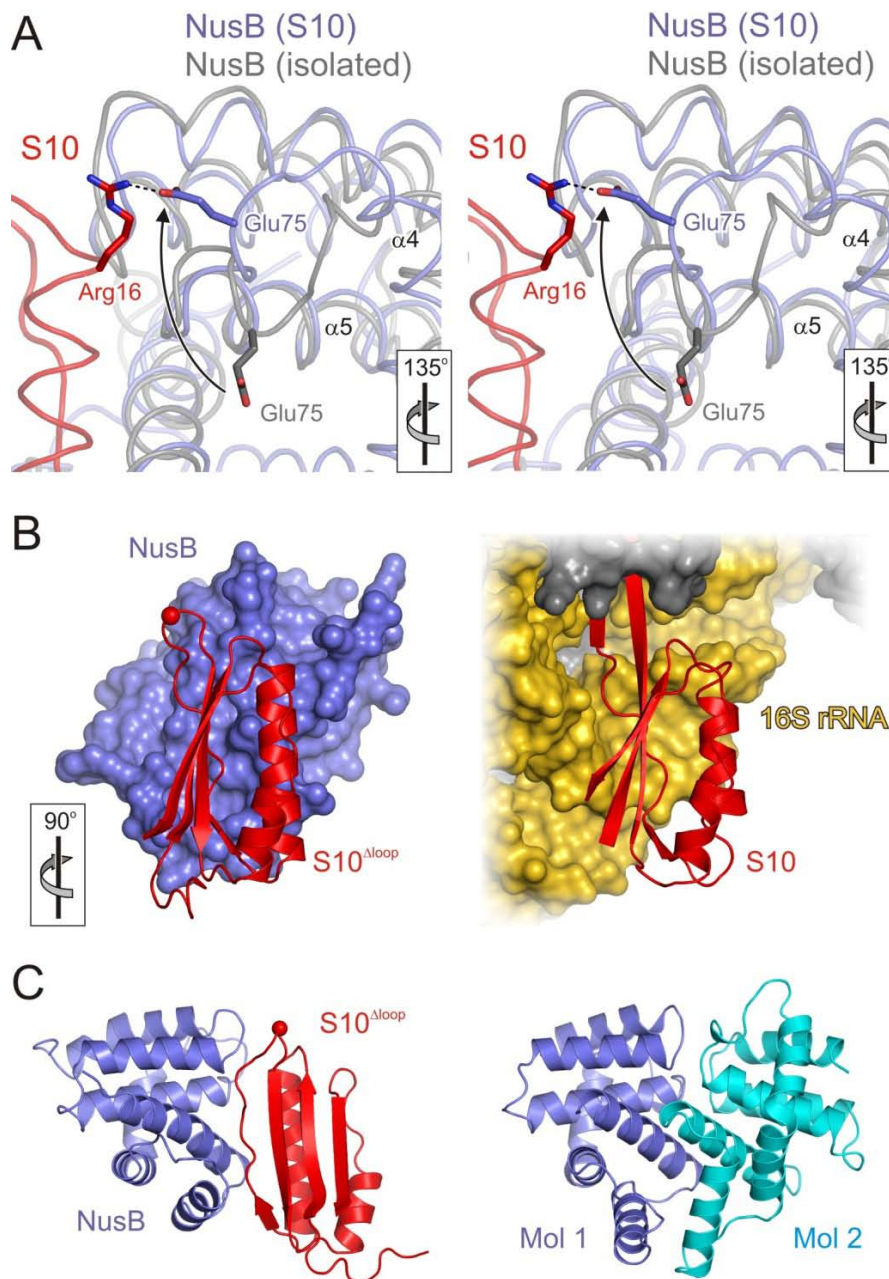


Figure 4.4 Aspects of the NusB-S10^{Δloop} interaction

(A) Stereo ribbon plot showing induced fit adjustment of the loop between helices $\alpha 4$ and $\alpha 5$ in NusB. An NMR structure of *E. coli* NusB (gray; PDB ID 1EY1; (Altieri et al., 2000)) was superimposed on the NusB subunit of the present NusB-S10^{Δloop} complex (blue and red, respectively). The view relative to Figure 4.3A is indicated. Glu75 of NusB changes its position dramatically (arrow) upon complex formation in order to engage in a salt bridge with Arg16 of S10^{Δloop}. Relevant residues are in sticks and colored by atom type as before.

(B) Comparison of S10^{Δloop} (red) binding to NusB (blue) and S10 binding to the remainder of the 30S subunit (rRNA, gold; r-proteins, gray). The orientation relative to Figure 4.3A is indicated.

(C) Comparison of the present NusB-S10^{Δloop} complex (blue and red, left) with the *M. tuberculosis* NusB dimer (blue and cyan, right; PDB ID 1EYV; (Gopal et al., 2000)). The blue NusB molecules of both complexes are in the same orientation.

distance; (Schuwirth et al., 2005)). As a consequence, the surface of S10 that binds NusB is occluded in the ribosome (Figure 4.4B). Thus, contrary to a previous hypothesis (Das et al., 1985; Gopal et al., 2001), S10 cannot participate with NusB in transcription antitermination as a part of the 30S subunit. This finding is in agreement with the observation that processive N-dependent antitermination can be reconstituted using purified S10 and other Nus factors (Das et al., 1985).

Mycobacterium tuberculosis NusB forms dimers (Gopal et al., 2000) whose significance for transcription antitermination has so far remained obscure. Comparison of these dimers to the NusB-S10^{Δloop} complex shows that NusB dimerization would interfere with S10 binding (Figure 4.4C). This observation is in agreement with the previous suggestion (Bonin et al., 2004b) that dimerization may be used as a packaging mechanism by some organisms to downregulate aberrant activities of isolated NusB. Similar autoinhibitory mechanisms have been demonstrated for other transcription antitermination factors (Belogurov et al., 2007; Mah et al., 2000).

4.2.4 Molecular basis of the conserved proline motif on S10

The solved crystal structure of the NusB-S10^{Δloop} complex sheds light on the binding interface between NusB and S10, in which NusB binds a proline motif (Gly38-Pro39-Ile40-Pro41-Leu42-Pro43-Thr44) on S10 (Figure 4.3C-4.3F and 4.5A). This proline motif is highly conserved in bacteria (Figure 4.5B). To investigate how the individual prolines on the proline motif affect the transcription activities, each of three prolines was mutated to alanine, and the N antitermination and Nun termination activities were tested directly by Max Gottesman's group (Table 4.4). All of the S10 proline mutants have a phenotype. Pro39Ala supports N antitermination but not Nun termination. Pro41Ala is toxic. Pro43Ala supports Nun termination (better than wt S10), but not N antitermination. The differences among these mutants suggest that they do not simply affect NusB binding. To further characterize whether the binding of S10 to NusB is affected by a partial proline motif, S10 bearing individual proline to alanine changes were co-expressed with NusB in *E. coli* and purified *via* a GST-tag on the S10^{Δloop}. All the S10^{Δloop} proteins still associated with NusB during purification (Figure 4.5C). Therefore, the full S10 proline motif is not mandatory for interaction with NusB.

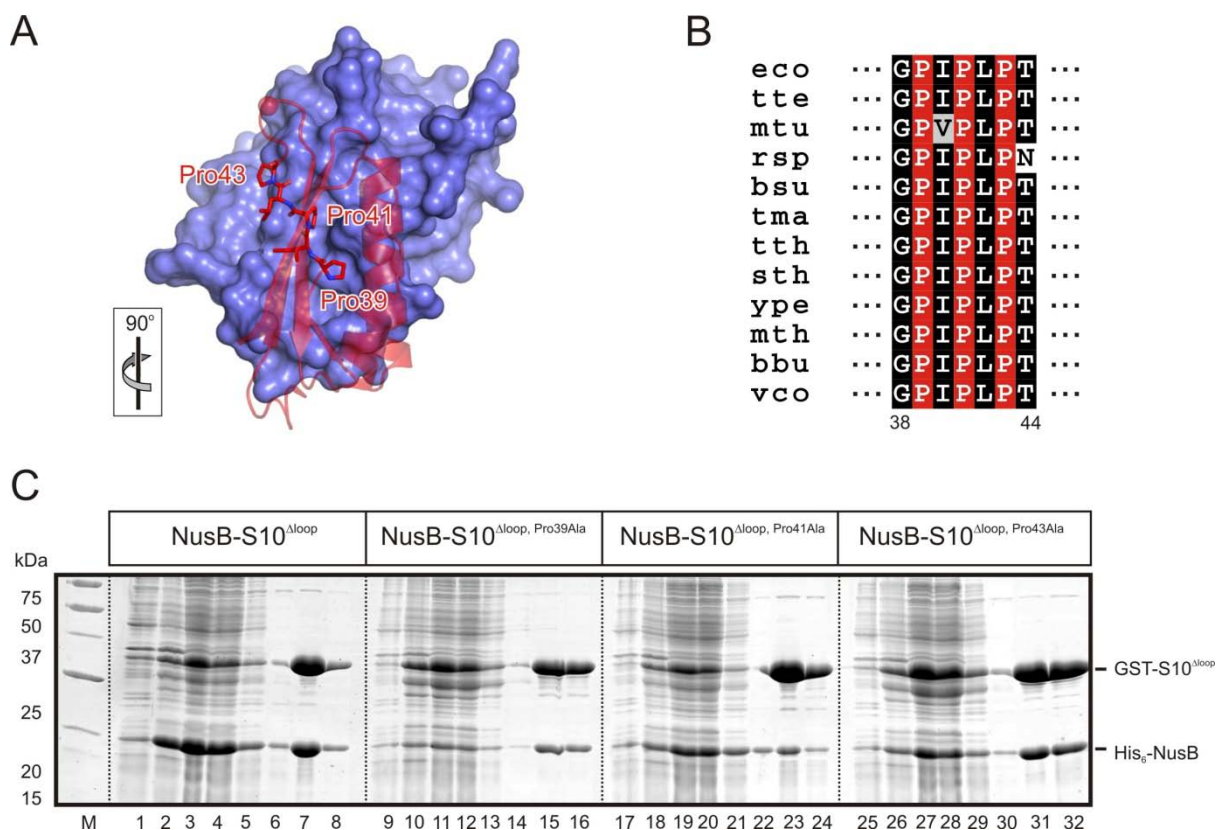


Figure 4.5 Aspects of proline motif of S10

(A) The highlight of proline motif of S10 for NusB binding. NusB is shown a surface view for clearance. Proline residues on S10^{Δloop} are labeled. The orientation relative to Figure 4.3A is indicated.

(B) Alignment of twelve bacterial S10 sequences. Numbering corresponds to *E. coli* S10. The background of conserved prolines is red, and that of other highly conserved amino acids is black. Abbreviations: *eco*, *Escherichia coli*; *tte* *Thermoanaerobacter tengcongensis*; *mtu*, *Mycobacterium tuberculosis*; *rsp*, *Rhodobacter sphaeroides*; *bsu*, *Bacillus subtilis*; *tma*, *Thermotoga maritima*; *tth*, *Thermus thermophilus*; *sth*, *Streptococcus thermophilus*; *ype*, *Yersinia pestis*; *mth*, *Moorella thermoacetica*; *bbu*, *Borrelia burgdorferi*; *vco*, *Vibrio cholera*.

(C) Copurification of GST-S10^{Δloop} or proline mutants with His₆-NusB. Group of eight lanes show the cell extract before induction (first lane), cell extract after induction (second lane), soluble extract (third lane), GSH-sepharose flow-through (fourth lane), the wash (fifth and sixth lanes), and the elution (seventh and eighth lanes). Coexpressed proteins are indicated above the group of lanes. M, molecular mass marker.

Table 4.4 Transcription activities tests by overproduction of S10 or S10 proline mutants

pBAD-NusE	Nun-Termination	N-Antitermination	Toxicity
wt	+	+	-
Pro39Ala	- (antitermination)	+	-
Pro41Ala	+/-	+/-	+
Pro43Ala	++	-	-

The complementation assay for N antitermination activity was carried out at 42 °C in an *E. coli* strain bearing a chromosomal *nusE71* defect. Nun termination activity was tested in the HK022 strain carrying the fusion *nusE R72G cI857 pR cro27 nutR tR1 cII::lacZ*. '+', '++': growth or toxic; '-': no growth or not toxic.

4.2.5 NusB does not influence the *cis/trans* equilibrium at Pro39 of S10

S10-Pro39 is molded into a *cis* conformation that organizes the neighboring Gly38 residue to interact with NusB, which has further long range effects *via* orienting other NusB residues that engages in additional S10 contacts (Figure 4.6B (left) and 4.6C (left)). While in the ribosome-bound S10, Pro39 adopts a *trans* conformation (Figure 4.6C (middle) (Schuwirth et al., 2005)). The phenomenon points out the question whether NusB influences the *cis/trans* equilibrium at Pro39 of S10. Firstly, S10^{Δloop, Pro39Ala} was still able to sustain NusB-S10 interaction in GST-pull-down assays (Figure 4.5C, lanes 9-16) that suggests Pro39 is not required for NusB binding. Secondly, the crystal structure of the NusB-S10^{Δloop, Pro39Ala} complex was solved by molecular replacement at 1.7 Å resolution with R_{work} and R_{free} factors of 20.0 % and 23.8 %, respectively (Table 4.2). The global structure of S10^{Δloop, Pro39Ala} in complex with NusB (Figure 4.6A) is virtually identical to that of the NusB-S10^{Δloop} complex (rmsd of 0.85 Å for 207 Cα atoms), indicating that the Pro39Ala mutation has no global conformational consequences. The S10-Ala39 is molded into a *trans* conformation (Figure 4.6C (right)) that allows formation of a longer β2-sheet (Figure 4.6A). In particular, the *trans* Ala39 still makes interaction with Ser19 of NusB as the *cis* Pro39 does (Figure 4.6B), demonstrating that a *cis* conformation at S10 residue 39 is not required for NusB binding.

The enzyme peptidyl-prolyl *cis/trans* isomerase (PPIase), which catalyses the *cis/trans* isomerization of proline imidic peptide bonds in oligopeptides (Lang et al., 1987), is considered as a folding chaperon to aid proteins folding. Cyclophilin A (CypA) is one of the PPIases that has been found in a variety of functional contexts (Piotukh et al., 2005). The preferred peptide sequence for CypA binding is Gly-Pro-X (any amino acid)-hydrophobic (Piotukh et al., 2005), which matches the proline motif on S10 (Gly38-Pro39-Ile40-Pro41) for NusB binding. The evidence expresses a tendency that NusB might work as a PPIase. In contrast, the proline on amino acid 39 of S10 is not necessary for NusB binding and further for λ N antitermination activity. A *cis* conformation is not required for NusB binding. Moreover, NusB was not found to catalyze *cis/trans*-prolyl isomerization of the model substrate *N*-Suc-Ala-Ala-Pro-Phe-*p*-nitroanilide (data not shown) in an assay used to identify the PPIase (Fischer et al., 1984). Therefore, little evidence was found that NusB could act as a PPIase to influence the *cis/trans* equilibrium at Pro39 of S10.

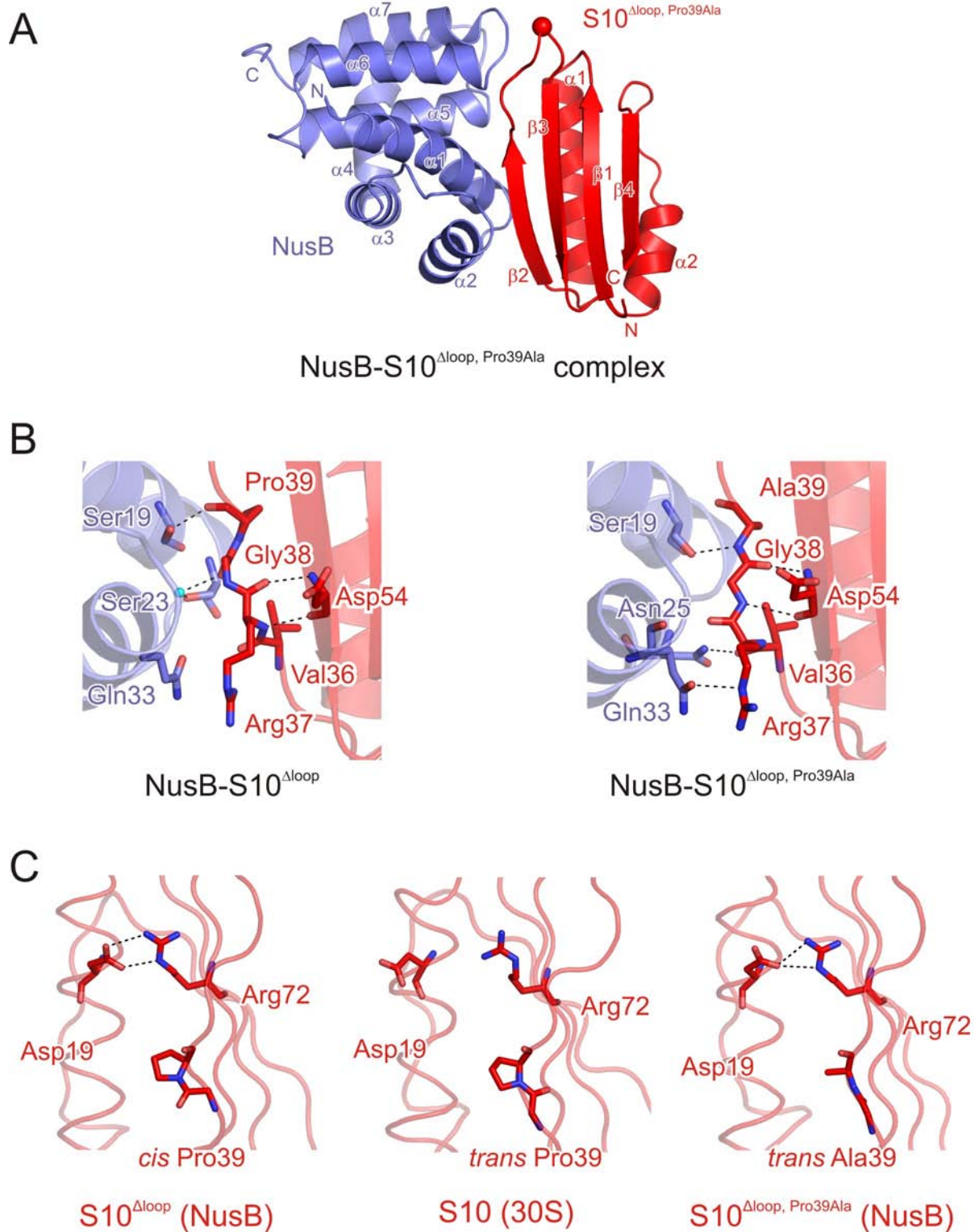


Figure 4.6 Structure of the NusB-S10^{Δloop, Pro39Ala} complex

(A) Ribbon plot of the *E. coli* NusB-S10^{Δloop, Pro39Ala} complex. NusB, blue; S10^{Δloop, Pro39Ala}, red. Secondary structure elements and termini are labeled. The red sphere makes the site at which the ribosome binding loop of S10^{Δloop, Pro39Ala} has been replaced by a single serine.

(B) Comparison of S10^{Δloop} binding to NusB with S10^{Δloop, Pro39Ala} binding to NusB. Interacting residues are labeled. Cyan sphere, water molecule. Dashed lines, hydrogen bonds or salt bridges.

(C) Ribbon plot showing residue 39 of S10 induces *cis* or *trans* conformation in various sources. Left: the *cis* conformation of Pro39 on S10^{Δloop} in complex with NusB; Middle: the *trans* conformation of Pro39 on S10 in the 30S subunit (PDB ID 2AVY; (Schuwirth et al., 2005)); Right: the *trans* conformation of Ala39 on S10^{Δloop, Pro39Ala} in complex with NusB.

4.2.6 Molecular basis of the *nusB5* and *nusE100* phenotypes

Mutations in *nusB* and *nusE* have served as important genetic tools to study processive antitermination. However, the biochemical basis for the dysfunction or suppressor activity of any of the mutant proteins was not defined. The *nusB5* allele gives rise to a Tyr18Asp mutation in NusB (Court et al., 1995) and leads to a defect in N-dependent antitermination, which blocks λ growth (Friedman et al., 1976). The *nusE100* mutation restricts N-termination but not N-antitermination (Robledo et al., 1991). An *E. coli* cell strain bearing a chromosomal *nusE100* defect was obtained from Max Gottesman's group. I have sequenced the *nusE100* allele and found that it encodes an S10^{Arg72Gly} variant. Remarkably, NusB-Tyr18 and S10-Arg72 are both involved in the same buried, hydrophilic, intermolecular interaction network at the center of the NusB-S10 interface (Figure 4.3C). Replacement of NusB-Tyr18 by Asp or replacement of S10-Arg72 by Gly is expected to interfere with this interaction network. Therefore, it is possible that the defects of the mutant alleles are in part caused by a weakened NusB-S10 affinity. This idea was tested by monitoring the ability of the mutant proteins to sustain NusB-S10 interaction in GST-pull down assays. Indeed, the NusB5 variant (Tyr18Asp) did not bind to S10^{Δloop} (Figure 4.1A, lanes 7-9) and the S10^{Δloop, Arg72Gly} mutant protein of *nusE100*^{Δloop} failed to interact with NusB (Figure 4.1A, lanes 10-12).

Previously, the lack of production of a stable gene product was thought to be the cause of the *nusB5* defect in N-antitermination (Mason et al., 1992). The results presented here, in contrast, suggest that *nusB5* gives rise to a gene product that is less active due to a weakened interaction with S10. In that case, the *nusB5* defect may be overcome simply by mass action. Therefore, the question is if high levels of NusB5 can restore N antitermination. Indeed, the complementation analysis by Max Gottesman's group showed that overexpression of the NusB5 protein in an *E. coli nusB* deletion strain partially rescued λ growth (Table 4.5). In agreement with this observation, multiple copies of a plasmid carrying the *nusB5* gene have previously been found to complement a chromosomal *nusB5* allele (Court et al., 1995). Thus, the data underscore the importance of a stable NusB-S10 interaction at physiological expression levels of these proteins for N and Nun activities.

Table 4.5 Overexpression of NusB5 overcomes the *nusB5* defect

pBAD	λ wt	λ r32
-	$<10^{-6}$	$<10^{-4}$
<i>nusB</i>	0.73	0.61*
<i>nusB5</i>	0.35	$<10^{-2}$

Values are Efficiency of Plating (EOP) relative to a *nusB*⁺ strain with no plasmid. Strains are *nusB::Cam* derivatives of W3102. Phage were titered at 42 °C. To induce the pBAD promoter, 0.1 % arabinose, was added to the plates. λ wt = λ imm434; λ r32 = λ imm434 r32. The r32 insertion increases the dependency of λ on N. * = small plaques.

4.2.7 The *nusE71* mutation defines an additional interaction surface on S10

The residue affected by the *nusE71* mutation (Ala86Asp), which blocks both N and Nun, is remote from the NusB interface and the RNA-binding region of S10. I have determined the crystal structure of the NusB-S10 ^{Δ loop, Ala86Asp} complex by molecular replacement at 2.6 Å resolution (Figure 4.7A). The structure was refined to R_{work} and R_{free} factors of 21.8 % and 28.0 %, respectively (Table 4.2). The structure of the NusB-S10 ^{Δ loop, Ala86Asp} complex is virtually identical to the wt complex (rmsd of ca. 0.8 Å for 206 C α atoms; Figure 4.7A). Therefore, dysfunction of S10^{Ala86Asp} is not due to a global effect on the structure of the protein. Rather, the mutation changed the local surface properties of S10 (Figure 4.7B). This finding suggests that yet another molecular interaction of S10 may be attenuated in S10^{Ala86Asp}. S10 is known to interact directly with RNAP (Mason and Greenblatt, 1991) and it is possible that the helix α 2 region encompassing residue 86 is involved in this association. Alternatively, the S10 region around Ala86 might mark an interface with N and Nun; N and S10 proteins are reported to co-purify in some preparations (Mogridge et al., 1995, 1998b). This notion would explain the otherwise puzzling observation that *nusE71* does not block *rrn* antitermination, which uses RNAP but not N or Nun (Zellars and Squires, 1999).

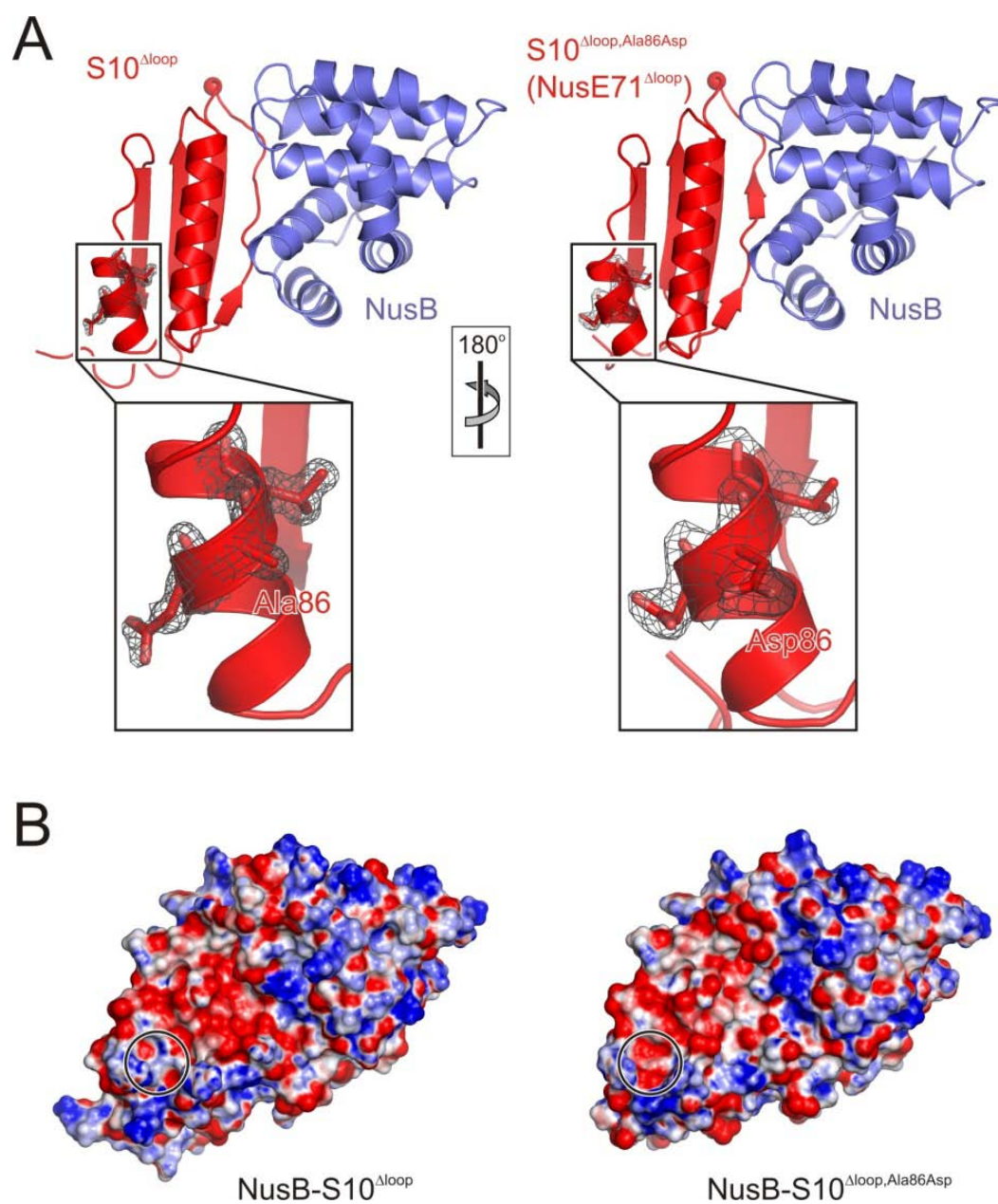


Figure 4.7 Structure of the NusB-S10^{Δloop, Ala86Asp} complex

(A) Comparison of the NusB-S10^{Δloop} complex (left) with the NusB-S10^{Δloop, Ala86Asp} complex (right). Gray meshes indicate the final 2F_o-F_c electron densities covering residue 86 and neighboring residues of the S10^{Δloop} molecules, contoured at the 1σ level. Insets show closeup views of the residue 86 regions. The orientations relative to Figure 4.3A are indicated.

(B) Comparison of the electrostatic surface potentials of the complexes. Blue, positive charge; red, negative charge. Left, NusB-S10^{Δloop} complex; Right, NusB-S10^{Δloop, Ala86Asp} complex. The positions of residue 86 are circled. The orientations are the same as in panel (A).

4.3 BoxA RNA binding by the NusB-S10 complex

4.3.1 The NusB-S10 complex exhibits an intermolecular, mosaic and contiguous BoxA RNA-binding surface

To explore the mechanism by which S10 enhances the affinity of NusB for BoxA RNA, the interactions of the NusB and the NusB-S10 complex with BoxA-containing RNA were investigated. This work was done by Henning Urlaub's group through mass spectrometric analysis of UV-induced crosslinking sites based on materials prepared by me. NusB-RNA or NusB-S10-RNA complexes were exposed to UV light and the zero-length crosslinks generated were analyzed by mass spectrometry (Table 4.6). The same two 19-nucleotide λ NutR BoxA or *rrn* BoxA RNAs, previously used in fluorescence-based interaction studies (Greive et al., 2005), were employed (Figure 4.8). Overall, Henning Urlaub's group identified four peptides in NusB (B1, B1', B2 and B3) and three peptides in full-length S10 (E1, E2 and E3) that crosslinked to distinct, short RNA elements (Figure 4.8; Table 4.6). UV-induced crosslinking in the absence of RNA oligos and mass analysis with complete mixtures but without UV-irradiation did not give rise to any peaks corresponding to those of the identified peptide-RNA crosslinks.

Table 4.6 Electrospray-ionization tandem mass spectrometry identification of protein-RNA crosslinks

NusB-S10			NusB					
λ BoxA ^a	<i>rrn</i> BoxA ^a	λ BoxA Core ^a	λ BoxA	<i>rrn</i> BoxA	Exp. M.W.	Crosslinked Peptide+RNA	Peptide M.W.	RNA M.W.
X	X		X	X	2488.20	NusB SDVPYKVAINEAIELAK (B1)+CU	1859.00	629.08
X	X		X	X	2644.25	NusB RSDVPYKVAINEAIELAK (B1')+CU	2015.11	629.08
X			X		2807.03	NusB SFGAEDSHKFVNGVLDK (B2)+UAC	1848.90	958.13
X	X		X	X	1520.56	NusB FVNGVLDK (B3)+UU	890.48	630.06
				X	1826.57	NusB FVNGVLDK (B3)+UUU	890.48	936.09
X	X				1520.62	S10 LKAFDHR (E1)+UA-18	885.48	653.09
X	X	X			1538.70	S10 LKAFDHR (E1)+UA	885.48	653.09
	X				2541.88	S10 FTVLISPHV NK (E2)+AAUU	1253.71	1288.17
	X				2235.85	S10 FTVLISPHV NK (E2)+AAU	1253.71	982.14
X	X				1804.62	S10 DQYEIR (E3)+AAU	822.38	982.14

^a λ BoxA: CAC CGC UCU UAC ACA AUU A;

rrn BoxA: CAC UGC UCU UUA ACA AUU A

λ BoxA core: CGC^{Br}UCU UAC ACA; ^{Br}U, 5-bromo uridine

^b “X” indicates observation of a particular crosslink in a given mixture.

Scrutinizing the sequences of the crosslinked peptides and RNA elements allowed me to deduce unequivocally the molecular neighborhoods (Section 3.2.4.5). These data show that NusB is in close proximity to residues 3-8 of *rrn* BoxA and residues 3-9 of λ BoxA, while S10 is in direct contact with residues 8, 9 and 12 of *rrn* BoxA, residues 7, 8 and 12 of λ BoxA and with residues just downstream in either of the RNAs (Figure 4.8; Table 4.6). The BoxA positions in direct contact with NusB and S10 are remarkably congruent with residues 2-9 of *rrn* BoxA and residues 2-7 of λ BoxA, which are essential for recruitment of NusB and S10 to the antitermination machineries (Mogridge et al., 1998b). In addition, parts of the S10 ribosome-binding loop (the entire peptide E2 and part of E3) crosslinked to RNA at the very 3'-end of the core BoxA elements and to nucleotides immediately downstream (Figure 4.8; Table 4.6). Thus, the ribosome-binding loop fosters auxiliary, but not essential (Table 4.1), mRNA contacts that might enhance processive antitermination.

Since identical crosslinks of NusB to the RNAs in the absence or in the presence of S10 (Table 4.6) were found, the specificity of the NusB-BoxA RNA contacts is influenced little if at all by S10. Thus, the direct S10-BoxA interactions detected herein are responsible for the increased BoxA RNA affinity of the NusB-S10 complex compared to NusB alone. Since isolated S10 binds RNA weakly and largely non-specifically (Greive et al., 2005), NusB apparently stabilizes an RNA-binding conformation of S10 and positions S10 on the BoxA RNA, suggesting that NusB loads S10 onto a specific RNA element.

The amino acid residues crosslinked to RNA in both NusB and S10 are dispersed in the primary sequences but nevertheless coalesce in 3D on one surface of the NusB-S10^{Alloop} complex (Figure 4.8). Peptides B1 (B1'), B2 and B3 form a contiguous surface on NusB and peptide E1 of S10 directly neighbors the C-terminus of peptide E3 at the base of the ribosome-binding loop. The tip of NusB peptide B2 is in weak direct contacts with S10 peptides E1 and E3 (Figure 4.8). Thus, NusB and S10 together present a contiguous, mosaic BoxA-binding surface. The results from structural and crosslinking analyses were combined to derive the overall topology of the NusB-S10-RNA complex (Figure 4.8, inset). The central region of a BoxA element is placed on the confluent binding surface of the protein complex. The RNA runs 5'-to-3' from the NusB to the S10 RNA-binding patches.

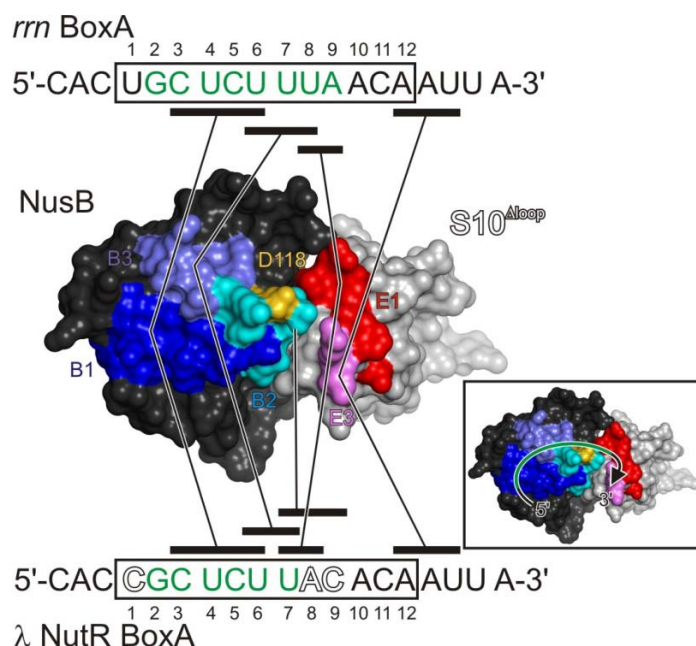


Figure 4.8 Mapping of crosslinked peptides on the surface of the NusB-S10^{Δloop} complex

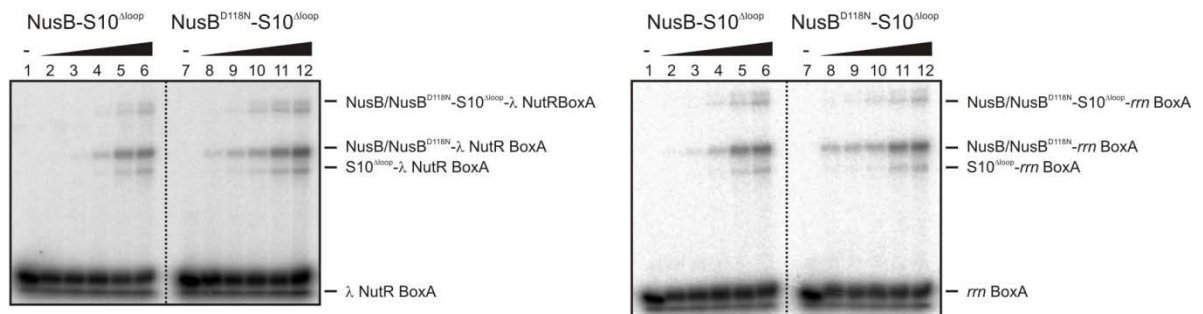
The view is from the top of Figure 4.3A. NusB, dark gray; S10, light gray. Crosslinked peptides of NusB (B1, B2, B3, B4; see Table 4.6 for peptide sequences) are light blue, dark blue, cyan and steel blue, respectively. Crosslinked peptides of S10 (E1 and E3) are red and violet, respectively. Asp118, which is mutated to Asn in the *nusB101* allele, is colored gold. RNAs encompassing the *rrn* and λ BoxA elements and used for crosslinking are given above and below the structure, respectively. Boxed regions with residue numbers indicate the core BoxA elements. Residues in green of *rrn* BoxA RNA and λ BoxA RNA have previously been implicated in recruitment of NusB and S10 to antitermination complexes by mutational analysis (Mogridge et al., 1998b). Outlined residues differ in λ BoxA compared to *rrn* BoxA. Black bars designate crosslinked regions of the RNAs. They are connected by lines to the peptides, to which they have been crosslinked (see Table 4.6 for a complete list of crosslinks). Inset 1 illustrates the deduced topology of the NusB-S10-BoxA RNA complexes.

4.3.2 *nusB101* represents a gain-of-function mutation with increased RNA affinity

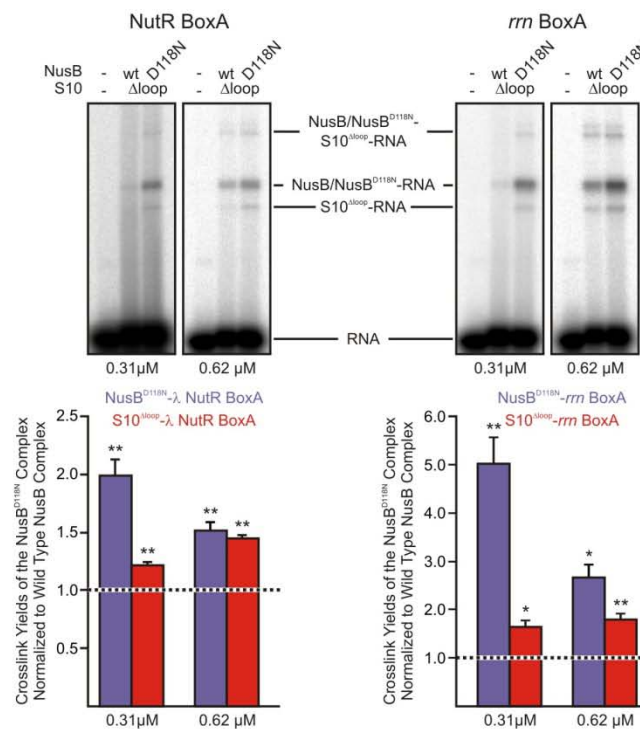
The *nusB101* mutation (Asp118Asn) suppresses the N-antitermination defects of NusA1 and NusE71 mutants at high temperatures (Ward et al., 1983). Notably, NusB-Asp118 is part of peptide B2, which lies at the center of the closely spaced RNA-binding patches on NusB and S10 (Figure 4.8). Removal of a negative charge at the NusB-118 position could conceivably increase the RNA affinity of NusB and of the NusB-S10 complex, in agreement with a previous proposal (Court et al., 1995). This idea was tested by crosslinking increasing amounts of NusB-S10^{Δloop} and of NusB101-S10^{Δloop} (NusB^{Asp118Asn}-S10^{Δloop}) to BoxA-containing RNAs under conditions where the crosslink yields reflect the binding equilibria. As predicted, NusB101-S10^{Δloop} exhibited increased affinities for either λ or *rrn* BoxA

sequences (Figure 4.9A and 4.9B). Thus, *nusB101* represents a gain-of-function mutation that increases the affinity of NusB for BoxA. Consistent with cooperativity among the component antitermination factors, enhanced RNA affinity in NusB101 might compensate for decreased RNA affinity of the suppressed NusA1 mutant, in which a core residue of the S1 RNA-binding domain is altered (Worbs et al., 2001).

A



B



C

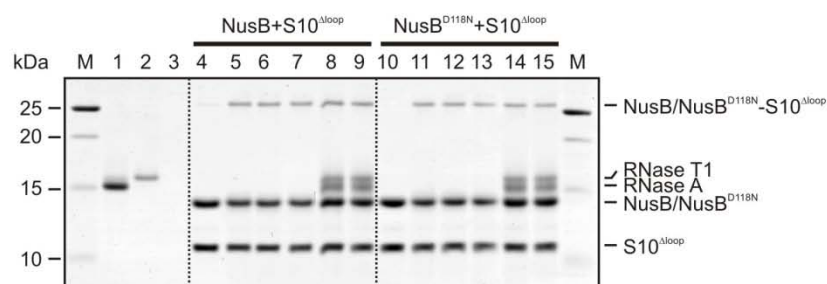


Figure 4.9 Protein-RNA and protein-protein crosslinking analysis

(A) Comparative crosslinking of λ BoxA RNA (left) or *rrn* BoxA RNA (right) to increasing amounts of NusB-S10^{Δloop} complex (lanes 1-6) or NusB^{Asp118Asn}-S10^{Δloop} (NusB101-S10^{Δloop}) complex (lanes 7-12). Protein concentrations were 0 (lanes 1, 7), 0.15 (lanes 2, 8), 0.31 (lanes 3, 9), 0.62 (lanes 4, 10), 1.25 (lanes 5, 11) and 2.5 μ M (lanes 6, 12). After 5 minutes of UV irradiation (all lanes), crosslinked species were resolved on an SDS gel and visualized by autoradiography. Crosslinking yields scaled with the concentration of the proteins, showing that they reflect the equilibrium binding situations. The yields of all crosslinked species (S10^{Δloop}-RNA, NusB/NusB^{Asp118Asn}-RNA, NusB/NusB^{Asp118Asn}-S10^{Δloop}-RNA indicated on the left) were higher in the presence of NusB101 (NusB^{Asp118Asn}; lanes 8-12) compared to wild type NusB (lanes 2-6). Crosslinking with either RNA gave rise to significant amounts of NusB-S10^{Δloop}-BoxA RNA or NusB101-S10^{Δloop}-BoxA RNA ternary crosslinks. Additionally, bands assigned to the ternary crosslinks were heterogeneous. These observations can be explained by NusB (or NusB101) and S10^{Δloop} (or S10) undergoing efficient UV-induced protein-protein crosslinking (see panel C). Thus, under the experimental conditions, covalent ternary complexes with different topologies (such as NusB/NusB101 crosslinked to S10^{Δloop} plus one of the proteins crosslinked to RNA; both proteins crosslinked to RNA and not crosslinked to each other; both proteins crosslinked to each other and both crosslinked to RNA) are obtained. These topologically different ternary complexes will exhibit different mobilities on SDS gels.

(B) Top: Representative crosslinking of λ NutR BoxA RNA (left two panels) or *rrn* BoxA RNA (right two panels) to NusB-S10^{Δloop} or NusB101-S10^{Δloop} (NusB^{Asp118Asn}-S10^{Δloop}). Two concentrations of protein complex (0.31 μ M and 0.62 μ M) were crosslinked, resolved on SDS gels and visualized by autoradiography. In each panel, RNA alone is in the left lane, NusB-S10^{Δloop} complex in the central lane and NusB101-S10^{Δloop} complex in the right lane. Bottom: Quantification of crosslink yields. Values are the crosslink yields of the protein components of the NusB101-S10^{Δloop} samples, relative to the crosslink yields of the corresponding components of the NusB-S10^{Δloop} samples. The crosslink yields of the components of the NusB-S10^{Δloop} samples were set at 100 % (dashed lines). Values represent the means of three independent experiments +/- the standard errors of the mean. Asterisks - $p \leq 0.032$; double asterisks - $p \leq 0.020$.

(C) UV-induced protein-protein crosslinking. 20 μ M of NusB-S10^{Δloop} or NusB101-S10^{Δloop} (NusB^{Asp118Asn}-S10^{Δloop}) were treated either with 1 μ l Benzonase (Novagen) or with 1 μ l each of RNase A and RNase T1 (Ambion) in 10 μ l reaction volumes at 37 $^{\circ}$ C for 1 hour or at 4 $^{\circ}$ C for 16 hours in order to remove any potential traces of contaminating RNA. Samples were exposed to 254 nm ultraviolet light for 5 min at 4 $^{\circ}$ C. Reactions were analyzed by 15 % SDS polyacrylamide gel electrophoresis. The gel was Coomassie stained. The same results were obtained when S10 was used instead of S10^{Δloop}. Lane 1, RNase A; Lane 2, RNase T1; Lane 3, Benzonase; Lane 4, NusB-S10^{Δloop}, no UV; Lane 5, NusB-S10^{Δloop}, 5 min UV; Lane 6, NusB-S10^{Δloop} plus Benzonase, 37 $^{\circ}$ C, 5 min UV; Lane 7, NusB-S10^{Δloop} plus Benzonase, 4 $^{\circ}$ C, 5 min UV; Lane 8, NusB-S10^{Δloop} plus RNases A and T1, 37 $^{\circ}$ C, 5 min UV; Lane 9, NusB-S10^{Δloop} plus RNases A and T1, 4 $^{\circ}$ C, 5 min UV; Lane 10, NusB^{Asp118Asn}-S10^{Δloop}, no UV; Lane 11, NusB^{Asp118Asn}-S10^{Δloop}, 5 min UV; Lane 12, NusB^{Asp118Asn}-S10^{Δloop} plus Benzonase, 37 $^{\circ}$ C, 5 min UV; Lane 13, NusB^{Asp118Asn}-S10^{Δloop} plus Benzonase, 4 $^{\circ}$ C, 5 min UV; Lane 14, NusB^{Asp118Asn}-S10^{Δloop} plus RNases A and T1, 37 $^{\circ}$ C, 5 min UV; Lane 15, NusB^{Asp118Asn}-S10^{Δloop} plus RNases A and T1, 4 $^{\circ}$ C, 5 min UV. M, molecular weight marker. The presence of both NusB or NusB^{Asp118Asn} and S10^{Δloop}

in the slower migrating bands obtained upon UV crosslinking was verified by tryptic mass spectrometric peptide fingerprinting.

4.3.3 The structure of NusB^{Asp118Asn}-S10^{Δloop} closely resembles the structure of NusB-S10^{Δloop}

To investigate whether the increased BoxA RNA affinity of the NusB^{Asp118Asn}-S10^{Δloop} complex correlates with structural rearrangements compared to the NusB-S10^{Δloop} complex, the crystal structure of the NusB^{Asp118Asn}-S10^{Δloop} complex was solved by molecular replacement at 2.5 Å resolution (Figure 4.10A). The structure was refined to R_{work} and R_{free} factors of 20.4 % and 25.6 %, respectively (Table 4.2). An asymmetric unit of the crystal contains three molecules each of NusB^{Asp118Asn} and S10^{Δloop}, which were assembled as three NusB^{Asp118Asn}-S10^{Δloop} complexes (Table 4.2). Two of the complexes exhibited well defined electron density. The electron density map of the third complex was fragmentary. In that complex, residues 60-77 and 127-139 of NusB^{Asp118Asn} and residues 45-47 and 60-72 of S10^{Δloop} could not be unambiguously traced. The following discussion refers to the structures of the two well defined complexes, which are very similar (rmsd of 0.53 Å for 220 C α atoms).

The global structure of NusB^{Asp118Asn} in complex with S10^{Δloop} is very similar to that of wt NusB in isolation (PDB ID 1EY1; (Altieri et al., 2000); rmsd of 2.54 Å for 110 C α atoms; Figure 4.10B and 4.10C). Furthermore, the structure of the NusB^{Asp118Asn}-S10^{Δloop} complex is virtually identical to that of the NusB-S10^{Δloop} complex (rmsd of 0.82 Å for 220 C α atoms; Figure 4.10C), demonstrating that the D118N mutation has no global conformational consequences. In particular, the locations of NusB residues N118 in the mutant and of residue D118 in the parent complex are virtually identical. Irrespective of the amino acid at position 118, the neighboring region undergoes identical adjustments upon S10^{Δloop} binding, during which the C α position of residue 118 is repositioned by ca. 2.8 Å (Figure 4.10C, inset). However, the Asp118Asn exchange leads to a significant difference in the local electrostatic surface properties of the complex (Figure 4.10D). Thus, the increased BoxA RNA affinity of NusB^{Asp118Asn} or its complex with S10 is due to the replacement of a negatively charged residue with an uncharged residue at the RNA binding site, which reduces the repulsion with the negatively charged sugar-phosphate backbone of the RNA. Alternatively or in addition, an asparagine compared to an aspartate at position 118 may engage in additional hydrogen bonds to the RNA.

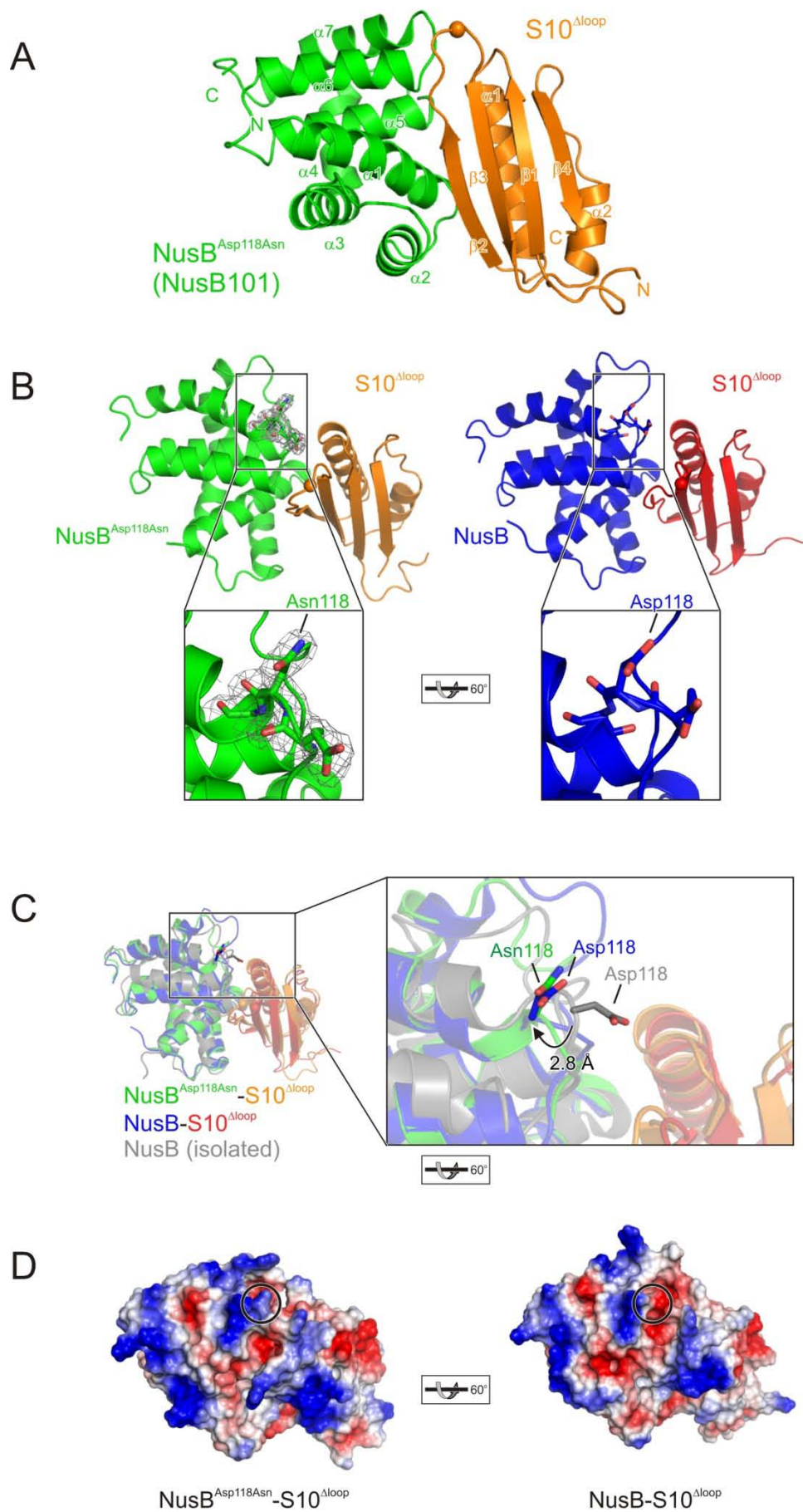


Figure 4.10 Structure of the NusB^{Asp118Asn}-S10^{Δloop} complex

(A) Ribbon plot of the *E. coli* NusB^{Asp118Asn}-S10^{Δloop} complex. NusB^{Asp118Asn}, green; S10^{Δloop}, orange. Secondary structure elements and termini are labeled. The orange sphere makes the site at which the ribosome binding loop of S10 has been replaced by a single serine.

(B) Comparison of the NusB^{Asp118Asn}-S10^{Δloop} complex (left) with the NusB-S10^{Δloop} complex. Insets: closeup views of the residue 118 regions. The orientation relative to (A) is indicated. Gray mesh, final 2F_o - F_c electron density of the NusB^{Asp118Asn}-S10^{Δloop} structure contoured at the 1σ level and covering Asn118 and neighboring residues. The orientation relative to (A) is indicated.

(C) Superimposition of the NusB-S10^{Δloop} complex (blue and red) and of NusB (gray, PDB ID 1EY1; (Altieri et al., 2000)) on the NusB^{Asp118Asn}-S10^{Δloop} complex (green and orange). Residues at position 118 are shown as sticks and a magnified view of the residue 118 region is provided (carbon, as the respective molecule; oxygen, red; nitrogen, blue). The orientation relative to (A) is indicated.

(D) Comparison of the electrostatic surface potentials of the complexes. Blue, positive charge; red, negative charge. Left, NusB^{Asp118Asn}-S10^{Δloop} complex; Right, NusB-S10^{Δloop} complex. The positions of residue 118 are circled. The orientations are the same as in (B).

4.4 Roles of S10 and NusB in transcription and translation

4.4.1 S10 supports transcription antitermination in the absence of NusB

N and *rrn* antitermination and Nun-termination involve appropriate tethering of BoxA and BoxB RNA sites to RNAP *via* N or Nun and the Nus factors (Nodwell and Greenblatt, 1991). Since S10 directly contacts RNA (this work; Figure 4.8) and RNAP (Mason and Greenblatt, 1991), it is possible that S10 may be the functional antitermination factor in the NusB-S10 complex at Nut sites. To test this idea, Max Gottesman's group overexpressed S10 and S10^{Δloop} in an *E. coli* strain lacking the *nusB* gene. Strikingly, overexpression of either S10 or S10^{Δloop} rescued λ growth, restoring functional N-antitermination in the absence of NusB (Table 4.7). Similarly, S10 or S10^{Δloop} expression rescued Nun-dependent termination in the *nusB*-deletion strain, as determined by the expression of a *lacZ* gene promoter-distal to λ Nut (Table 4.7). These results are at variance with the traditional view that the role of S10 is to recruit NusB to RNAP (Mason and Greenblatt, 1991; Mason et al., 1992). Therefore, NusB, although it engages in more extensive BoxA contacts than S10, merely serves as a loading factor that ensures efficient entry of S10 into these transcription complexes, while S10 constitutes the critical antitermination component of the NusB-S10 complex.

Table 4.7 Overproduction of S10 or S10^{Δloop} allows λ to grow on a Δ*nusB* strain (*nusB::Cam*)

N-Antitermination			
Chromosomal <i>nusB</i>	pBAD Plasmid	Arabinose	λ EOP
+	-	-	1.0
Δ	-	-	<10 ⁻⁵
Δ	<i>nusE</i> ⁺	-	<10 ⁻²
Δ	<i>nusE</i> ^{Δloop}	-	<10 ⁻²
Δ	<i>nusE</i> ⁺	+	0.30
Δ	<i>nusE</i> ^{Δloop}	+	0.71
Nun-Termination			
HK022	pBAD Plasmid	- Arabinose	+ Arabinose
-	<i>nusE</i>	2260	1453
-	<i>nusE</i> ^{Δloop}	2373	1856
+	<i>nusE</i>	951 (58)	213 (85)
+	<i>nusE</i> ^{Δloop}	1208 (49)	406 (78)

N antitermination: *nusE*⁺ and *nusE*^{Δloop} are carried by a pBAD plasmid and, where indicated, induced with 0.1 % arabinose. *λimm434*, which is insensitive to λ repressor, was plated at 37 °C on LB plates with or without (+/-)

50 µg/ml ampicillin to determine EOP. Strains are W3102 derivatives carrying the fusion $\lambda cI857 - pR - cro$ (ΔRBS) - *nutR* - *tR1* - *cII::lacZ*. Nun termination: *nusB::Cam* cells additionally carried a HK022 prophage as indicated. Cells were grown at 37 °C until early log phase and then shifted to 42 °C for 2 hours without (“- Arabinose”) or with (“+ Arabinose”) 0.1 % arabinose. Numbers in parentheses indicate % termination.

4.4.2 NusB delivers S10 into other molecular environment

When crystallization trials for the NusB-S10 ^{Δ loop} complex were performed, the complex could be crystallized in an orthorhombic (P2₁2₁2₁) or a hexagonal (P6₁) space group under different conditions (Table 3.5). The crystal form with space group P2₁2₁2₁ (unit cell: 40.7 Å, 49.0 Å, 122.8 Å; 90°, 90°, 90°) allowed me to determine the structure of the NusB-S10 ^{Δ loop} complex, which presents one hetero-dimer in an asymmetric unit with solvent content of 43 % (Table 4.2). The asymmetric unit of the second crystal form with space group P6₁ (unit cell: 63.2 Å, 63.2 Å, 61.2 Å; 90°, 90°, 120°) is too small to harbor the NusB-S10 ^{Δ loop} complex. Since there is no indication of twinning, it is possible that one protein might have crystallized alone. Therefore, the crystals were checked by SDS-PAGE gel. Indeed, NusB protein was lost during crystallization (Figure 4.11B).

The structure of S10 ^{Δ loop} in isolation was solved by molecular replacement from the data of the second crystal form (Table 4.2; Figure 4.11A). The asymmetric unit contains one S10 ^{Δ loop} (Table 4.2). The structure of S10 ^{Δ loop} in isolation was only refined to R_{work} and R_{free} factors of 25.3 % and 29.3 %, respectively (Table 4.2), which are a little higher at 1.9 Å resolution, since there are some unsolved features related with screw axis in the density map. These unsolved features cannot be generated by the presence of residual NusB in crystals, because S10 ^{Δ loop} residues, Arg37, Pro39, Ile40, Pro41 and Pro43, which directly interact with NusB, also directly participates in S10 ^{Δ loop} alone crystal packing (Figure 4.11C and 4.11F). As a consequence, the NusB-binding surface of S10 ^{Δ loop} is occluded by crystal packing in the crystalline S10 ^{Δ loop} dimer (Figure 4.11C and 4.11F) that rules out the presence of NusB in crystals. Thus, I reasoned that unsolved features might be from the presence of partially aggregated S10 ^{Δ loop} in crystals. The global structure of S10 ^{Δ loop} in isolation is very similar to that of the S10 ^{Δ loop} in complex with NusB (rmsd of 0.81 Å for 79 C α atoms; Figure 4.11D). The C-terminus of S10 ^{Δ loop} in isolation is slightly oriented to the left relative to that of the S10 ^{Δ loop} subunit of NusB-S10 ^{Δ loop} complex (Figure 4.11E). In addition, Pro39 remains a *cis* conformation in the S10 ^{Δ loop} in isolation (Figure 4.11G), evidencing that the *cis* conformation at Pro39 is an intrinsic property of S10. These data suggest that *in vitro* NusB can play a role of a loading factor in delivering S10 into the other molecular environment (crystals).

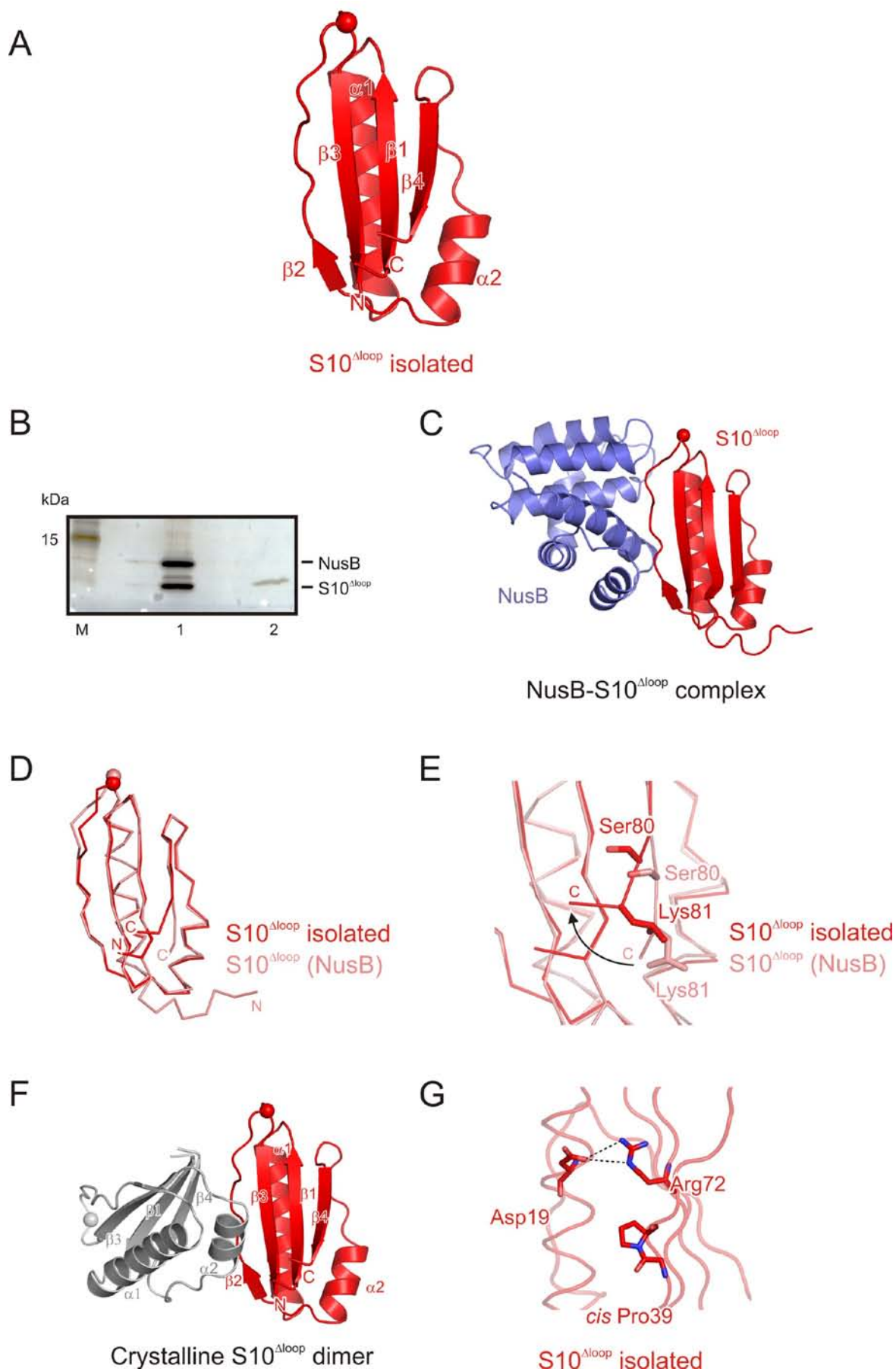


Figure 4.11 Aspects of the S10^{Δloop} in isolation

- (A) Ribbon plot of the S10^{Δloop} in isolation. S10^{Δloop}, red. Secondary structure elements and termini are labeled. The red sphere makes the site at which the ribosome binding loop of S10 has been replaced by a single serine.
- (B) SDS-PAGE gel analysis of S10^{Δloop} in isolation crystals. M, molecular mass marker. Lane 1, the NusB-S10^{Δloop} complex used for crystallization; Lane 2, crystals grew in conditions that yielded S10^{Δloop} alone. Crystals were washed three times with mother liquor and dissolved in the 1× Laemmli loading dye.
- (C) Ribbon plot of the NusB-S10^{Δloop} complex. NusB, blue; S10^{Δloop}, red.
- (D) Superimposition of S10^{Δloop} in isolation (red) on the S10^{Δloop} (pink) from NusB-S10^{Δloop} complex.
- (E) Ribbon plot showing detailed comparison of C-terminus of S10^{Δloop} in isolation (red) with that of the S10^{Δloop} (pink) from the NusB-S10^{Δloop} complex. The structure of the S10^{Δloop} in isolation was superimposed on the S10^{Δloop} subunit of the NusB-S10^{Δloop} complex. The C-terminus of S10^{Δloop} in isolation is slightly oriented to the left (arrow) relative to that of the S10^{Δloop} subunit of the NusB-S10^{Δloop} complex.
- (F) Ribbon plot of the crystalline S10^{Δloop} dimer. S10^{Δloop}, red; Crystalline S10^{Δloop}, grey.
- (G) Ribbon plot showing the *cis* conformation of Pro39 on S10^{Δloop} in isolation.

4.4.3 Does NusB escort S10 into ribosomes?

It was shown that NusB plays a role of a loading factor in transcription antitermination and *in vitro* (where NusB delivers S10 into crystals). Mutations of NusB in *E. coli* slow down the cell growth rate that suggests its role in translation (Taura et al., 1992). Thus, I speculated that NusB could also function as a loading factor that facilitates the entry of S10 into ribosomes. By this means, S10 would not be detected in assembled ribosomes from an *E. coli* strain bearing the absence of NusB. To test this idea, I employed the same ribosome binding assay as I used before (Section 4.1.2), and directly monitored the association of the GST-S10 to ribosomes from wt *E. coli* strain 9739 and ribosomes from *E. coli* strain 9976 (isogenic to strain 9739) bearing a chromosomal *nusB* gene knockout (*nusB::Cam*). Associations of GST-S10 with crude and salt washed ribosomes were evaluated by Western blot. In wt strain the GST-S10 goes into ribosomes and is well detectable in crude and salt washed ribosomes (Figure 4.12, lanes 3 and 4). To my surprise, GST-S10 in crude and salt washed ribosomes from mutant strain is also detectable and its contents are comparable with those from wt strain (Figure 4.12, lanes 7 and 8). The results presented here cannot clearly answer the question that whether NusB escorts S10 into ribosomes and therefore, it has to be further investigated by other means.

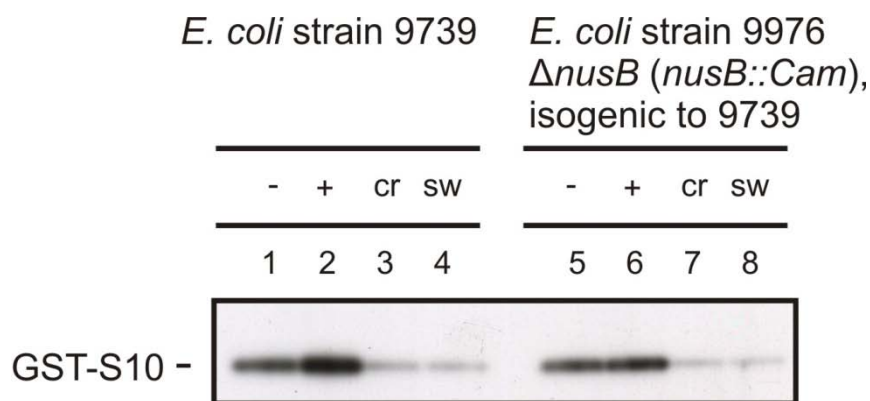


Figure 4.12 Ribosome binding of S10

Western blot probing the binding of GST-S10 to ribosomes from wt *E. coli* strain 9739 and from *E. coli* strain 9976 (isogenic to 9739) bearing a chromosomal *nusB* gene knockout (*nusB::Cam*). Equal amounts of cells before (-; lanes 1 and 5) and after (+; lanes 2 and 6) induction with IPTG as well as equal amounts (0.1 A_{260} equivalents) of crude (cr; lanes 3 and 7) and salt-washed (sw; lanes 4 and 8) ribosomes from *E. coli* strain 9739 expressing GST-S10 (lanes 1-4) or from *E. coli* strain 9976 expressing GST-S10 (lanes 5-8) were analyzed on a 12 % SDS-PAGE gel, transferred to a nitrocellulose membrane and analyzed by Western blot.

4.4.4 NusG couples transcription and translation via S10

In the N and *rrn* antitermination and Nun-termination complexes, NusG is a key component that inhibits transcription pausing and increases the rate of elongation, where NusG-NTD directly binds to RNAP (Burova et al., 1995; Li et al., 1992; Squires and Zaporozhets, 2000). A role for NusG in translation has been identified by the finding that the peptide elongation rate *in vivo* is reduced in *nusG*-depleted cells by measuring the rate of synthesis of a *lacZ* construct (Zellars and Squires, 1999).

The genetic evidence that *nusG4* (S163F) mutation restores λ N antitermination activity in an *E. coli* strain bearing a chromosomal *nusE71* defect indicates a functional interaction between NusG and S10. The genetic interaction might reflect a direct physical contact between the proteins *in vitro*. In order to test this notion, the transcriptionally active NusB-S10 complex (NusB-S10 ^{Δ loop}) was mixed with NusG, and mixtures were subjected to analytical size exclusion chromatography. The mixture of NusB-S10 and NusG eluted 1-2 fractions earlier compared to either NusB-S10 or NusG alone, indicating formation of a stable NusG-NusB-S10 complex (Figure 4.13).

To further characterize the detailed interactions within the complex, Paul Rösch's group investigated the complex formation by NMR based on the NusB-S10 ^{Δ loop} complex, which is more stable than wt NusB-S10 complex. NMR titrations revealed that S10 ^{Δ loop} is the binding component for NusG in the NusB-S10 ^{Δ loop} complex and NusG-CTD directly binds to

S10^{Δloop}. The interaction surface of S10 with NusG is still accessible when S10 forms part of 30S ribosomal subunit. Thus, S10 could mediate simultaneous formation of a NusG-NTD-RNAP complex and a NusG-CTD-ribosome complex. This analysis suggests that NusG establishes a link to ribosome-bound S10 or the S10-NusB complex.

The notion that NusG is a molecular link between transcription and translation is presented by Paul Rösch's group. Detailed results and discussions are not provided here as the majority of the work has been done by them.

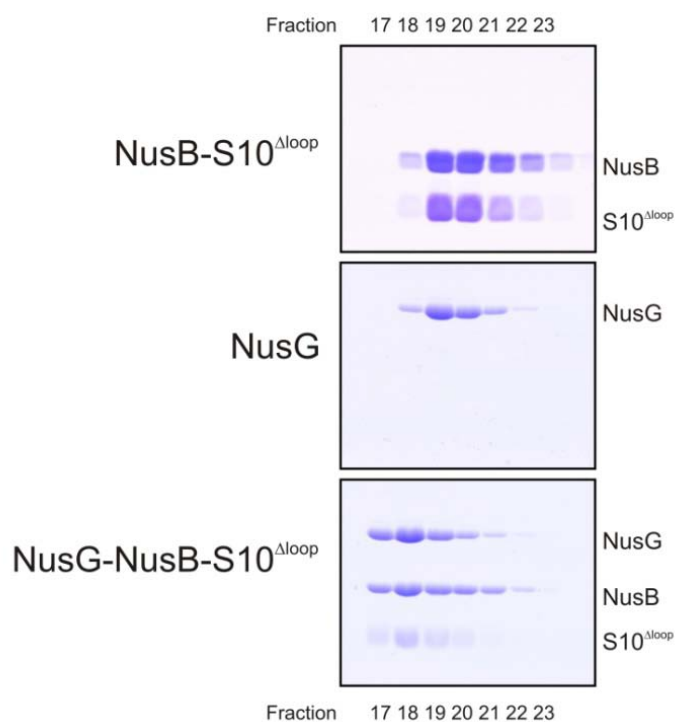


Figure 4.13 Size exclusion chromatography analyses

Size exclusion chromatography analyses of the NusB-S10^{Δloop} complex, NusG and of a mixture of NusB-S10^{Δloop} and NusG. The panels show the migration of NusB-S10^{Δloop} complex (top), NusG (middle) and of a mixture of NusB-S10^{Δloop} and NusG (bottom). Fractions collected from each run were analyzed by SDS-PAGE. Equivalent fractions (17-23) are aligned below each other in the three panels. The mixture of NusB-S10^{Δloop} and NusG elutes 1-2 fractions earlier compared to either NusB-S10^{Δloop} or NusG alone, indicating formation of a stable complex.

5 Discussion

As one strategy to increase the functional diversity of their proteomes, organisms make use of “moonlighting” proteins that can function in more than one cellular context (Jeffery, 1999). In many cases, the molecular basis for the dual activity of these proteins is unknown. Transcription and translation are two highly coupled processes during prokaryotic gene expression. One of mechanisms by which these two cellular contexts communicate directly with one other is sharing proteins. S10 and NusB are two of these shared proteins. Their molecular mechanisms of dual activity in transcription and translation are incompletely understood. Here, I have presented work that defines the roles of a ribosomal protein, S10, which is part of antitermination complexes, and of another antitermination factor, NusB, which forms a stable sub-complex with S10. The results have repercussion for the generation of functional diversity in proteomes by employing one protein for multiple activities.

5.1 S10^{Δloop} is a tool to dissect transcriptional and translational functions of S10

The long ribosome-binding loop of S10 is delineated that is exclusively required for S10 function in the ribosome but not in transcription. The phenomenon that S10^{Δloop} fails to bind to ribosomes (Figure 4.2B) demonstrates that the loop of S10 is the main binding part with other r-proteins and the 16S rRNA in the 30S ribosomal subunit. The finding that the long loop of S10 is dispensable in transcription is consistent with the observation that mutants defective in transcriptional functions (*nusE71*, *nusE100*) map to the globular part of S10. Therefore, transcriptional and translational functions are attributed to distinct regions of S10. The functional architecture of S10 is paralleled by that of r-protein L4, which also has a second activity as a transcriptional attenuator (Lindahl et al., 1983). In L4, a similar ribosome-binding loop was also found dispensable for attenuation (Worbs et al., 2000; Zengel et al., 2003). The findings show that evolution made economic use of r-proteins by diverting regions not under strict selection by ribosomal functions to other purposes.

The S10 loop is of obvious architectural importance for the 30S subunit (Figure 2.5; (Schluzen et al., 2000; Wimberly et al., 2000)), consistent with the observation that the loop is essential for cell viability (Figure 4.2). These results suggest that the transcription antitermination activity of S10 is independent of ribosomes or ribosome-bound S10, in perfect agreement with the finding that S10 cannot bind to NusB and the 30S subunit at the same time. The above results also demonstrate that under normal growth conditions, rRNA

transcription antitermination is not essential for the cells and that proteins involved in antitermination are essential because of their function in other cellular processes, as recently also shown for other antitermination factors (Bubunencko et al., 2007; Phadtare et al., 2007).

5.2 S10 is adapted to different functional contexts without global structural remodeling

S10 has been suggested to represent a largely intrinsically unstructured protein, whose structure could adapt to different functional contexts (Gopal et al., 2001). S10 expressed alone exhibits low solubility and tends to aggregate. NusB confers increased solubility on S10 (Figure 4.1A), suggesting that S10 may preferentially exist in complex with NusB off the ribosome. I observed that the bulk of S10 adopts the same global fold in complex with NusB as in the ribosome. Thus, the results presented here rigorously exclude the possibility that the structure of S10 is extensively remodeled in order to recruit the protein as a transcription factor. Indeed, the long ribosome-binding loop is most likely the only intrinsically unfolded region of S10.

5.3 Mutually exclusive binding of S10 to the 30S subunit or NusB may provide for feedback control of ribosome biogenesis

Balancing the levels of ribosomal building blocks is critical for bacteria, since ribosome biosynthesis can consume half of the available metabolic energy (Bremer and Dennis, 1987). A number of negative feedback loops have been characterized that act to ensure stoichiometric levels of ribosomal constituents. When expressed in surplus of their rRNA binding sites, several r-proteins restrict their own expression and that of other proteins in their operons by binding to their own mRNAs, thereby sequestering the messages from translation (Lindahl and Zengel, 1986). In addition to such translational feedback, r-protein L4 also down-regulates transcription of its operon (Lindahl et al., 1983). Evidence presented here suggests that such negative feedback may be complemented by positive feedback through r-protein S10 (Figure 5.1). Since crystal structures show that S10 cannot participate in transcription antitermination on RNA polymerase and translation on the ribosome at the same time, only S10 produced in excess of ribosomes will elicit antitermination of *rrn* operons and thus a higher rate of rRNA biosynthesis. As a consequence, surplus S10 would act to increase the rRNA level. With respect to rRNA production, NusB and BoxA may therefore be envisioned as enhancers of an S10-based feedback regulation.

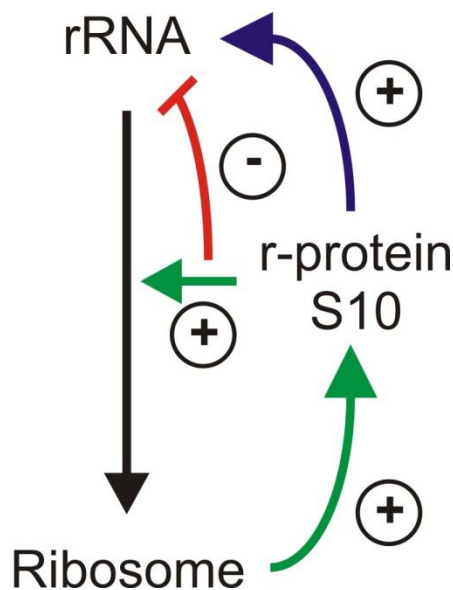


Figure 5.1 A feedback control circuit by S10

S10 cannot participate in rRNA transcription and ribosome assembly at the same time (shown by the red curve and '-'). S10 functions as an architectural element for ribosome assembly when it expresses on the level of fulfilling available ribosomal binding sites (shown by green arrows and '+'). Only S10 produced in the excess of available ribosome binding sites up-regulates the antitermination modular for rRNA transcription (shown by the blue arrow and '+'). Black arrow, transcription is coupled with translation.

5.4 S10 and NusB form a functional module for recognition of BoxA

Much of the proteome is organized as functional modules (Hartwell et al., 1999), which support an autonomous function as, for example, devices within a macromolecular machine. Here it is shown how the NusB-S10 complex acts as a functional RNA-binding module of the transcriptional machineries with which it is associated. Both subunits of the NusB-S10 complex contribute to a mosaic yet contiguous binding surface for a crucial RNA signaling element, BoxA. An analogous situation was encountered intramolecularly in NusA, in which different NusA RNA-binding domains come together to create an enlarged, composite RNA-binding site (Beuth et al., 2005; Worbs et al., 2001). Cooperation between two or more subunits to generate a composite binding surface for an additional factor is an important architectural principle in macromolecular assemblies (Liu et al., 2007).

5.5 S10 is the active antitermination factor of the NusB-S10 complex and NusB serves as an adaptor in the transcription process

In the traditional view of processive antitermination, S10 serves as an auxiliary factor that recruits the antitermination factor NusB to RNAP (Mason and Greenblatt, 1991; Mason et al.,

1992). Contrary to that view, work presented here shows that S10 supports N-antitermination and Nun-termination even in the absence of NusB. According to results presented here, NusB has supportive functions, while the fundamental antitermination activity of the complex relies on S10. What are the supportive functions of NusB and what constitutes the fundamental antitermination activity of S10?

S10 is a truly multi-functional protein even within the transcriptional complexes. Apart from interacting with NusB ((Mason et al., 1992) and this work), it directly binds RNAP (Mason and Greenblatt, 1991). Furthermore, S10 has been suggested to contact phage λ protein N (Mogridge et al., 1995, 1998b). My work shows that S10 is even more versatile and also binds to the BoxA mRNA element. Thus, S10 constitutes a hub within the N- and *rrn* antitermination and Nun-termination complexes, through which the functions of other factors may be integrated. However, isolated S10 binds RNA with low specificity (Greive et al., 2005). Under physiological conditions, positioning of S10 on the mRNA by RNAP or the phage proteins is presumably inefficient. NusB is therefore required as an adaptor that ensures efficient loading of S10 on the mRNA at BoxA and subsequent contact with RNA polymerase.

5.6 Hypothesis: NusB may deliver S10 into ribosomes

NusB and S10 together form a functional module in the transcription antitermination, where NusB serves as a loading factor that ensures efficient entry of S10 into the transcription complexes. In addition, NusB can deliver S10 into the other molecular environment *in vitro*, e.g. into crystals. Thus, I speculated that in translation NusB may still function as a loading factor that delivers S10 into ribosomes. Namely, S10 would not be detected in assembled ribosome from an *E. coli* strain bearing the absence of NusB. However, ribosomes binding assay showed little difference between the content of S10 in ribosomes from wt strain and that of S10 in ribosomes from *nusB* gene knockout strain (Figure 4.12). Little difference in contents of S10 might be due to the fact that fully assembled ribosomes from which S10 was obtained may yield the high background for Western blot detection. As S10 is a crucial protein involved in the 30S ribosomal subunit assembly, further investigations have to be made to monitor S10 from the 30S ribosomal subunit in order to get rid of effects from fully assembled ribosomes.

The hypothesis that NusB may deliver S10 into ribosomes in translation must be considered whether there are reasonable amounts of NusB so that it could carry out two tasks

(loading S10 into both transcription antitermination complex and ribosomes). A minimum of ~2000 S10 molecules, in which there are ~36 *rrn* operons per cell and ~55 RNAP molecules per operon, would be required for transcription under the fast growth conditions in the cell (Squires and Zaporozets, 2000). There must be one S10 molecule per ribosome and there could be $\leq 70,000$ S10 molecules involved in ribosome assembly under same fast growth conditions in the cell (Squires and Zaporozets, 2000). It is estimated that about 3000-6000 NusB molecules are present depending on the growth conditions in the cell (Swindle et al., 1988). Therefore, this is a quantity that might be expected for NusB to load ~2000 S10 molecules into transcription antitermination complex. In the meantime the remaining (~1000-3000) NusB molecules are apparently not enough to make stoichiometric binding with 70,000 S10 molecules involved in ribosome assembly and subsequently, cannot load all S10 molecules into ribosomes at one time. Thus, I assume that the loading of S10 into ribosomes by NusB is a repetitive process. This repetitive loading process regulates the rate of ribosome formation in order to meet the cell's need for protein synthesis capacity. This assumption is in an agreement with suggestions by Squires et al that the role NusB plays in translation is likely to be transient rather than the permanent and stoichiometric role that the r-proteins play in ribosome assembly and function (Squires and Zaporozets, 2000).

6 References

- Altieri, A.S., Mazzulla, M.J., Horita, D.A., Coats, R.H., Wingfield, P.T., Das, A., Court, D.L., and Byrd, R.A. (2000). The structure of the transcriptional antiterminator NusB from *Escherichia coli*. *Nat Struct Biol* 7, 470-474.
- Ban, N., Nissen, P., Hansen, J., Moore, P.B., and Steitz, T.A. (2000). The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* 289, 905-920.
- Belogurov, G.A., Vassilyeva, M.N., Svetlov, V., Klyuyev, S., Grishin, N.V., Vassilyev, D.G., and Artsimovitch, I. (2007). Structural basis for converting a general transcription factor into an operon-specific virulence regulator. *Mol. Cell* 26, 117-129.
- 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.
- Beuth, B., Pennell, S., Arnvig, K.B., Martin, S.R., and Taylor, I.A. (2005). Structure of a *Mycobacterium tuberculosis* NusA-RNA complex. *EMBO J.* 24, 3576-3587.
- Blum, H., and Beier, H. (1987). Improved silver staining of plant proteins, RNA and DNA polyacrylamid gels. *Electrophoresis* 8, 93-99.
- Bonin, I., Muhlberger, R., Bourenkov, G.P., Huber, R., Bacher, A., Richter, G., and Wahl, M.C. (2004a). 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.
- Bonin, I., Robelek, R., Benecke, H., Urlaub, H., Bacher, A., Richter, G., and Wahl, M.C. (2004b). Crystal structures of the antitermination factor NusB from *Thermotoga maritima* and implications for RNA binding. *Biochem. J.* 383, 419-428.
- Bremer, H., and Dennis, P.P. (1987). Modulation of chemical composition and other parameters of the cell by growth rate. In *Escherichia coli* and *Salmonella thyphimurium*: Cellular and molecular biology, F.C. Neidhardt, J.L. Ingraham, K.B. Low, B. Magasanik, M. Schaechter, and H.E. Umbarger, eds. (Washington, DC, USA: American Society for Microbiology), pp. 1527-1542.
- Brock, S., Szkaradkiewicz, K., and Sprinzl, M. (1998). Initiation factors of protein biosynthesis in bacteria and their structural relationship to elongation and termination factors. *Mol Microbiol* 29, 409-417.
- Broennimann, C., Eikenberry, E.F., Henrich, B., Horisberger, R., Huelsen, G., Pohl, E., Schmitt, B., Schulze-Briese, C., Suzuki, M., Tomizaki, T., *et al.* (2006). The PILATUS 1M detector. *J. Synchr. Radiat.* 13, 120-130.
- Bubunencko, M., Baker, T., and Court, D.L. (2007). Essentiality of ribosomal and transcription antitermination proteins analyzed by systematic gene replacement in *Escherichia coli*. *J Bacteriol* 189, 2844-2853.

- Burova, E., Hung, S.C., Sagitov, V., Stitt, B.L., and Gottesman, M.E. (1995). Escherichia coli NusG protein stimulates transcription elongation rates in vivo and in vitro. *J Bacteriol* 177, 1388-1392.
- CCP4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50, 760-763.
- Ciampi, M.S. (2006). Rho-dependent terminators and transcription termination. *Microbiology* 152, 2515-2528.
- Court, D.L., Patterson, T.A., Baker, T., Costantino, N., Mao, X., and Friedman, D.I. (1995). Structural and functional analyses of the transcription-translation proteins NusB and NusE. *J Bacteriol* 177, 2589-2591.
- 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.
- Das, A., Ghosh, B., Barik, S., and Wolska, K. (1985). Evidence that ribosomal protein S10 itself is a cellular component necessary for transcription antitermination by phage lambda N protein. *Proc. Natl. Acad. Sci. USA* 82, 4070-4074.
- Das, R., Loss, S., Li, J., Waugh, D.S., Tarasov, S., Wingfield, P.T., Byrd, R.A., and Altieri, A.S. (2008). Structural biophysics of the NusB:NusE antitermination complex. *J. Mol. Biol.* 376, 705-720.
- Davis, I.W., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2004). MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res.* 32, W615-619.
- Dell, V.A., Miller, D.L., and Johnson, A.E. (1990). Effects of nucleotide- and aurodox-induced changes in elongation factor Tu conformation upon its interactions with aminoacyl transfer RNA. A fluorescence study. *Biochemistry* 29, 1757-1763.
- Drenth, J. (1994). *Principles of Protein X-ray Crystallography* (Springer-Verlag).
- Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* 60, 2126-2132.
- Fischer, G., Bang, H., Berger, E., and Schellenberger, A. (1984). Conformational specificity of chymotrypsin toward proline-containing substrates. *Biochim Biophys Acta* 791, 87-97.
- Friedman, D.I., Baumann, M., and Baron, L.S. (1976). Cooperative effects of bacterial mutations affecting lambda N gene expression. I. Isolation and characterization of a nusB mutant. *Virology* 73, 119-127.
- Friedman, D.I., and Court, D.L. (1995). Transcription antitermination: the lambda paradigm updated. *Mol Microbiol* 18, 191-200.

- Friedman, D.I., and Gottesman, M.E. (1983). Lytic mode of lambda development. In *Lambda II*, R.W. Hendrix, J.W. Roberts, F.W. Stahl, and R.A. Weisberg, eds. (Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press), pp. 21-51.
- Friedman, D.I., Schauer, A.T., Baumann, M.R., Baron, L.S., and Adhya, S.L. (1981). Evidence that ribosomal protein S10 participates in control of transcription termination. *Proc Natl Acad Sci U S A* 78, 1115-1118.
- Gopal, B., Haire, L.F., Cox, R.A., Jo Colston, M., Major, S., Brannigan, J.A., Smerdon, S.J., and Dodson, G. (2000). The crystal structure of NusB from *Mycobacterium tuberculosis*. *Nat. Struct. Biol.* 7, 475-478.
- Gopal, B., Papavinasasundaram, K.G., Dodson, G., Colston, M.J., Major, S.A., and Lane, A.N. (2001). Spectroscopic and thermodynamic characterization of the transcription antitermination factor NusE and its interaction with NusB from *Mycobacterium tuberculosis*. *Biochemistry* 40, 920-928.
- Greive, S.J., Lins, A.F., and von Hippel, P.H. (2005). Assembly of an RNA-protein complex. Binding of NusB and NusE (S10) proteins to boxA RNA nucleates the formation of the antitermination complex involved in controlling rRNA transcription in *Escherichia coli*. *J. Biol. Chem.* 280, 36397-36408.
- Gribskov, M. (1992). Translational initiation factors IF-1 and eIF-2 alpha share an RNA-binding motif with prokaryotic ribosomal protein S1 and polynucleotide phosphorylase. *Gene* 119, 107-111.
- Gualerzi, C.O., and Pon, C.L. (1990). Initiation of mRNA translation in prokaryotes. *Biochemistry* 29, 5881-5889.
- Gusarov, I., and Nudler, E. (1999). The mechanism of intrinsic transcription termination. *Mol Cell* 3, 495-504.
- Hartwell, L.H., Hopfield, J.J., Leibler, S., and Murray, A.W. (1999). From molecular to modular cell biology. *Nature* 402, C47-52.
- Janosi, L., Hara, H., Zhang, S., and Kaji, A. (1996). Ribosome recycling by ribosome recycling factor (RRF)--an important but overlooked step of protein biosynthesis. *Adv Biophys* 32, 121-201.
- Jeffery, C.J. (1999). Moonlighting proteins. *Trends Biochem Sci* 24, 8-11.
- Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.* 26, 795-800.
- Kapanidis, A.N., Margeat, E., Ho, S.O., Kortkhonjia, E., Weiss, S., and Ebright, R.H. (2006). Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science* 314, 1144-1147.

- Karimi, R., Pavlov, M.Y., Buckingham, R.H., and Ehrenberg, M. (1999). Novel roles for classical factors at the interface between translation termination and initiation. *Mol Cell* 3, 601-609.
- Kisselev, L.L., and Buckingham, R.H. (2000). Translational termination comes of age. *Trends Biochem Sci* 25, 561-566.
- Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680-685.
- Lang, K., Schmid, F.X., and Fischer, G. (1987). Catalysis of protein folding by prolyl isomerase. *Nature* 329, 268-270.
- Laskowski, R.A., Moss, D.S., and Thornton, J.M. (1993). Main-chain bond lengths and bond angles in protein structures. *J Mol Biol* 231, 1049-1067.
- Laursen, B.S., Sorensen, H.P., Mortensen, K.K., and Sperling-Petersen, H.U. (2005). Initiation of protein synthesis in bacteria. *Microbiol Mol Biol Rev* 69, 101-123.
- Legault, P., Li, J., Mogridge, J., Kay, L.E., and Greenblatt, J. (1998). NMR structure of the bacteriophage lambda N peptide/boxB RNA complex: recognition of a GNRA fold by an arginine-rich motif. *Cell* 93, 289-299.
- 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, J., Mason, S.W., and Greenblatt, J. (1993). Elongation factor NusG interacts with termination factor rho to regulate termination and antitermination of transcription. *Genes Dev* 7, 161-172.
- Li, S.C., Squires, C.L., and Squires, C. (1984). Antitermination of E. coli rRNA transcription is caused by a control region segment containing lambda nut-like sequences. *Cell* 38, 851-860.
- Lindhahl, L., Archer, R., and Zengel, J.M. (1983). Transcription of the S10 ribosomal protein operon is regulated by an attenuator in the leader. *Cell* 33, 241-248.
- Lindhahl, L., and Zengel, J.M. (1986). Ribosomal genes in Escherichia coli. *Annu Rev Genet* 20, 297-326.
- Lingel, A., Simon, B., Izaurralde, E., and Sattler, M. (2003). Structure and nucleic-acid binding of the Drosophila Argonaute 2 PAZ domain. *Nature* 426, 465-469.
- Liu, S., Li, P., Dybkov, O., Nottrott, S., Hartmuth, K., Luhrmann, R., Carlomagno, T., and Wahl, M.C. (2007). Binding of the human Prp31 Nop domain to a composite RNA-protein platform in U4 snRNP. *Science* 316, 115-120.
- Luttgen, H., Robelek, R., Muhlberger, R., Diercks, T., Schuster, S.C., Kohler, P., Kessler, H., Bacher, A., and Richter, G. (2002). Transcriptional regulation by antitermination.

- Interaction of RNA with NusB protein and NusB/NusE protein complex of *Escherichia coli*. *J. Mol. Biol.* *316*, 875-885.
- 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.
- 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.
- Mason, S.W., Li, J., and Greenblatt, J. (1992). Direct interaction between two *Escherichia coli* transcription antitermination factors, NusB and ribosomal protein S10. *J Mol Biol* *223*, 55-66.
- Mizushima, S., and Nomura, M. (1970). Assembly mapping of 30S ribosomal proteins from *E. coli*. *Nature* *226*, 1214.
- 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., Schweimer, K., Rosch, P., Gottesman, M., and Landick, R. (2009). Two structurally independent domains of *E. coli* NusG create regulatory plasticity via distinct interactions with RNA polymerase and regulators. *J Mol Biol* *391*, 341-358.
- Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* *53*, 240-255.
- Nierhaus, K.H. (1991). The assembly of prokaryotic ribosomes. *Biochimie* *73*, 739-755.

- Nissen, P., Hansen, J., Ban, N., Moore, P.B., and Steitz, T.A. (2000). The structural basis of ribosome activity in peptide bond synthesis. *Science* 289, 920-930.
- 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.
- Noller, H.F., Hoffarth, V., and Zimniak, L. (1992). Unusual resistance of peptidyl transferase to protein extraction procedures. *Science* 256, 1416-1419.
- 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.
- Phadtare, S., Kazakov, T., Bubunenko, M., Court, D.L., Pestova, T., and Severinov, K. (2007). Transcription antitermination by translation initiation factor IF1. *J. Bacteriol.* 189, 4087-4093.
- Piotukh, K., Gu, W., Kofler, M., Labudde, D., Helms, V., and Freund, C. (2005). Cyclophilin A binds to linear peptide motifs containing a consensus that is present in many human proteins. *J Biol Chem* 280, 23668-23674.
- 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.
- Ramakrishnan, V. (2002). Ribosome structure and the mechanism of translation. *Cell* 108, 557-572.
- Rhodes, G. (2006). *Crystallography Made Crystal Clear* (Elsevier Academic Press, third edition).
- 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.W., Shankar, S., and Filter, J.J. (2008). RNA polymerase elongation factors. *Annu Rev Microbiol* 62, 211-233.
- Robledo, R., Atkinson, B.L., and Gottesman, M.E. (1991). *Escherichia coli* mutations that block transcription termination by phage HK022 Nun protein. *Journal of Molecular Biology* 220, 613-619.
- Rodnina, M.V., Savelsbergh, A., Katunin, V.I., and Wintermeyer, W. (1997). Hydrolysis of GTP by elongation factor G drives tRNA movement on the ribosome. *Nature* 385, 37-41.
- Sambrook, J., and Fritsch, E.F. (1989). *Molecular cloning-A laboratory manual*. Cold Spring Harbour, New York, Cold Spring Harbour Laboratory Press.

- Schluzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bartels, H., Agmon, I., Franceschi, F., and Yonath, A. (2000). Structure of functionally activated small ribosomal subunit at 3.3 angstroms resolution. *Cell* *102*, 615-623.
- Schmeing, T.M., and Ramakrishnan, V. (2009). What recent ribosome structures have revealed about the mechanism of translation. *Nature* *461*, 1234-1242.
- Schuwirth, B.S., Borovinskaya, M.A., Hau, C.W., Zhang, W., Vila-Sanjurjo, A., Holton, J.M., and Cate, J.H. (2005). Structures of the bacterial ribosome at 3.5 Å resolution. *Science* *310*, 827-834.
- Selmer, M., Dunham, C.M., Murphy, F.V.t., Weixlbaumer, A., Petry, S., Kelley, A.C., Weir, J.R., and Ramakrishnan, V. (2006). Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* *313*, 1935-1942.
- Shiba, K., Ito, K., and Yura, T. (1986). Suppressors of the secY24 mutation: identification and characterization of additional ssy genes in Escherichia coli. *J Bacteriol* *166*, 849-856.
- Skordalakes, E., and Berger, J.M. (2003). Structure of the Rho transcription terminator: mechanism of mRNA recognition and helicase loading. *Cell* *114*, 135-146.
- Squires, C.L., Greenblatt, J., Li, J., and Condon, C. (1993). Ribosomal RNA antitermination in vitro: requirement for Nus factors and one or more unidentified cellular components. *Proc Natl Acad Sci U S A* *90*, 970-974.
- Squires, C.L., and Zaporozets, D. (2000). Proteins shared by the transcription and translation machines. *Annu Rev Microbiol* *54*, 775-798.
- Steiner, T., Kaiser, J.T., Marinkovic, S., Huber, R., and Wahl, M.C. (2002). Crystal structures of transcription factor NusG in light of its nucleic acid- and protein-binding activities. *EMBO J* *21*, 4641-4653.
- Studier, F.W. (2005). Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* *41*, 207-234.
- Sullivan, S.L., and Gottesman, M.E. (1992). Requirement for E. coli NusG protein in factor-dependent transcription termination. *Cell* *68*, 989-994.
- Swindle, J., Zylicz, M., Georgopoulos, C., Li, J., and Greenblatt, J. (1988). Purification and properties of the NusB protein of Escherichia coli. *J Biol Chem* *263*, 10229-10235.
- Taura, T., Ueguchi, C., Shiba, K., and Ito, K. (1992). Insertional disruption of the nusB (ssyB) gene leads to cold-sensitive growth of Escherichia coli and suppression of the secY24 mutation. *Mol Gen Genet* *234*, 429-432.
- Torres, M., Balada, J.M., Zellars, M., Squires, C., and Squires, C.L. (2004). In vivo effect of NusB and NusG on rRNA transcription antitermination. *J Bacteriol* *186*, 1304-1310.
- Torres, M., Condon, C., Balada, J.M., Squires, C., and Squires, C.L. (2001). Ribosomal protein S4 is a transcription factor with properties remarkably similar to NusA, a protein

- involved in both non-ribosomal and ribosomal RNA antitermination. *EMBO J* 20, 3811-3820.
- Vassylyev, D.G., Vassylyeva, M.N., Zhang, J., Palangat, M., Artsimovitch, I., and Landick, R. (2007). Structural basis for substrate loading in bacterial RNA polymerase. *Nature* 448, 163-168.
- Ward, D.F., DeLong, A., and Gottesman, M.E. (1983). *Escherichia coli* nusB mutations that suppress nusA1 exhibit lambda N specificity. *J Mol Biol* 168, 73-85.
- Weisberg, R.A., and Gottesman, M.E. (1999). Processive antitermination. *J. Bacteriol.* 181, 359-367.
- Whalen, W.A., and Das, A. (1990). Action of an RNA site at a distance: role of the nut genetic signal in transcription antitermination by phage-lambda N gene product. *New Biologist* 2, 975-991.
- Wimberly, B.T., Brodersen, D.E., Clemons, W.M., Morgan-Warren, R.J., Carter, A.P., Vornrhein, C., Hartsch, T., and Ramakrishnan, V. (2000). Structure of the 30S ribosomal subunit. *Nature* 407, 327-339.
- Winn, M.D., Isupov, M.N., and Murshudov, G.N. (2001). Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr. D* 57, 122-133.
- Wong, I., and Lohman, T.M. (1993). A double-filter method for nitrocellulose-filter binding: application to protein-nucleic acid interactions. *Proc. Natl. Acad. Sci. USA* 90, 5428-5432.
- 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.
- Worbs, M., Huber, R., and Wahl, M.C. (2000). Crystal structure of ribosomal protein L4 shows RNA-binding sites for ribosome incorporation and feedback control of the S10 operon. *Embo J* 19, 807-818.
- Worbs, M., Wahl, M.C., Lindahl, L., and Zengel, J.M. (2002). Comparative anatomy of a regulatory ribosomal protein. *Biochimie* 84, 731-743.
- Yusupova, G.Z., Yusupov, M.M., Cate, J.H., and Noller, H.F. (2001). The path of messenger RNA through the ribosome. *Cell* 106, 233-241.
- Zellars, M., and Squires, C.L. (1999). Antiterminator-dependent modulation of transcription elongation rates by NusB and NusG. *Molecular Microbiology* 32, 1296-1304.
- Zengel, J.M., Jerauld, A., Walker, A., Wahl, M.C., and Lindahl, L. (2003). The extended loops of ribosomal proteins L4 and L22 are not required for ribosome assembly or L4-mediated autogenous control. *RNA* 9, 1188-1197.

Zengel, J.M., and Lindahl, L. (1992). Ribosomal protein L4 and transcription factor NusA have separable roles in mediating terminating of transcription within the leader of the S10 operon of *Escherichia coli*. *Genes Dev* 6, 2655-2662.

Zengel, J.M., and Lindahl, L. (1994). Diverse mechanisms for regulating ribosomal protein synthesis in *Escherichia coli*. *Prog Nucleic Acid Res Mol Biol* 47, 331-370.

7 Appendixes

7.1 Principles of protein X-ray crystallography

X-ray crystallography is a principle technique in the determination of protein structures. Its process involves the crystal growth, data collection and processing, solutions of the phase problem and fitting, refinement and validation of crystal structures. Here, the basic principles of protein X-ray crystallography are briefly described. The following sections are compiled from crystallography textbooks (Drenth, 1994; Rhodes, 2006).

7.1.1 Crystal growth

The initial step in protein crystallography is the production of protein crystals. To form a crystal the pure protein solution at a concentration between 0.5 and 200 mg/ml is mixed with reagents to decrease the protein solubility close to the precipitation point. Protein-solvent interactions are disturbed by these reagents, which results in the formation of nucleation sites, to allow protein molecules to assemble into a periodic lattice from supersaturated solutions. Crystal growth is then followed by expansion and cessation when the crystal reaches a certain size.

The most commonly used experimental method to form crystals from protein solution is vapor diffusion by sitting drop or hanging drop techniques. The common format involves setting up a droplet containing equal amounts of protein solution and precipitant solution in a sealed chamber. The droplet is equilibrated against the precipitant solution in the reservoir of chamber. Due to the mixture of protein solution and precipitant solution, the precipitant concentration in the droplet is lower than that in the reservoir. Thus, the water molecules leave the droplet and dissolve in the reservoir by the evaporation in order to achieve the equilibrium. The equilibration over time in this seal environment leads to supersaturating concentrations that allow protein crystallization in the droplet.

7.1.2 Data collection and processing

Crystals are exposed under X-rays for diffraction data collection. Crystals are frequently flash-cooled to liquid nitrogen temperature (~100 K) for reducing thermal vibrations of atoms in crystals, and avoiding the radiation damage of crystals in order to collect complete datasets from a single crystal. Nowadays synchrotron is used as the main source of X-rays that produces high intensity X-ray radiation and allows selection of radiation with wavelengths in

a wide range. The speed and quality of structure solution have improved dramatically compared with conventional X-ray sources, e.g. the sealing tube and the rotating anode. When the flash-cooled crystal held within a loop is mounted in a goniometer in the path of X-ray beam, X-rays are scattered into many discrete beams by the electrons in molecules that form a crystal lattice. Scattered X-rays are observed (diffracted) when the angle (θ) of incidence at lattice planes is equal to the angle (θ) of scattering and the path length difference is equal to an integer number (n) of wavelengths (λ) according to Bragg's law:

$$n \lambda = 2d \sin\theta$$

in which d is the spacing between lattice planes. Interferences among scattered X-rays generated at lattice points in parallel planes produce distinct spots that can be recorded on a detector to yield the diffraction pattern. Each spot represents a reflection, which is a sum of individual scattering of all of the electrons in the unit cell along a particular direction. The sum that describes diffracted X-rays at position hkl (reciprocal space coordinates) is called structure factor $F(h, k, l)$, which is a function of the electron density distribution in the unit cell. Therefore, the structure factor is a vector defined by intensity F_{hkl} and phase α_{hkl} . The goal of crystallography is to calculate the electron density ρ at every position x, y, z in the unit cell. This can be done by Fourier transform (FT). The FT is a transition between two different, but equivalent ways of describing an object or a process. The structure factor $F(h, k, l)$ is the Fourier transform of $\rho(x, y, z)$ but the reverse is also true: $\rho(x, y, z)$ is the Fourier transform of $F(h, k, l)$ and therefore, $\rho(x, y, z)$ can be written as a function of all $F(h, k, l)$:

$$\rho(xyz) = 1/V \sum_h \sum_k \sum_l |F(hkl)| \exp[-2\pi i(hx + ky + lz) + i\alpha(hkl)]$$

where V is the total volume of the unit cell and i is the contribution of each atom.

The data processing from recorded spots yields a list of reflections (positions) and their intensities. The intensity of the diffracted X-rays is proportional to the square of the amplitudes (I_{hkl}), which could be measured from diffraction pattern. It now seems easy to calculate the electron density $\rho(x, y, z)$ at every position (x, y, z) in the unit cell. However, there is a problem. The phase angle (α_{hkl}) cannot be obtained directly from the diffraction pattern.

7.1.3 Solutions of the phase problem

The importance of phases in producing the correct structure is demonstrated by Kevin Cowtan's FT model of a duck and of a cat: the electron density map derived by combining amplitudes for the duck diffraction and phases from the cat diffraction leads to a cat (<http://www.ysbl.york.ac.uk/~cowtan/fourier/magic.html>). To solve the phase problem several techniques were developed: Direct Method, Molecular Replacement, Single Isomorphous Replacement (SIR), Multiple Isomorphous Replacement (MIR), Single-wavelength Anomalous Dispersion (SAD), Multiple-wavelength Anomalous Dispersion (MAD), and combination of above approaches which gives rise to Single Isomorphous Replacement Anomalous Scattering (SIRAS) and Multiple Isomorphous Replacement Anomalous Scattering (MIRAS). One can refer to any crystallography textbooks for the theories behind these methods. As molecular replacement method was applied to my work, the method is briefly explained below.

Molecular Replacement (MR) can be useful to deduce the phase if a homology model is available. As a rule of thumb, a sequence identity > 35 % is normally required between a homology protein and the unknown protein, or the two proteins are expected to have a very similar fold of the polypeptide chain (rmsd of α C atoms < 2.0 Å). Placement of the homology protein in the target unit cell requires its proper orientation and precise position that involves two steps: rotation and translation. In the rotation step the spatial orientation of the known and unknown protein with respect to each other is determined while in the next step the translation needed to superimpose the now correctly oriented protein onto the other protein is calculated. The basic principle of the MR can be understood by regarding the Patterson function of a protein crystal structure. The Patterson function $P(u, v, w)$ is a Fourier summation with intensities as coefficients and without phase angles:

$$P(uvw) = 1/V \sum_h \sum_k \sum_l |F(hkl)|^2 \cos[2\pi(hu + kv + lw)]$$

u, v, w are relative coordinates in the unit cell.

The Patterson map generated by Patterson function is a vector map: vectors between atoms in the real structure show up as vectors from the origin to maxima in the Patterson map. If the pairs of atoms belong to the same molecule, then the corresponding vectors are relatively short and their end-points are found not too far from the origin in the Patterson map; they are called self-Patterson vectors (intramolecular vectors), which can provide us

with the rotational relationship between the known and the unknown structures. In the rotation step the intramolecular vectors for the known molecule are calculated in a P_1 unit cell. These calculated intramolecular vectors are rotated in accordance with an Eulerian angle system until they match the observed Patterson functions from the unknown molecule. For the final solution of the MR method the translation required to overlap one molecule onto the other in real space must be determined, after it has been oriented in the correct way with the rotation function. The translation function is calculated that gives the correlation between a set of cross-Patterson vectors (intermolecular vectors) for a model structure and the observed Patterson function. Intermolecular vectors mean vectors in the Patterson map derived from vectors between atoms in two molecules in the model structure related by a crystallographic symmetry operation. With the translation function one can determine the position of molecule 1 with the respect to the symmetry related molecule 2, and subsequently for any other pairs of symmetry related molecules. When the correct position is located, the phases of the model in this position can be used to deduce the phases for unknown protein.

7.1.4 Fitting, refinement and validation of crystal structures

From molecular replacement an approximate model of the protein structure can be obtained in which the broad features of the molecular architecture are apparent. To adjust an initial model such that the best possible agreement with electron density map is achieved while maintaining a reasonable stereochemistry, an iterative model building and refinement are carried out. The calculated structural factor (F_{cal}) after each cycle of refinement is compared with the observed structural factor (F_{obs}) to yield an R factor, which is one of quality assessment factors of the final structure. Refinement is the process of adjusting the model to find a closer agreement between F_{cal} and F_{obs} . Several methods have been developed and, if applied, they lower the R factor substantially, reaching values in the 10 to 20 % range or even lower. The adjustment of the model consists of changing the positional parameters (x, y, z) and the temperature factors (B factors) for all atoms in the structure. B factor generated after the refinement is used to judge the mobility of the structure within the crystal. Attention should be given to residues or parts of residues with conspicuously high B factor values. It has been shown that the R factor can reach surprisingly low values in the refinement of protein structural models that appear later to be incorrect, for instance, because the number of model parameters is taken too high. Therefore, R_{free} factor was suggested to improve this situation. In this method reflections are divided into a test set and a working set. The test set is a random selection of 5 % of the observed reflections. The refinement is carried out with

the working set only, and the R_{free} factor is calculated with the test set only. R_{free} factor is unbiased by the refinement process and therefore, it reflects the accuracy of the structural model. In addition, the accuracy of structural model can be estimated by other methods, such as Ramachandran Plot. The result of protein structure determination is the generation of a file that lists x, y and z coordinates for all atoms present in the crystal.

7.2 Abbreviations

°C	Degree celsius
3D	3 Dimensional
Å	Angstrom (1Å = 10 ⁻¹⁰ m)
AA/aa	Amino acid
<i>aae</i>	<i>Aquifex aeolicus</i>
AR	Acidic repeats
ATP	Adenosine-5'-triphosphate
BLAST	Basic local alignment search tool
bp	Base pair
CC	Correlation coefficient
CCD	Charged coupled device
CTD	C-terminal domain
ddH ₂ O	Double distilled water
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
<i>E. coli/eco</i>	<i>Escherichia coli</i>
EDTA	Ethylene-diamine-tetraacidic acid
EF	Elongation factor
EF-G	Elongation factor G
EF-Ts	Elongation factor Ts
EF-Tu	Elongation factor Tu
F	Structure factor
FOM	Figure of merit
GTP	Guanosine-5'-triphosphate
h	Hour
HEPES	N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid
I	Intensity
IF	Initiation factor
IPTG	Isopropyl-β-D-thiogalactopyranoside
K	Kelvin

kb	Kilo base
kD	Kilo dalton
KH	K-homology
KOW	Kyprides, Ouzounis, Woese
l	Liter
LB	Luria bertani medium
M	Molarity
MAD	Multi-wavelength anomalous dispersion
mAU	milli-Absorption unit
MES	2-(N-morpholino)-ethanesulfonic acid
min	Minute
MIR	Multiple isomorphous replacement
MPD	2-Methyl 2, 4-pentanediol
MR	Molecular replacement
mRNA	messenger RNA
Ni-NTA	Nickel-Nitrilotriacetate
nm	Nanometer
NMR	Nuclear magnetic resonance
NTD	N-terminal domain
Nus	N-utilization substance
Nut	N-utilization site
OD	Optical density
P	Phosphate
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PDB	Protein data bank
PEG	Polyethylene glycol
RF	Release factor
rmsd	Root mean square deviation
RNA	Ribonucleic acid
RNAP	RNA polymerase
rpm	Revolutions per minute
r-protein	Ribosomal protein
RRF	Ribosome recycling factor

RRM	RNA recognition motif
rRNA/rrn	Ribosomal RNA
RT	Room temperature
s	Second
S	Svedberg
SAD	Single-wavelength anomalous diffraction
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SeMet	Selenomethionine
SIR	Single isomorphous replacement
SIRAS	Single isomorphous replacement anomalous scattering
SLS	Swiss light source
Taq	<i>Thermus aquaticus</i>
TEV	Tobacco etch virus protease
<i>T. mar/tma</i>	<i>Thermotoga maritima</i>
Tris	Tris-(hydroxymethyl) aminomethane
tRNA	transfer RNA
<i>T. th/tth</i>	<i>Thermus thermophilus</i>
U	Unit
UV	Ultraviolet
V	Volume
x g	Times gravity

Nucleic acid bases

Adenine	A
Cytosine	C
Guanine	G
Thymine	T
Uracil	U

Amino acids

Alanine	A	Ala
Arginine	R	Arg
Asparagine	N	Asn
Aspartic acid	D	Asp

Cysteine	C	Cys
Glutamine	Q	Gln
Glutamic acid	E	Glu
Glycine	G	Gly
Histidine	H	His
Isoleucine	I	Ile
Leucine	L	Leu
Lysine	K	Lys
Methionine	M	Met
Phenylalanine	F	Phe
Proline	P	Pro
Serine	S	Ser
Threonine	T	Thr
Tryptophan	W	Trp
Tyrosine	Y	Tyr
Valine	V	Val

7.3 Acknowledgements

I am greatly indebted to my supervisor, Prof. Markus Wahl, whose support, encouragement and advice throughout my Ph.D. life are always invaluable. I should like to thank all the former and present members in his group for the great working atmosphere, all their technical help and scientific discussions. Special thanks go to Mrs. Elke Penka for her help on cloning, to Dr. Gert Weber, Dr. Sunbin Liu and Dr. Christian Stegmann for their experimental demonstrations.

The majority of my work was carried out in the Prof. Reinhard Lührman's department. I would like to express my gratitude to him and his group for assistance. Special thanks go to Dr. Olexandr Dybkov for the data analysis of crosslinking assay, to Mrs. Irene Öchsner for her excellent technical support and personal encouragement and to Mrs. Juliane Moses for her help in day to day life.

I do appreciate my thesis committee members, Prof. Ralf Ficner and Dr. Dirk Fasshauer, for their useful suggestion and guidance. Also, I wish to thank GGNB officers for all organizational questions.

I am much obliged to the project collaborators, Prof. Max Gottesman, Prof. Paul Rösch, Dr. Henning Urlaub, Prof. Donald Court and their groups, who shared their results and contributed to this project.

I would like to thank the staff at beamlines PXI/II (SLS) and BL14.2 (BESSY) for support during diffraction data collection.

Lastly, I must thank my family and friends for their continuous support, especially Feifei for her commitment to our private life and her unconditional love.

7.4 Curriculum Vitae

Personal Information:

Name: Xiao Luo
Date of birth: 5th December, 1980
Place of birth: HeiLongJiang. China

Education and Research Experience:

- 10/2006 – Ph.D., GGNB, University of Göttingen, Germany.
(The research was carried out at the Max-Planck Institute for Biophysical Chemistry, Göttingen and at the Free University Berlin, Germany)
Project title: Structural and functional analysis of the NusB-S10 complex shared between transcription and translation
Supervisor: Prof. Markus Wahl
- 09/2005 – 09/2006 M.Re. in Structural Proteomics, University of St-Andrews, UK.
Project I title: CC1, a novel crenarchaeal DNA binding protein
Supervisor: Prof. Malcolm White
Project II title: The structural proteomics of *Sulfolobus solfataricus*
Supervisor: Prof. James Naismith
- 10/1999 – 09/2003 B.Sc. in Biochemistry, University of HeiLongJiang, China.

Publications:

Burmann B. M., Schweimer K., Luo X., Wahl M. C., Stitt B. L., Gottesman M. E. and Rösch P. (2010). A NusE:NusG complex links transcription and translation. *Science*, **328**: 501-504.

Burmann B. M., Luo X., Rösch P., Wahl M. C. and Gottesman M. E. (2010). Fine tuning of the *E. coli* NusB:NusE complex affinity to BoxA RNA is required for processive antitermination. *Nucleic Acids Research*, **38**: 314-326.

Luo X., Hsiao H. H., Bubunenko M., Weber G., Court D. L., Gottesman M. E., Urlaub H. and Wahl M. C. (2008). Structural and functional analysis of the *E. coli* NusB-S10 transcription antitermination complex. *Molecular Cell*, **32**: 791-802.

Luo X., Schwarz-Linek U., Botting H. C., Hensel R., Siebers B. and White M. F. (2007). CC1, a novel crenarchaeal DNA binding protein. *Journal of Bacteriology*, **189**: 403-409.