

**Structural insights into RNA binding by NusA
and interaction studies of Nun with *E. coli* Nus factors**

Dissertation

Zur Erlangung des Doktorgrades
der Fakultät Biologie, Chemie und Geowissenschaften
der Universität Bayreuth

vorgelegt von

M. Sc.

Pagadala Santhanam Sujatha

Bayreuth 2008

Die vorliegende Arbeit wurde von August 2005 bis Juli 2008 am Lehrstuhl für Struktur und Chemie der Biopolymere der Universität Bayreuth unter der Leitung von Prof. Dr. Paul Rösch angefertigt.

Vollständiger Abdruck der von der Fakultät für Biologie, Chemie und Geowissenschaften der Universität Bayreuth genehmigten Dissertation zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften (Dr. rer. Nat.).

Promotionsgesuch eingereicht am :	16.07.2008
Tag des wissenschaftlichen Kolloquiums:	22.10.2008

Erster Gutachter :	Prof. Dr. Paul Rösch
Zweiter Gutachter:	Prof. Dr. Matthias Ullmann
Vorsitzender:	Prof. Dr. Andreas Fery
	Prof. Dr. Rainer Schobert

With fond memories of my mom

Contents

1. Introduction.....	1
1.1 Bacteriophages.....	1
1.2 Life cycle of lambda bacteriophage.....	2
1.3 Transcription mechanism.....	5
1.4 <i>nut</i> – RNA of bacteriophage λ	9
1.5 Transcription antitermination and termination in <i>E. coli</i>	11
1.5.1 Phage λ N antitermination.....	11
1.5.2 Mechanism of Nun mediated termination.....	12
1.6 Elongation factor NusA.....	15
1.7 Thesis objectives.....	21
2. Materials and Methods.....	22
2.1 Culture media.....	22
2.1.1 Luria Bertani medium.....	22
2.1.2 Minimal medium (M9).....	22
2.1.3 P-5052 medium.....	23
2.2 Estimation of protein concentration.....	24
2.3 SDS – polyacrylamide gel electrophoresis.....	24
2.4 Schagger and Jagow gel electrophoresis.....	25
2.5 HK022 Nun protein.....	26
2.5.1 Expression of Nun.....	26
2.5.2 Cell lysis and purification of HK022 Nun.....	26

2.6	Nun carboxy terminal domain (Nun CTD).....	28
2.6.1	Cloning and expression of Nun CTD.....	28
2.6.2	Cell lysis and purification of Nun CTD.....	31
2.7	NusA full length.....	32
2.7.1	Expression of NusA.....	32
2.7.2	Cell lysis and purification of NusA.....	33
2.8	NusA acidic repeat 1 (NusA ar1).....	34
2.8.1	Expression of NusA ar1.....	34
2.8.2	Cell lysis and purification of NusA ar1.....	34
2.9	NusA acidic repeat 2 (NusA ar2).....	35
2.9.1	Expression of NusA ar2.....	35
2.9.2	Cell lysis and purification of NusA ar2.....	36
2.10	S1+KH1+KH2 domain of NusA (SKK).....	37
2.10.1	Expression of SKK.....	37
2.10.2	Cell lysis and purification of SKK domain.....	38
2.11	NusG.....	39
2.11.1	Expression of NusG.....	39
2.11.2	Cell lysis and purification of NusG.....	40
2.12	NusB.....	41
2.12.1	Expression of NusB.....	41
2.12.2	Cell lysis and purification of NusB.....	41
2.13	RNA oligonucleotide.....	42
2.14	NMR spectroscopy.....	43
2.14.1	NMR sample preparation.....	43
2.14.2	NMR spectrometers and measurements.....	44
2.14.3	NMR data processing and analysis.....	44

2.14.4	Experiments necessary for Biomolecular NMR.....	45
2.14.4.1	The HSQC experiment.....	45
2.14.4.2	TROSY experiment.....	46
2.14.4.3	Sequence specific assignments.....	48
2.14.5	Protein-Protein, Protein-RNA interaction studies.....	52
2.14.6	Chemical shift mapping.....	54
2.14.7	Dissociation constant.....	55
3.	Experiments and Results.....	56
3.1	Expression and purification of Nun constructs.....	56
3.1.1	Expression and purification of Nun (1-112).....	56
3.1.2	Expression and purification of Nun (45-112).....	59
3.2	Expression and purification of NusA and its domains.....	61
3.2.1	Overexpression and purification of NusA.....	61
3.2.2	Expression and purification of NusA ar1.....	62
3.2.3	Expression and purification of NusA ar2.....	64
3.2.4	Expression and purification of SKK domain.....	65
3.3	Expression and purification of NusG.....	68
3.4	Expression and purification of NusB.....	69
3.5	Interaction of Nun with <i>E. coli</i> host factors.....	71
3.5.1	Interaction of Nun with NusA ar1 domain.....	71
3.5.2	Interaction of Nun with NusG.....	76
3.5.3	Interaction of Nun with NusB.....	78
3.6	Backbone resonance assignment of SKK domain of NusA.....	79
3.7	Interaction studies of NusA RNA binding domains (SKK).....	85
3.7.1	Binding of SKK with <i>nut</i> RNA.....	85
3.7.2	Normalized chemical shift changes.....	89

3.7.3	Dissociation constant for SKK- λ <i>nutL</i> complex.....	91
3.7.4	Interaction with λ <i>nutL</i> spacer.....	92
3.8	Autoinhibition effect of NusA ar2.....	93
3.8.1	Interaction of NusA ar2 with SKK domain.....	93
3.8.2	Displacement by α -CTD subunit of RNA polymerase.....	98
4.	Discussions.....	100
4.1	Effect of Nus factors on HK022 Nun.....	100
4.2	Backbone assignment of RNA binding domains of NusA (SKK).....	103
4.3	RNA binding studies on SKK domain.....	108
4.4	Mapping of RNA binding interface.....	110
4.5	Regulation of NusA-RNA interaction by autoinhibition.....	112
5.	Summary.....	114
6.	Zusammenfassung.....	116
7.	Abbreviations.....	118
8.	References.....	120
9.	Appendix.....	136
9.1	Script for K_D fitting.....	136
9.2	Nucleotide sequence of HK022Nun.....	137
9.3	Nucleotide sequence of Nun CTD.....	137
9.4	Nucleotide sequence of NusA (1-495).....	138
9.5	Nucleotide sequence of NusA ar1.....	139
9.6	Nucleotide sequence of NusA ar2.....	140
9.7	Nucleotide sequence of SKK domain.....	140
9.8	Nucleotide sequence of NusG.....	141
9.9	Nucleotide sequence of NusB.....	142

9.10 Backbone resonance assignment for SKK domain.....	143
10. Acknowledgements.....	150
11. Erklärung.....	152

1 Introduction

1.1 Bacteriophages

The molecular biology of gene regulation has come to an interesting juncture. On the one hand, the molecular mechanisms used to control gene expression often seem to involve relatively simple ‘adhesive’ interactions, whereas more complex, allosteric processes being less common [Ptashne 2002]. On the other hand, understanding the properties of even small networks of regulatory genes is surprisingly difficult, making analysis of large genome networks daunting. The ‘simple’ bacteriophage is at the forefront of these advances and challenges [Dodd et al., 2005; Roger 2003].

Bacteriophages have played an important role in the development of molecular biology. At the present time, phages are the best organisms for investigation because of their lesser complexity than bacteria and the availability of an enormous number of mutants. Phages have also been useful in the study of replication, transcription and regulation [Freifelder 2001; Hershey et al., 1971; Ptashne 1992].

One reason for the interest in bacteriophages and their relatives is that they recombine genetically with their host, “mingling cellular and viral inheritance” in ways that are fascinating to contemplate and, very likely, of practical importance to humans [Weisberg et al., 1999; Hershey 1967]. In recent years, it has been recognized that bacteriophages have several potential applications in modern biotechnology industry. They have been proposed as delivery vehicles for protein and DNA vaccines, as gene therapy delivery vehicles, as alternatives to antibiotics, for the detection of pathogenic bacteria and as tools for screening libraries of proteins, peptides and antibodies [Jason et al., 2006].

Bacteriophage can persist by itself, but can neither grow nor replicate except within a bacterial cell. Most phages possess genes encoding a variety of proteins. However, all known phages use the ribosomes, protein-synthesizing factors, amino acids, and energy-generating systems of the host cell, and hence a phage can grow only in a metabolizing bacterium. Each

phage must perform some minimal functions for continued survival, like protection of its nucleic acid from environmental chemicals that could alter the molecule, delivery of its nucleic acid to the inside of a bacterium, conversion of an infected bacterium to a phage-producing system which yields a large number of progeny phage and release of progeny phage from an infected bacterium. These functions are carried out in a variety of ways by different phage species. All of the species have certain features in common but differences in detail show the many ways in which specific biological functions can be accomplished [Voet et al., 2nd edition].

Lambda phage

The origins of molecular biology are deeply enmeshed with the discovery and characterization of the temperate coliphage λ [Gottesman 2004]. The isolation of λ was first reported in 1951 by Esther Lederberg and then later it was described in greater detail, in 1953 by Esther and Joshua Lederberg [Lederberg 1951; Lederberg et al., 1953].

The temperate coliphage λ has served as a model for the fields of gene regulation and temporal programming of gene expression [Gottesman 1999]. The λ phage has been useful because, following infection, λ development can proceed along two alternative pathways. Some cells enter a productive cycle, in which phage DNA replicates autonomously and is packaged into progeny phage particles, which are then liberated by lysis. Other cells survive infection to become lysogenic and harbor the phage DNA inserted into the chromosome as a prophage, which remains transcriptionally quiescent for genes of the productive cycle [Campbell 1994].

1.2 Life cycle of lambda bacteriophage

The lysis-lysogeny decision of λ bacteriophage is a paradigm for developmental genetic networks [Oppenheim et al., 2005; Court et al., 2007]. There are three key features which characterize the network. First, after infection of the host bacterium, a decision between lytic or lysogenic development is made that is dependent upon environmental signals and the number of infecting phages per cell. Second, the lysogenic prophage state is very stable. Third, the prophage enters lytic development in response to DNA-damaging agents.

The CI (activates *pRM* promoter) and Cro regulators define the lysogenic and lytic states, respectively, as a bistable genetic switch. Whereas CI maintains a stable lysogenic state, recent studies indicate that Cro sets the lytic course not by directly blocking CI expression but indirectly by lowering levels of CII which activates *cI* transcription (Fig 1.1).

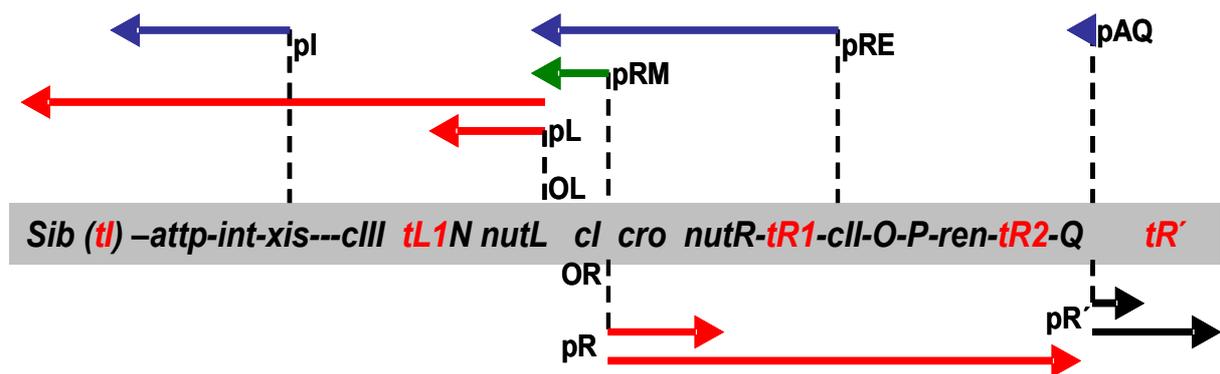


Figure 1.1 Gene and transcription map of λ bacteriophage regulatory region. Genes are shown in the shaded rectangle. The early transcripts for *pL* and *pR* promoters are shown as red arrows. The late transcript from *pR'* is indicated with black arrows. The CII-activated *pI*, *pRE*, and *pAQ* transcripts are indicated with blue arrows. The *pRM* transcript activated by CI is a green arrow. Transcription terminators (t) are shown as red letters among the genes. The *tI* terminator is indicated in parenthesis because it is contained within the larger *sib* processing site. The operators *OL* and *OR* where CI and *Cro* bind are shown next to the *pL* and *pR* promoters.

Lytic development

Transcription is initiated with the synthesis of the early transcripts from the *pL* and *pR* promoters (Fig 1.1). Early transcripts which encode two regulators, N and Cro, are attenuated at the *tL1* and *tR1* terminators. These transcriptional terminators play an important role in controlling the cascade of gene expression. By acting as a weak repressor for both *pL* and *pR* promoters, Cro facilitates the lytic mode. For example, the N protein which is an antitermination factor promotes the assembly of a transcription complex [Barik et al., 1987; Greenblatt et al., 1998]. This assembly occurs on the RNA at *nutL* and *nutR* sites and is made up of RNA polymerase and a number of host proteins called Nus.

The N and Nus-modified RNA polymerase can overcome the *tL1* and *tR1* transcription terminators, resulting in expression of the distal delayed early functions. The delayed early functions include the lysogenic regulators CII and CIII, as well as the lytic DNA replication genes O and P, and the late gene regulator Q [Friedman et al., 1995].

After sufficient accumulation, the Q protein modifies RNA polymerase that has just initiated transcription from the *pR'* late promoter. This modification causes the RNA polymerase to become resistant to transcription terminators present downstream to *pR'*, allowing the expression of the late genes, which encode proteins for phage morphogenesis and host cell lysis. During the last stage of the cascade, the late gene products assemble phage virions and lyse the host.

Lysogenic development

Two phage proteins, Int and CI, are required to form stable lysogens. Int allows the integration of the phage genome into the bacterial chromosome, and CI represses the two early phage promoters to prevent any lytic phage gene expression. When λ bacteriophage first infects, Int and CI are not initially made, and the λ phage initiates gene expression along a set of events that are common to both the lytic and lysogenic pathways. If conditions are favorable during this initial phase, Int and CI synthesis can be switched on, to enable lysogenic development. This activation depends primarily upon the phage CII function.

The *cII* gene is located between the *tR1* terminator and the replication genes and, thus, is transcribed with the early lytic genes. However, CII protein is required only for lysogenic development of infecting phages. Another gene required for lysogenic development is *cIII*, located beyond *tL1* in the *pL* operon.

Mutations in these genes as well as in the *cI* gene encoding the repressor function cause λ bacteriophage plaques to be clear, unlike the normal turbid plaques where the turbidity indicates growth of lysogenic cells. Whereas the CI repressor is required to maintain the repressed lysogenic state, the CII and CIII proteins are only required to initially activate CI synthesis [Kaiser 1957; Friedman et al., 2001]. Once CI has been made, the CII and CIII functions are no longer required because, CI can maintain its own synthesis.

1.3 Transcription mechanism

RNA polymerase

Transcription is regulated at several biochemical steps by protein factors through genetic signals recognized in the form of DNA or RNA [Das 1993; Burgess et al., 1987]. Transcription of all *E. coli* (*Escherichia coli*) genes is carried out by a single form of core RNA polymerase, constituted by four subunits ($\alpha_2\beta\beta'$) that are encoded by *rpoA*, *rpoB* and *rpoC* genes, respectively. The core RNA polymerase binds to one of several sigma factors that directs the RNA polymerase holoenzyme to the promoter of distinct classes of genes. σ^{70} , encoded by *rpoD*, serves as the initiator for most *E. coli* genes. Hence, the complete holoenzyme has 6 subunits: $\alpha_2\beta\beta'\sigma\omega$ (~480 kDa).

The catalytic center is constituted by both β and β' subunits, which share homology to the largest subunits of eukaryotic RNA polymerases. The β subunit can be cross-linked to nucleotides, DNA and the RNA product. Mutations in *rpoB* encoding β affect virtually every aspect of transcription like sensitivity to antibiotics, promoter recognition and interaction with σ [Gross et al., 1992], abortive initiation, elongation kinetics, intrinsic termination and regulation by termination and antitermination factors [von Hippel 1998; Rhodius et al., 1998].

Likewise, mutations in *rpoC* encoding β' alter promoter-specificity and sensitivity to rifampicin and regulation by elongation control proteins. The alpha subunit, required for core assembly, plays a pivotal role in the positive control of initiation by activators, such as the cAMP-CRP complex.

α -CTD of RNA polymerase

Activation of gene transcription in a prokaryote system is triggered by several kinds of transcription activators [Ishihama 1988]. In the *E. coli* RNA polymerase holoenzyme, one of the regions responsible for transcription activation has been localized to the α subunit at COOH terminal. The carboxyl-terminal domain α subunit (α -CTD), is regarded as the contact site for transcription activator proteins and for the promoter UP element.

Deletion of this region does not interfere with the assembly of the core or the holoenzyme, but reconstituted RNA polymerase containing C-terminal truncated alpha subunits cannot be activated by a group of transcription activator proteins [Ishihama 1992; Igarashi et al., 1991].

This group of proteins contains the class I transcription factors, and their contact sites have been placed at various positions in the C-terminal domain. The isolated α subunit and its C-terminal domain protect the UP element region from deoxyribonuclease I (DNase I) digestion, which indicates that the C-terminal portion of the α subunit is responsible for the contact with *cis*-acting UP elements as well as with *trans*-acting transcription factors. The solution structure of α CTD, a 98-amino acid COOH-terminal fragment (residues 233 to 329 plus methionine at the NH₂-terminus) was determined by NMR [Jeon et al., 1995]. The structure is compactly folded and comprises four helices and two long loops at the terminals of the domain (PDB code - 1COO).

Transcription initiation

In all organisms, transcription performed by DNA-dependent RNA polymerases can be divided into three mechanistically and structurally distinct stages: initiation, elongation, and termination. In the first step, RNA polymerase binds to the promoter, forming a metastable “closed promoter” complex. The σ^{70} holoenzyme recognizes two conserved hexamers centered around -10 (TATAAT) and -35 (TTGACA) positions relative to the start site, utilizing two conserved DNA-binding domains in σ^{70} . The closed complex then undergoes several structural transformations to isomerize into a heparin-resistant “open promoter” complex that contains a single stranded DNA bubble encompassing the -12 to +4 region. Finally, the open complex couples two specific ribonucleotides forming a dinucleotide tetraphosphate that serves as the primer for subsequent RNA chain elongation.

Transcription elongation

During elongation, RNA polymerase performs thousands of nucleotide addition cycles. Each cycle must culminate in forward translocation by one nucleotide (nt) to allow for the incorporation of the next substrate; this step entails the separation of 1 bp of the dwDNA accompanied by the displacement of one nt of the nascent RNA from the DNA template at the upstream edge of the RNA/DNA hybrid and subsequent annealing of the upstream DNA duplex [von Hippel et al., 2002]. Though the elongation complex is capable of the uninterrupted synthesis of RNA chains thousands of nucleotides long, yet, the complex becomes abruptly destabilized at terminators that demarcate the RNA end, in many cases with single nt precision. The interplay between processive synthesis, transient halting at numerous

'roadblocks' and RNA release depends on the intricate network of interactions between RNA polymerase, the nucleic acid signals and/or auxiliary transcription factors within the elongation complex [Vassylyev et al., 2007; Uptain et al., 1997; Nudler 1999].

Transcription termination [Gusarov et al., 1999; Greive et al., 2005]

Transcript elongation by RNA polymerase involves a processive mechanism. Yet, the RNA chain is not extended at a fixed rate along the DNA. Until recently, two prevalent types of sites were known to impede elongation: (a) the so-called "pause" sites, which induce a temporary, reversible block to nucleotide addition and (b) terminators, which cause the release of RNA and/or RNA polymerase, either intrinsically, or upon activation by a diffusible factor such as NusA, Rho or Tau. The dissociation of the sigma factor is thought to mark the entry of NusA protein, a key elongation modulator that couples termination and antitermination factors to RNA polymerase and itself promotes pausing and termination at specific template sites. As the elongation complex is the target of many more diffusible factors, which also includes transcript cleavage factors, which help RNA polymerase to overcome the "dead-end" sites, and additional Nus factors help RNA polymerase to transcribe processively through both factor-dependent and intrinsic terminators.

Bacteria use two main modes of terminating transcription: Rho-independent or 'intrinsic' termination, mainly requiring elements located on the mRNA, and Rho-dependent termination, relying on both mRNA elements and trans-acting factors. About half of the transcription terminators identified in *E. coli* are Rho-dependent. These terminators lie at the natural end of genes or within cistrons and in control regions preceding the coding sequences of genes, where they play an important role in the regulation of gene expression.

Rho is a homohexameric protein with RNA-dependent ATPase and helicase activities which binds to the mRNA. Essential sites on the mRNA are the Rho-binding site, known as the Rho utilization site (rut), and a distal region where the transcripts are terminated. During the transcription termination process the Rho factor which is a hexameric RNA-DNA helicase of *E. coli* binds to the nascent transcript at a 'loading site' that is rich in cytosine residues and also relatively unstructured. Once bound, Rho interacts with the *E. coli* transcription factor (NusG) and translocates directionally (5'→3') along the RNA chain by an ATP-driven process, moving towards the transcribing RNA polymerase (Fig 1.2).

Most Rho-dependent termination positions on the template are also pause sites, and therefore function (at least in part) by allowing Rho to ‘catch up’ with the paused RNA polymerase, which leads to termination by allowing Rho to use its RNA–DNA-helicase activity to unwind the RNA–DNA hybrid within the transcription bubble. The stability of elongation complexes, can be modulated by transcription factors that bind directly or indirectly (by *cis* RNA looping) to the RNA polymerase after binding to the nascent RNA. These transcription factors include NusA and NusG, which increase and decrease termination efficiency at intrinsic terminators, respectively, and increase termination efficiency (NusG) at Rho-dependent terminators. However, when assembled into complexes that contain antitermination proteins (such as the N protein that is encoded by phage λ) — which are often bound in conjunction with host proteins NusB and NusE — NusA and NusG operate together to decrease termination efficiency.

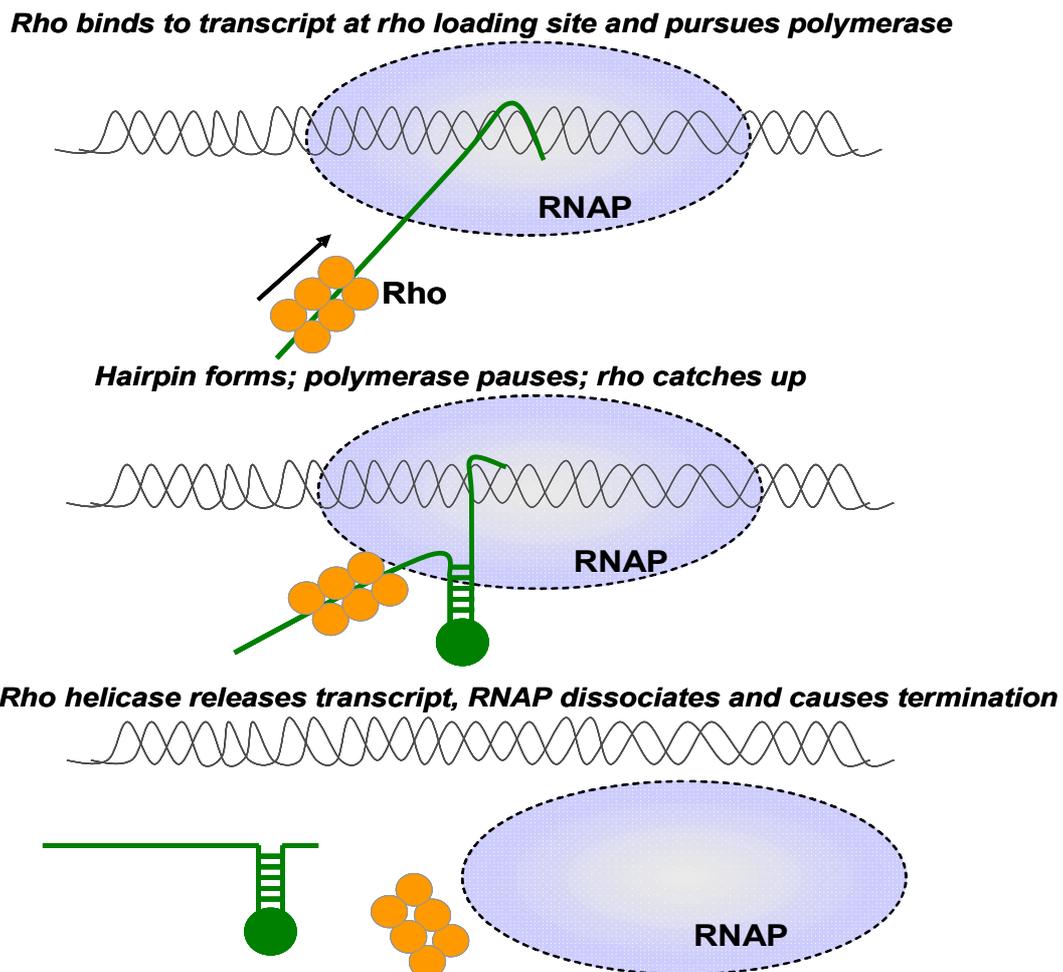


Figure 1.2 A model of rho-dependent termination.

1.4 *nut* - RNA of bacteriophage λ

The genome of the phage λ codes for two cis-acting element called *nut* site (N-utilization): *nutL* and *nutR* each of 60 basepairs long and lies in between 50-250 basepairs of the 3' side of the promoter P_L and P_R [Rosenberg et al., 1978].

The *nutL* and *nutR* sites are composed of three conserved motifs, including *BoxA* (8 nucleotides [nt] located upstream from the *BoxB* stem loop structure), *BoxB* (15 nt stem loop structure), and *BoxC* (8 nt) [Washburn et al., 2006; Das et al., 1996].

Transcription of the *nut* site provides a locus on which N and Nus factors can nucleate the formation of a specific ribonucleoprotein complex [Mogridge et al., 1998]. *BoxA* RNA recruits NusB and NusE into an antitermination complex that includes RNA polymerase, NusA, and NusG. *BoxB* RNA forms a stem-loop that binds N or HK022 Nun. As it binds *BoxB*, N associates with NusA, NusG, and RNA polymerase. It is proposed that N, Nus factors, and *nut* interact and complex with RNA polymerase while tethered on the same RNA [Nodwell et al., 1991]. Although N is the essential factor for antitermination, *nut* and the Nus factors confer stability and full activity to the antitermination complex.

BoxB RNA alone binds N and Nun with similar affinities. This equivalent affinity for *BoxB* RNA does not reflect the inability of N to compete with Nun at *nutL* *in vivo*. A third conserved motif, *BoxC* (8 nt), lies downstream of *nutL* and *nutR* and does not appear to play a role in antitermination. The two *nut* sites differ in the spacer regions between *BoxA*, *BoxB*, and *BoxC* and by a single nucleotide change in the *BoxB* loop and the sixth nucleotide in *BoxC*. The spacing between *BoxA* and *BoxB* is seven and eight nucleotides for *nutL* and *nutR* respectively (Fig 1.3) [Patterson et al., 1994].

It has been proposed [Washburn et al., 2003] that the phenotypic difference between *nutL* and *nutR* might be explained by the relative distances of the two *nut* elements from their respective promoters. *nutL* is 34 nucleotides from p_L , whereas *nutR* lies 227 nucleotides from its cognate promoter. *nutL* also differs from *nutR* in that, it lies immediately promoter proximal to RNase III cleavage sites (rIII).

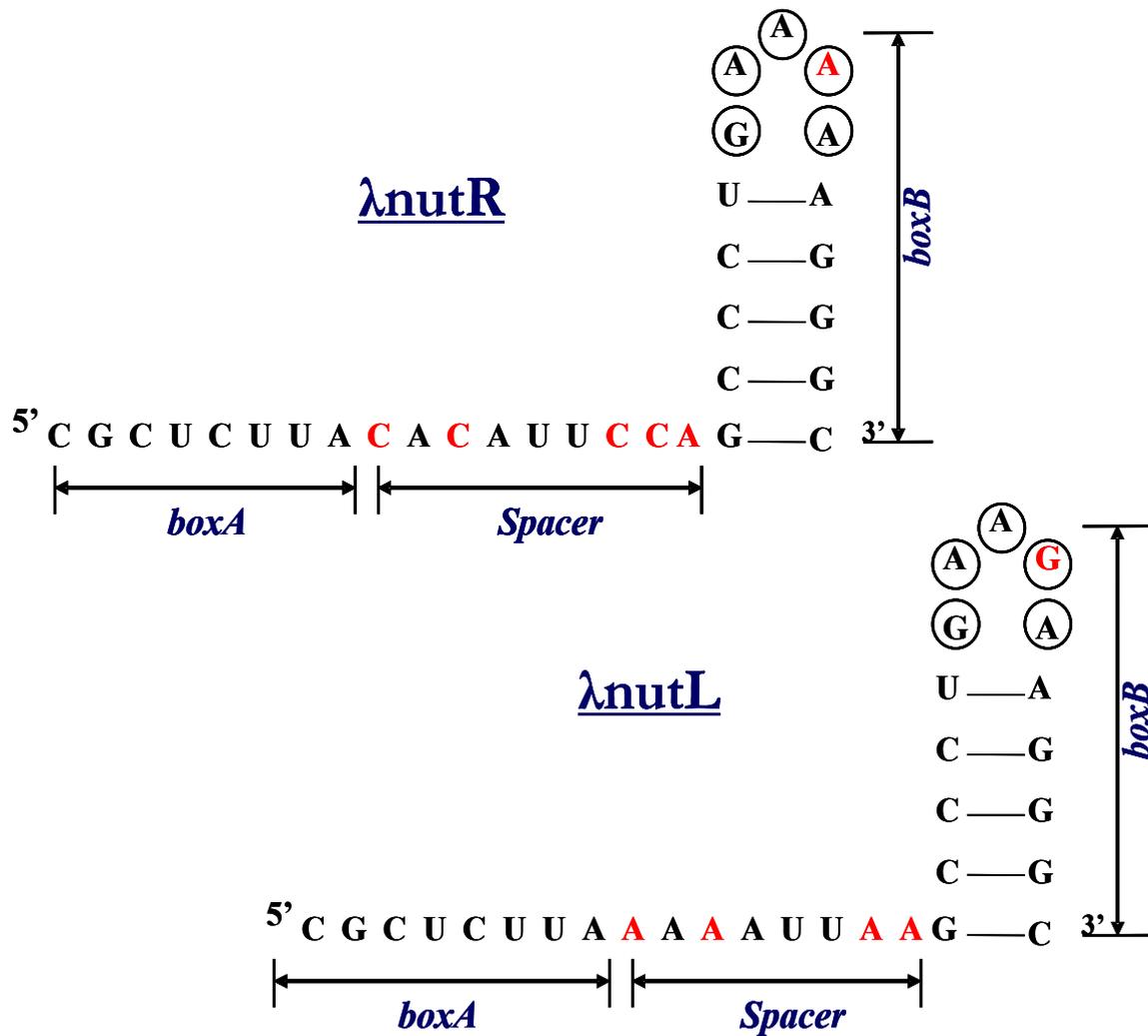


Figure 1.3 *nut* sequence showing three elements: *BoxA*, *BoxB* and the spacer region located between *BoxA* and *BoxB*. The differences are highlighted by red colored alphabets.

Although the basic elements of the λ *nut* region are conserved in many lambdoid phages, there are lambdoid phages that vary from λ paradigm. The most divergent example is phage HK022, which in place of *nut* sites, has *put* sites [Weisberg et al., 1999]. Unlike *nut* sites, which serve as a nucleation site for protein antitermination complexes, the *put* RNA structure itself appears to be necessary and sufficient for modification of RNA polymerase into an antitermination mode. However, HK022 encodes a protein, Nun, that appears to be an ortholog of N, but does not modify transcription of HK022. Instead, Nun, which is expressed from the HK022 prophage, acts as an exclusion function by binding at λ *nut* sites to arrest transcription.

1.5 Transcription antitermination and termination in *E. coli*

1.5.1 Phage λ N antitermination

Antitermination is a critical event for genetic regulation of transcription in both eukaryotic and prokaryotic cells. Antitermination involves the interplay of protein host factors with RNA and the RNA polymerase transcription complex to allow transcription through early termination sites [Greenblatt et al., 1993]. The transcriptional regulation process in bacteriophage λ can be viewed as a paradigm for antitermination. In phage λ antitermination, the N protein gene product from bacteriophage λ plays an essential role in transcriptional antitermination in the two phage early operons, which are critical for phage development. The inhibition of termination at intrinsic and Rho-dependent terminators by λ N depends upon the recognition of *nut* RNA on the nascent phage transcript [Das 1992].

The key component of the antitermination complex is the highly basic 107 amino acid λ N protein, which is largely unfolded in solution [Mogridge et al., 1998]. λ N consists of three functionally distinct regions with different interaction partners: aminoacid residues from 1-22 binds the *nutBoxB* RNA, 34-47 binds the carboxy terminal part of *E. coli* NusA acidic repeat domain 2, 73-107 forms the RNA polymerase binding region [Weisberg et al., 1999; Whalen et al., 1988; Devito et al., 1994]. Highly efficient, processive N mediated antitermination requires *E. coli* transcription elongation factors NusA, NusB, NusG, and NusE (S10), as well as *nutBoxA* [Friedman et al., 1990; Henkin et al., 2002; Agnieszka et al., 2003].

NusA, a 56 kDa essential protein, was subsequently shown to affect transcriptional pausing, termination, and antitermination. NusB, a 14 kDa protein essential for cell growth only at low temperatures, may be involved in translation as well as transcription. The *nusE71* mutation, which defined the NusE product, is an allele of *rpsJ*, encoding ribosomal protein S10. NusG, first identified through a *nusG* mutation that suppressed the effects of the *nusA1* and *nusE71* mutations, is a required factor for the N antitermination *in vitro* as well as an enhancer of termination factor Rho [Friedman et al., 1995]. The N and Nus proteins function as a complex modifying RNA polymerase to a termination-resistant form. After this complex has been formed, it leads to efficient *in vitro* and *in vivo* suppression of terminators located thousands of base pairs downstream of the *nut* site [Mogridge et al., 1995]. A model of N-dependent antitermination is shown in Fig 1.4 [Greive et al., 2005; von Hippel et al., 1996].

closer view, significant differences emerge. The differences reveal alternative strategies used by related phages to cope with similar problems and illuminate previously unknown regulatory and structural motifs [Dhillon et al., 1980; Dhillon et al., 1981].

The HK022 genome, a dsDNA molecule of 40,751 bp, has been completely sequenced (GenBank accession no. AF069308) and a majority of the genes revealed by the sequence have been assigned functions. Nun is a small 13 kDa (109 aa) arginine rich, RNA binding protein from the bacteriophage HK022. Nun terminates the RNA transcripts of phage λ when λ tries to infect Hong Kong's *E. coli* host. Nun is not essential for any part of HK022 life cycle, but only seems to prevent super-infection of *E. coli* by certain lambdoid phages.

As the λ *nut* sites are required for antitermination, surprisingly, these sites are also components of a transcription termination pathway. In this pathway, the HK022 Nun protein replaces N in the transcription elongation complex and converts the antitermination pathway into a termination reaction [Oberto et al., 1989]. In addition to the *nut* sites, the two pathways also use the host NusA, NusB, NusE, and NusG proteins [Robledo et al., 1991; Nudler et al., 2002].

The structure of the Nun/*BoxB* complex has been solved by NMR spectroscopy [Faber et al., 2001]. The *BoxB*-Nun complex is quite similar in structure to the corresponding *BoxB*-N complex. Both proteins form a bent alpha-helix upon binding to *BoxB*. However, Nun amino acids Leu-22, Ile-30, Trp-33, Ile-37 and Leu-41 form a hydrophobic surface which is not present in N (1-36) bound to *BoxB*. This surface could be a recognition site for host factors. This complex is thought to interact with RNA polymerase and Nus factors to bring about the termination reaction. Nun competes for the same *nutR* and *nutL* *BoxB* rna/stem loops effectively shutting down λ 's antitermination system.

Model for Nun action

The three histidine residues in the C-terminus of Nun, form a Zn^{2+} coordinating motif. Zn^{2+} inhibits Nun binding to *BoxB* [Watnick et al., 1998; Watnick et al., 2000]. Mutation of any of the histidines to alanine enhances *BoxB* binding and makes it insensitive to Zn^{2+} . Thus, the carboxyl-terminus of wild-type Nun acts to interfere with the N-terminal RNA-binding motif.

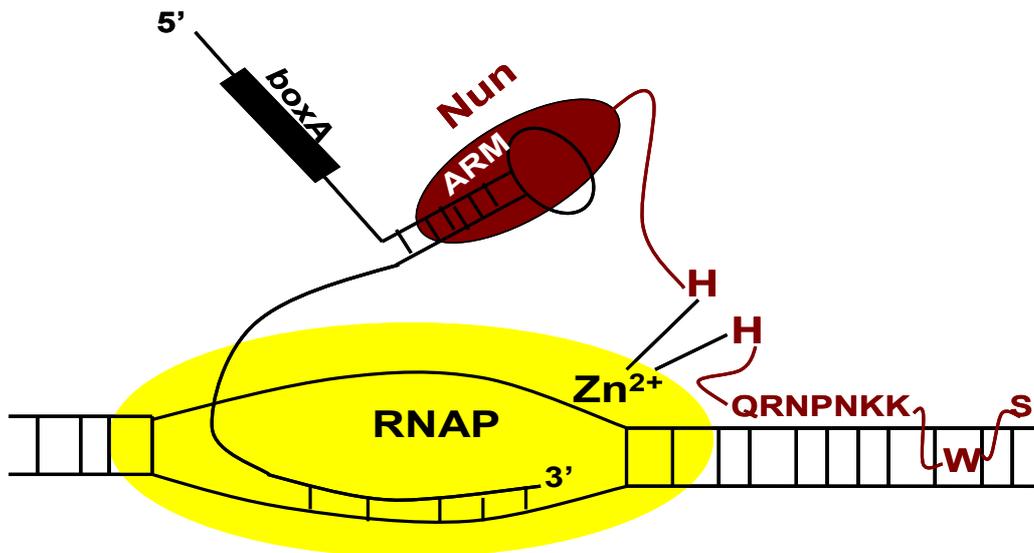


Figure 1.5 Model for transcription arrest by Nun. The amino terminal ARM of Nun interacts with *BoxB*, tethering Nun in proximity to the transcription elongation complex. Histidine residues in the C-terminal region permit Nun to contact RNA polymerase in a Zn^{2+} dependent manner. The C-terminus contacts DNA, possibly by intercalation of the penultimate tryptophan residue into the DNA template.

All three Nun C-terminal histidines are necessary for Zn^{2+} to inhibit RNA binding. However, H93A and H100A single mutants retain termination activity, which implies that only two histidines in the C-terminus are required for transcription termination. The additional residues needed for coordinating Zn^{2+} could be supplied by RNA polymerase [Garber et al., 1998].

The C-terminal location of W108 is unusual, which implies that W108 plays a role other than stabilizing the structure of the protein. W108 might intercalate into the DNA template, blocking the translocation of RNA polymerase. Nun termination, therefore, may involve two modifications of the elongating complex as shown in Fig 1.5: a tightening of the RNA polymerase clamp, and an interaction with DNA template.

NusA can also modulate RNA binding by Nun. NusA binds to the C-terminal domain of Nun and stimulates *BoxB* binding. The binding of NusA presumably sequesters the Nun CTD, allowing the Nun-NTD to bind RNA (Fig. 1.6). This inhibition is relieved by NusG, NusB, and NusE, which may translocate NusA from the CTD to another location in the transcription complex and thereby stabilizing the Nun complex and stimulate transcription arrest.

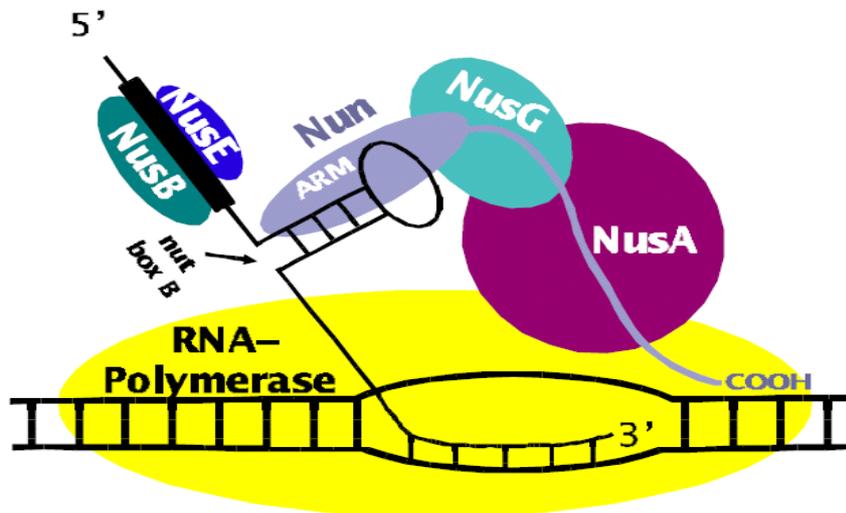


Figure 1.6 Model of Nun action in the presence of Nus factors.

1.6 Elongation factor NusA

The NusA (N utilization substance A) transcription elongation protein is one of the highly conserved host factors required by *E. coli* for the transcription termination/antitermination [Berg et al., 1989]. NusA is the largest of the Nus proteins ($M_r \sim 55,000$ in *E. coli*) and is present in the *Mycoplasma genitalium* genome, the presumed minimal set of genes required for bacterial life [Fraser et al., 1995].

NusA modulates transcription elongation by associating with the core component of RNA polymerase after promoter escape, and release of the σ^{70} subunit required for initiation at most promoters [Greenblatt et al., 1981a]. NusA influences elongation by increasing the dwell time for RNA polymerase at certain pause sites [Greenblatt et al., 1981], possibly by interacting with and stabilizing the RNA hairpin structure often associated with pause sites. These effects on the RNA polymerase elongation rate can influence gene regulation and may be necessary to couple transcription and translation and prevent premature termination caused by the Rho factor [Zheng et al., 1994].

NusA and the additional host proteins NusB, NusG, and ribosomal protein S10 are important for the N protein of bacteriophage λ to modify RNA polymerase into a termination-resistant state [Horwitz et al., 1987; Nudler et al., 2001]. N, the host proteins and *nut* site RNA assemble into a highly stable complex that associates with elongating RNA polymerase.

Within this complex, N binds *BoxB* RNA, and NusA binds to N. NusB and ribosomal protein S10 form a heterodimer that binds the *BoxA* portion of the *nut* site RNA. NusG and S10, as well as NusA, bind RNA polymerase. Even in the absence of a DNA template, a stable complex can be assembled on the *nut* site RNA that contains N, RNA polymerase and all the host cofactors. While the presence of all of the host antitermination factors allows for the formation of a stable and highly processive antitermination complex, high concentrations of N alone can cause *nut* site-independent antitermination *in vitro* [Rees et al., 1996]. As this effect is enhanced by the presence of NusA in the reaction, it was suggested that NusA functions to stabilize the N-NusA-RNA polymerase-*nut* site complex. Genetic studies on antitermination by the λ N protein have suggested that NusA may interact with the *BoxA* portion of the *nut* site as well. Based on all the evidence, it implies that NusA plays an important role in stabilizing the transcription complex (Fig 1.7). In any case, the interaction of NusA with an N-*nut* site complex is likely to be at least bipartite. First, there is a direct interaction of NusA with amino acids 34-47 of N. Second, there may be a direct interaction of NusA with *BoxA* [Greenblatt et al., 1998].

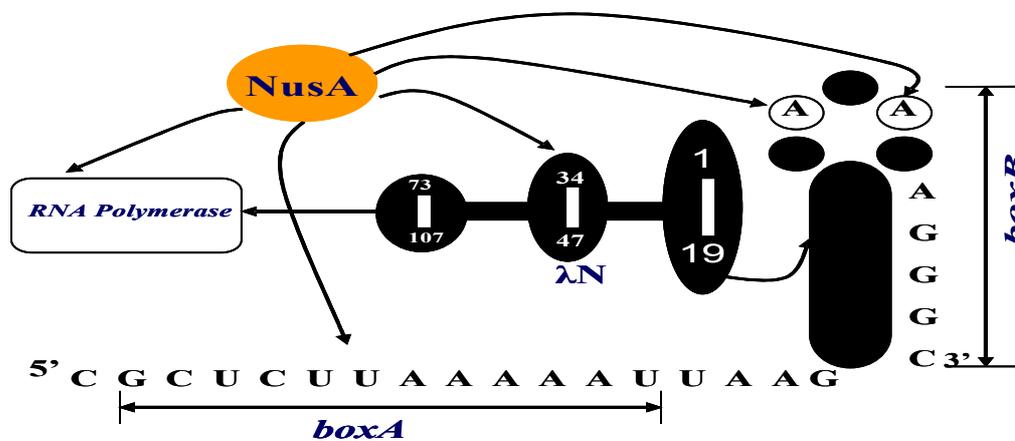


Figure 1.7 Diagram emphasizing the role of NusA in stabilizing the N-transcription complex. NusA interacts with *BoxA*, *BoxB* loop, a central region of N, and RNA polymerase.

Domain architecture of NusA

The 495 amino acid protein carries three RNA binding motifs, S1, KH1 and KH2 and two acid-rich regions, ar1 and ar2 [Mah et al., 1999]. NusA has two RNA polymerase binding domains, one at the amino terminus and the other at the C-terminal domain. The central S1 and the two KH domains are involved in the *nut*-RNA binding. The ar1 and ar2 domain at the C-terminus interacts with the λ N and α -CTD respectively (Fig 1.8).

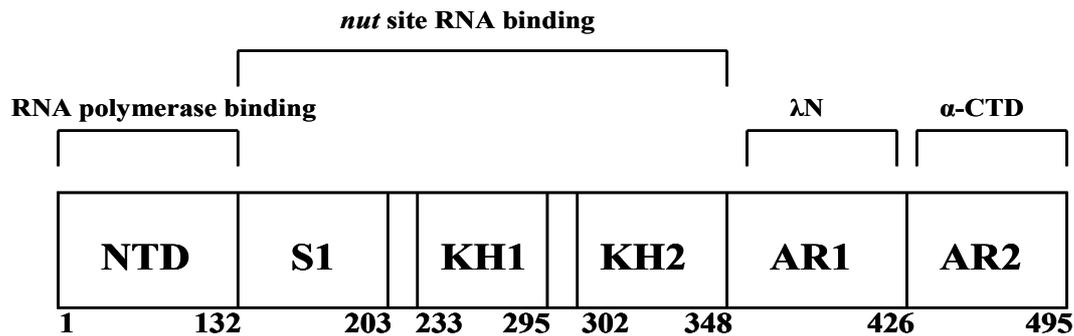


Figure 1.8 Schematic representation showing functional domains of NusA.

N-terminal domain (NTD) mediates the interaction of the protein with RNA polymerase. The NTD has two α helices lie alongside the first and last strand of a three-stranded antiparallel β sheet. The conformation of this part of the protein is therefore rather flexible and may be critically stabilized by crystal contacts. The NTD is linked by a long helix, α_3 , and coupled through a short flexible linker to three C-terminal binding domains, a single S1 domain followed by two copies KH domain (Fig 1.9). A model has been proposed where these two types of recognized RNA binding motif form an extended RNA binding interface [Worbs et al., 2001].

The S1 domain was first identified in the sequence of the *E. coli* S1 ribosomal protein and subsequently in the sequence of the NusA protein. The S1 domain folds into a five-stranded antiparallel β barrel with Greek key topology and a small 3_{10} helix following the third strand, β_6 . C terminally of the S1 domain, NusA features two consecutive K-homology motifs. They consist of three-stranded mixed β sheets packed against three (KH2) to four (KH1) α helices on one side (Fig 1.9). The KH RNA-binding domain was first identified in the human heterogeneous nuclear ribonucleoprotein K (hnRNP K) [Worbs et al., 2001]. Of particular interest is the area encompassing the S1 and KH motifs, as these domains occur in many nucleic acid binding proteins, both alone and tandemly repeated. The ubiquity of such arrays suggests that they are used as a general tool to adjust the specificity and strength of the RNA-protein interactions. KH domains are important functional components of the NusA protein [Zhou et al., 2002]. and a recently solved x-ray structure of the homologous NusA from *Mycobacterium tuberculosis* show that only these both domains act in concert to bind the RNA [Zhou et al., 2002].

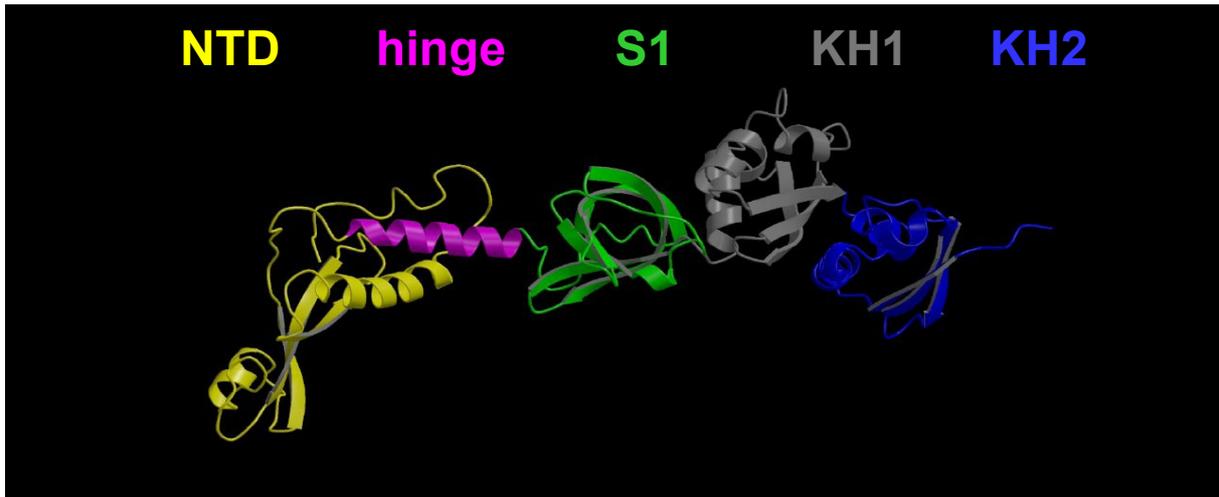


Figure 1.9 Stereo ribbon diagram of *Thermotoga maritima* NusA. Domains are indicated with different colors. PDB code – 1HH2.

The crystal structures of two non-*E. coli* NusA factors have been solved so far, *Thermotoga maritima* [Worbs et al., 2001] and *Mycobacterium tuberculosis* [Gopal et al., 2001]). These structures show a common domain organization (S1+KH1+KH2) as described before. This NusA core organization is conserved in most of the bacteria. An additional carboxy terminal region, NusA-CTD, comprising 160 residues [(NusA(353–416) and NusA(431–490)] is found in several α -, β -, and γ -proteobacteria like enterobacterium *E. coli*. Though NusA-CTD is not as highly conserved as the NusA core, the latter region is characterized by its acidity and frequently by an internal sequence repeat of 70 residues. The solution structure of NusA-CTD was solved with high-resolution by NMR [Eisenmann et al., 2005]. The two subdomains of NusA-CTD are connected by a linker region. Either subdomain contains two helix-hairpin-helix (HhH) motifs, each formed by two anti-parallel α helices connected by a short hairpin (Fig 1.10).



Figure 1.10 Structure of NusA ar1 (PDB code – 1WCL) and NusA ar2 (PDB code – 1WCN).

Autoinhibition effect of NusA

N protein alone is sufficient to bind and retard the mobility of RNA containing a wild-type *nut* site, whereas full-length NusA cannot shift the RNA on its own and needs N for its RNA binding activity (Fig 1.11-A and 1.11-C) [Mogridge et al., 1995]. None of the NusA fragments except NusA (1-416) (Fig 1.11-B) can bind the RNA directly in the absence of N. Recently, Greenblatt and coworkers showed that the extra CTD of *eco*NusA serves as an autoinhibitor of RNA binding [Mah et al., 2000]. A carboxy-terminal deletion mutant NusA (1–416), which retains the S1 and KH homology regions of NusA but only one of its two HhH motifs, can bind RNA in the absence of N. This suggests that one or more of the RNA-binding domains of NusA might be occluded by the second HhH motif or other determinants within the 79 carboxy-terminal amino acids of NusA (Fig 1.11).

It has been inferred that autoinhibition of RNA binding in *eco*NusA is mediated via a negative patch on the CTD [Mogridge et al., 1995; Mah et al., 2000]. Consistent with RNA binding to NusA being mediated by the composite positive flank, the CTD could nicely block this area or part thereof through its negative surface.

Role of α -CTD subunit of RNA polymerase

The inability of full-length NusA to bind RNA resembles the inability of the intact initiation subunit σ^{70} of RNA polymerase to bind DNA. In analogy to the way in which interaction of σ^{70} with RNA polymerase relieves the inhibitory effect of the amino terminus of σ^{70} on promoter-specific DNA binding [Dombroski et al., 1993], it is possible that the interaction of NusA with RNA polymerase relieves the inhibitory effect of the carboxyl terminus of NusA and allows NusA to bind RNA.

Nuclease protection experiments and protein-RNA cross-linking experiments [Liu et al., 1995], have already shown that NusA interacts with or is close to RNA nucleotides upstream of the 3' end of the nascent transcript in a transcription complex and these results were consistent with the observation [Mah et al., 2000] made by affinity chromatography experiments, that the α -CTD subunit of RNA polymerase stimulates RNA binding by NusA. Based on various observations, it has been suggested that during elongation, NusA uses its RNA polymerase-binding region (1–137) [Mah et al., 1999] to interact with RNA polymerase subunits β and β' , and its carboxy-terminal region to interact with α -CTD subunit.

The interaction with α -CTD subunit may then cause a conformational change in NusA such that its RNA-binding domains either fold or become exposed and competent to bind the nascent RNA (Fig 1.11-D).

Thus, as part of the transcription complex, NusA would be in a position to bind and stabilize pause and termination motifs in the nascent RNA, leading to enhancement of pausing and termination at certain sites. Hence, the interaction of the α -CTD with NusA is essential for NusA to stimulate termination only if the inhibitory carboxy-terminal region of NusA is present and not if it is deleted.

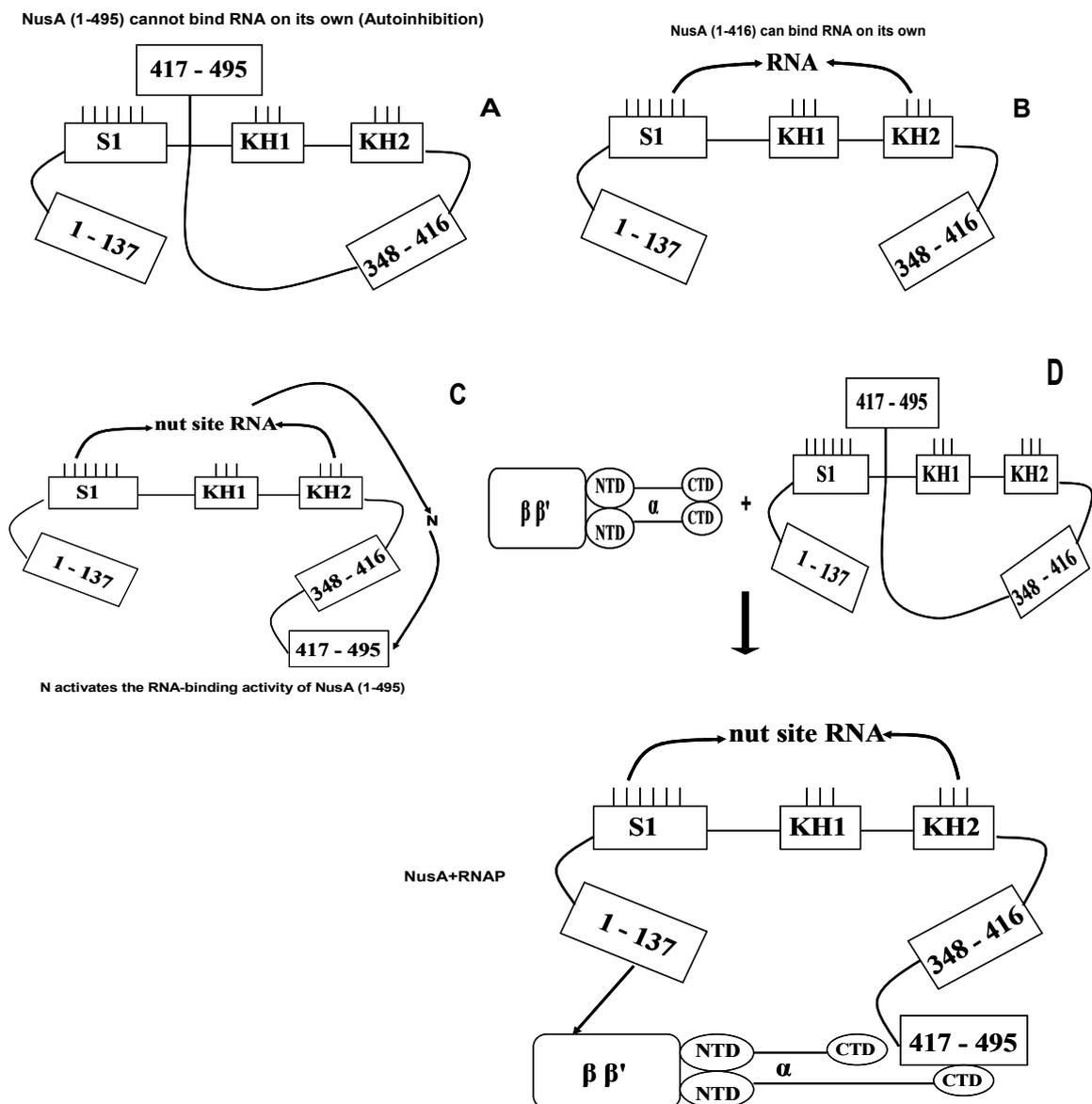


Figure 1.11 Model showing the autoinhibition effect of NusA on RNA binding.

1.7 Thesis objectives

Transcription is the primary regulatory process that is used by cells, tissues and organisms to facilitate and control the complex programmes of gene expression, cellular metabolism, and organ and tissue development. In the mechanisms of transcription termination and antitermination, participation of various Nus host factors and their interactions plays a significant role. Four N-utilization substances, NusA, NusB, NusE, and NusG, are important elongation/termination modulators. Transcription regulation through these Nus factors has been intensively studied in the expression of genes from lambdoid phages.

In this context, a unique mechanism of transcription elongation control was found in Nun protein of bacteriophage HK022. One part of my work is focused on understanding the role of Nun in the termination complex by studying the interaction of HK022-Nun with various Nus host factors on a structural level by NMR.

The other part of my project was mainly aimed at a better understanding of the regulation of RNA binding by NusA and the autoinhibition effect of NusA. Thus, this part of the work is targeted on

- ^1H , ^{13}C , and ^{15}N backbone resonance assignment of RNA binding domains of NusA (SKK domain).
- Analysis of the interaction between RNA binding domains of NusA and *nut* site RNA by NMR spectroscopy.
- NMR spectroscopy assessment of RNA binding inhibition by autoinhibition domain of NusA.

2 Materials and Methods

2.1 Culture media

The culture media was prepared using the ultra pure water (Filtering unit Milli-Q Biocel, 0.22 μm , Millipore, Eschborn) and heat sterilized at 121 °C (30 min, 1.2 bar, autoclave type 23, Varioklav, Melag, Berlin or Varioklav Dampfsterilisator, H+P labortechnik, Oberschleißheim). Prior to use, sterile filtered (0.2 μm filter, Sartorius, Goettingen) antibiotics were added to the medium as required.

2.1.1 Luria Bertani medium [Sambrook et al., 1989]

Luria Bertani (LB) medium was prepared by dissolving 10 g of peptone, 5 g of yeast extract, and 10 g of NaCl in 1000 mL of water and set to autoclave.

2.1.2 Minimal medium (M9) [Sambrook et al., 1989; Meyer et al., 1983]

Uniform labeling of proteins with ^{15}N and ^{13}C isotopes were achieved by growing the cells in minimal medium. To prepare a liter of M9 medium, 200 mL of 5 x M9 medium was diluted with 800 mL of autoclaved H_2O and supplemented with 2 mL of TS2 - trace element solution [Meyer et al., 1983], 2 mL of 1 M MgSO_4 , 1 mL of 10 mM Fe (III)-citrate, 0.1 mL of 1 M CaCl_2 , 20 mL of 20 % (w/v) glucose and 10 mL of 100 x MEM vitamin solution (Gibco, Invitrogen, Karlsruhe). To produce ^{15}N and ^{13}C labeled proteins, ^{15}N NH_4Cl or ^{15}N $(\text{NH}_4)_2\text{SO}_4$ and ^{13}C glucose were used as the sole nitrogen and carbon source respectively.

Table 2.1 Components in the 5 x M9 and trace element stock solution.

5 x M9–Medium:

$\text{Na}_2\text{HPO}_4 \cdot 12 \text{H}_2\text{O}$	85.5 g
KH_2PO_4	15.0 g
NaCl	2.5 g
NH_4Cl <i>add</i> 1000 mL of H_2O	5.0 g

TS2: (components to be added to 1000 mL of H₂O)

ZnSO ₄ . 7 H ₂ O	100 mg
MnCl ₂ . 4 H ₂ O	30 mg
H ₃ BO ₃	300 mg
CoCl ₂ . 6 H ₂ O	200 mg
NiCl ₂ . 6 H ₂ O	20 mg
CuCl ₂ . 2 H ₂ O	10 mg
Na ₂ MoO ₄ . 2 H ₂ O	900 mg
Na ₂ SeO ₃	20 mg

2.1.3 P-5052 medium [Robert et al., 2005; Studier et al., 2005]

All components for the medium were prepared using millipore water and were either filter-sterilized (0.2 µm filter, Sartorius, Goettingen) or heat-sterilized (121 °C, 30 min, 1.2 bar, autoclave type 23, Varioklav, Melag, Berlin). P-5052 medium contains 2 mM magnesium sulphate, 1 x trace metals solution, 1 x 5052 solution, 1 x NPS solution, 1 x vitamin solution, and respective antibiotic.

Stock solutions:

5000 x trace elements:

The trace metal solution contains 50 mM FeCl₃, 20 mM CaCl₂, 10 mM MnCl₂, 10 mM ZnSO₄, 2 mM CoCl₂, 2 mM CuCl₂, 2 mM NiSO₄, 2 mM Na₂MoO₄, 2 mM Na₂SeO₃ and 2 mM H₃BO₃. The solution was wrapped in an aluminum foil and stored at room temperature.

50 x NPS solution:

1.25 M Na₂HPO₄, 1.25 M KH₂PO₄ and 0.25 M Na₂SO₄. This preparation was prepared fresh, heat-sterilized and used within a week. [(An unlabeled 50 x NPS solution can be prepared by substitution of ammonium chloride as the nitrogen source (2.5 M NH₄Cl)]. For labeling, the media contains 5 g of ¹⁵N NH₄Cl (Cambridge Isotope Laboratories, Andover, USA) per liter.

50 x 5052 solution:

25 % glycerol (v/v), 2.5 % glucose (w/v) and 10 % lactose (w/v). This preparation was prepared fresh and heat-sterilized.

2.2 Estimation of protein concentration

A measure of protein and nucleic acid concentration was obtained upon monitoring the absorbance at 280 nm and 260 nm respectively. Absorption was measured either in a black wall quartz cuvette (Hellma, Müllheim) with a thickness of 1 cm, using a Helios γ spectrophotometer (Thermo spectronic, Cambridge, UK) or else in the Nanodrop ND-1000 instrument (Peqlab Biotechnology GmbH, Germany). After measuring, the protein concentration was then determined by applying Beer-Lambert's law [Ingle et al., 1988] as follows which is related to the amount of light absorbed by the sample.

$$A(\lambda) = \epsilon \cdot c \cdot b \quad [2.1]$$

Where $A(\lambda)$ = measured absorbance at 280 nm

ϵ = molar extinction coefficient ($M^{-1} \text{ cm}^{-1}$)

c = concentration of the substance that absorbs light (M)

b = path length of the sample cell (cm)

Molar extinction coefficients [Gill et al., 1989] of proteins and nucleic acids were obtained from amino acid analysis using *ProtParam Tool (ExPASy, Expert Protein Analysis System proteomics server, Swiss institute of Bioinformatics, www.expasy.ch, Switzerland)*.

2.3 SDS – polyacrylamide gel electrophoresis [Laemmli 1970]

The analytical electrophoresis of proteins was carried out in polyacrylamide gels under the conditions that ensure dissociation of the proteins into their individual polypeptide subunits. Thereby the molecular mass of the protein was determined by electrophoresing it together with “marker” protein of known molecular masses that bracket that of the protein of interest [Voet et al., 2nd edition].

In this technique, the proteins were separated on a porous supporting material prepared by cross linking acrylamide by N-N methylene bis-acrylamide. Initially the proteins were denatured by heating them at 95 °C for 10 min in a buffer containing sodium dodecyl sulphate [(SDS)-an ionic detergent] and β -mercaptoethanol (reducing reagent). The gels (10 x 8 x 0.75 cm) were electrophoresed in Mighty small SE250/260 gel electrophoresis chambers (Hoefler, San Francisco, CA, USA) at a constant voltage of 30 mA.

The SDS gel has two distinct zones, a stacking gel overlaying a separation gel, both of Tris buffered system. The two zones are characterized by their porosity and pH conditions. The compositions for the preparation of two gels are mentioned in Table 2.2.

Table 2.2 Ingredients for 19 % Sodium dodecyl sulphate polyacrylamide gel (SDS-PAGE).

19 % SDS gel	Separation gel	Stacking gel
Na ₂ SO ₃	50 mg	-
H ₂ O	2.13 mL	8.70 mL
0.5 M Tris/HCl; pH 6.8	-	5.0 mL
3 M Tris/HCl; pH 8.8	3.75 mL	-
30 % (w/v) Acrylamide rotiphorese® Gel A (Roth, Karlsruhe)	18.50 mL	2.6 mL
2 % (w/v) Bisacrylamide rotiphorese® Gel B (Roth, Karlsruhe)	7.70 mL	1.08 mL
10 % SDS	0.3 mL	0.2 mL
TEMED	20 µL	20 µL
10 % (w/v) APS	200 µL	200 µL

The separated proteins were visualized by Coomassie Brilliant Blue staining [Wilson 1983] and destaining with concentrated methanol : acetic acid solution. The gel after destaining was photographed using a Gel document system (GEL DOC 2000, Biorad, Munich).

2.4 Schagger and Jagow gel electrophoresis [Schagger et al., 1987]

The most generally used technique to visualize the polypeptides with masses below about 15 kDa is the one developed by Schagger and von Jagow. This technique employs a discontinuous gel system containing SDS. However, the interference of SDS with the stacking and separation of small polypeptides is diminished by changing the trailing ion (in the cathode buffer) from glycine to the more mobile Tricine (N-tris[hydroxymethyl]-methylglycine) [Wisdom 1997] and by lowering the pH of the separating gel.

The separation gel consists of 16.5 % T and 6 % C in 1 M Tris/HCl, pH 8.45, 0.1 % (w/v) SDS and 6 M urea. The stacking gel comprises 4 % T and 6 % C in 0.775 M Tris/HCl, pH 8.45 and 0.1 % (w/v) SDS. The polymerization was initiated by addition of 10 µL of TEMED and 100 µL of 10 % (w/v) APS for 20 mL gel solution.

The upper part of the electrophoresis instrument was filled with cathode buffer (0.1 M Tris/HCl, pH 8.25, 0.1 M Tricine, 0.1 % (w/v) SDS and the base is filled with anode buffer (0.2 M Tris/HCl, pH 8.9). A constant voltage of 28 mA was applied until the samples reached the end of the stacking gel and then the voltage is increased to 40 mA. For the estimation of the molecular weight in gels, Fluka molecular weight standard (Fluka, New-Ulm) or Peptide marker kit (GE Healthcare life sciences, Germany) was used. Staining and destaining of the gels were carried out in the same manner as that of SDS-PAGE.

2.5 HK022 Nun protein

2.5.1 Expression of Nun

Nun full length construct (1-112) was expressed in the *E. coli* BL21(DE3). Pre-inoculum was developed by inoculating the glycerol stock with 200 mL of LB medium containing ampicillin antibiotic (1 µg/mL) and incubating overnight at 37 °C. Expression of protein was carried out by inoculating 1.2 liter of LB medium containing ampicillin antibiotic with overnight culture so as to have an initial concentration of cells corresponding to OD₆₀₀ (Optical density at 600) of about 0.1. The culture was then incubated with shaking (170 rpm) (C25KC Incubator shaker, New Brunswick Scientific, Edison, NJ, USA) at 37 °C, until the culture has reached the mid-log phase of the growth (OD₆₀₀ ~0.8). The expression was then induced by addition of 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) (GERBU, Gaiberg). The cells were allowed to grow till the stationary phase (~4 h) and were then harvested by centrifuging (Centrikon T-124, Rotor A 6.9, Kontron, Eching) at 6000 rpm (5000 g), 4 °C for 30 min. The cell pellets were washed with 50 mM Tris-HCl, pH 8.0 and were harvested again by centrifugation. The cells obtained are then stored at -80 °C until further use.

2.5.2 Cell lysis and purification of HK022 Nun

Preparation of cell extract

Frozen cell pellets were resuspended in lysis buffer (50 mM Tris-HCl, pH 8.0, 2 mM ethylenediaminetetraacetic acid (EDTA), 1 mM phenylmethylsulfonylflouride (PMSF), 5 mM Dithiothreitol (DTT), 0.2 mg/mL deoxyribonuclease I (DNase I), one protease inhibitor tablet-EDTA free (Roche, Mannheim). The cells were freeze and thawed three times. After freeze

and thaw, the suspension was stirred on ice for 30 min. The cell suspension was sonified for 2 times of each 1 minute with ultrasound (Duty cycle 0.5, 200 Watt, Sonifier Labsonic U, B. Braun Biotech International, Melsungen) with 10 min pause in between each step. Cellular debris was cleared from the cell lysate by centrifuging (Biofuge Stratus, Rotor 3334, Heraeus) at 4 °C, 13000 rpm (19000 g) for 30 min. The cleared lysate was filtered (Minisart Sterilfilter, 0.45 µm, Sartorius, Goettingen) and used for the purification.

Purification

All the buffers used were filtered and degassed before purification.

Binding buffer : 50 mM Tris-HCl, pH 8.0, 2 mM EDTA

Elution buffer : 50 mM Tris-HCl, pH 8.0, 2 mM EDTA, 2 M NaCl

HK022-Nun was purified by employing cation exchange chromatography. The ÄKTA *purifier* 10-system (Amersham Biosciences, Freiburg) was used to purify the protein with a HiTrap™ Heparin column (5mL ~ 1 CV) (Amersham Biotech). The column was equilibrated with 10 CV of Binding buffer. The combined supernatant after centrifugation were applied on the heparin column at a flow rate of 1mL/min. Flow through was collected in each step. The column was washed with buffer A for 5 CV or more to remove unbound proteins. The bound protein was then eluted by buffer B with a step gradient of 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100 % respectively. Fractions collected during the elution were analyzed by SDS-PAGE (2.3).

High-performance liquid chromatography (HPLC) [Horvath et al., 1967]

Fractions containing Nun which were identified on a 19 % SDS-polyacrylamide gel were pooled and subjected to HPLC (Kontron, Eiching) for further purification.

Buffer A: 0.1 % TFA in millipore water

Buffer B: 0.1 % TFA, 80 % acetonitrile in millipore water

Purification was carried out by using a HPLC C-18 column (PrepLC™ 25 mm Module, Waters, Milford, Massachusetts, USA) by applying a slow linear gradient. Fractions containing Nun were kept in the speedvac (ABM Greiffenberger, Marktredwitz) to dry. The dried samples were stored in the cold room till further use.

2.6 Nun carboxy terminal domain (Nun CTD)

2.6.1 Cloning and expression of Nun CTD

The Nun C-terminal domain gene was amplified by polymerase chain reaction (PCR) with a 5' primer including a NcoI restriction site and a 3' primer with a BamHI site. The fragment was cloned in to the expression vector pET-GB1 (G. Stier, EMBL, Heidelberg, Germany) via NcoI and BamHI. The expressed protein is N-terminally fused to the 56 residues long streptococcal immunoglobulin-binding domain of protein G [Huth et al., 1997; Zhou et al., 2001]. The 16.6 kilodalton (kDa) fusion protein contains a 6 x His tag at the N-terminus and a tobacco etch virus (TEV) protease cleavage site between GB1 and Nun CTD, and carrying resistance to kanamycin antibiotic.

DNA isolation and quantification

For isolating plasmid DNA, a 10 fold concentrated cell suspension was used. Isolation of DNA was performed according to the manufacturer's instruction using the NucleoSpin[®] Plasmid Kit (MACHEREY-NAGEL GmbH & Co. KG, Düren, Germany). The purified DNA was eluted with 50 µL of H₂O. After the isolation of DNA, the DNA product was analyzed by 0.8 % agarose gel by loading the plasmid DNA (3 µL of plasmid DNA + 3 µL of 6 x loading dye). As a DNA marker, λ DNA/Hind III marker (MBI-Fermentas, Germany) was used.

Polymerase chain reaction [Kramer et al., 2001]

The Nun C-terminal domain (45 - 109) gene was amplified by PCR with a 5' primer including a NcoI restriction site (underlined): 5'-GCC CAT GGG TGT GAC ACC TGG ATT TAA TGC TAT AGA TGA C-3' and a 3' primer with a BamHI site (underlined): 5'-TGG GAT CCT TAA CTC CAT TTT TTG TTC GGG TTG-3'. The oligonucleotides were obtained from Biomers.net GmbH, Ulm, Germany. All the enzymes and buffers used in this study were obtained from New England Biolabs and MBI-Fermentas.

The relevant coding regions of DNA were amplified using Primus 25 advanced Thermocycler (Peqlab Biotechnology GmbH, Germany). 50 µL of PCR mixture contained, 1 µL genomic DNA, 10 x Thermopol-buffer 5 µL, dNTPs (each 200 µM) 1 µL, primer forward (100 pmol/µL) 1 µL, primer reverse (100 pmol/µL) 1 µL, VENT-polymerase 0.5 µL and sterile water 40.5 µL. Reaction mixture was subjected to 35 amplification cycles, consisting of

denaturation at 94 °C for 2 minutes, and further extension of denaturation was carried out by subjecting to 35 cycles of 30 seconds at 94 °C. Followed by denaturation, primer annealing at 50 °C for 30 s and elongation at 72 °C for 4 min. A final extension of 4 min at 72 °C was performed. The PCR product was analyzed by electrophoresis using 1.5 % agarose gels. After the electrophoresis the PCR fragment was excised from the gel and extracted as per the instructions provided in the QIAquick Gel Extraction Kit protocol (Qiagen, Hilden).

Restriction digestion analysis (BamHI)

Restriction digestion analysis was performed with PCR amplified fragments as well as with pET GB1 vector to check the sequence similarity and the arrangement of Nun CTD genes in the amplified locus. Restriction enzyme (BamHI) with recognition site 5'-GGATCC-3' was used according to the recommendations of the manufacturer. A reaction mixture of 60 µL in total containing the PCR fragment or GB1 vector (47 µL), BamHI (3µL), 10 x buffer for BamHI (6µL) and sterile water (4µL). The reaction mixture was incubated at 37 °C (Incubator type BE 200, Memmert, Schwabach) for 1.5 hour. After the digestion reaction, the respective fragments were analyzed by agarose gel electrophoresis.

Ethanol precipitation

To the BamHI digested samples, added 2 µL of Glycogen, 1/10 volume of 3 M Lithium chloride and 3 volume of 100 % ethanol. Incubated the samples for 1 h at -80 °C. After incubation, centrifuged at 13,000 rpm (10,000 g) for 15 min at 4 °C, decanted the supernatant, and drain inverted on a paper towel. Add 70 % ethanol (corresponding to about three volume of the original sample), incubated at room temperature (RT) for 5-10 min and centrifuged again for 10 min, and decanted and drained the tube, as above. Placed the tube in a Speed-Vac (ABM Greiffenberger Antriebstechnik, Marktredwitz) and dried the DNA pellet for about 5-10 min, or until dried. Dissolved the dried DNA in 20 µL of sterile water.

Restriction digestion analysis (NcoI)

Followed by ethanol precipitation, the samples were subjected to NcoI digestion. The restriction enzyme NcoI with a recognition site 5'-CCATGG-3' was used according to the recommendations of the manufacturer. The reaction mixture contains ethanol precipitated sample (20 µL), NcoI (2µL), 10 x NE buffer 4 (3µL) and sterile water (5µL) with a total volume of 30 µL. The mixture was incubated for 15 min at 37 °C.

Dephosphorylation of vector plasmid [Ahmad et al., 1981]

After digesting the plasmid with the desired enzyme it was collected on a preparative gel by gel extraction procedure. To the digested vector (57 μ L), added 2 μ L of Antarctic phosphatase, 7 μ L of 10 x buffer for the enzyme and 4 μ L of sterile water. The mixture was incubated for 20 min at 37 °C. Heat inactivated (or as required to inactivate the restriction enzyme) for 5 minutes at 65 °C.

Quick ligation

To proceed with the ligation, Quick ligation Kit (New England Biolabs) was used. The reaction mixture was incubated at RT by setting up with 2 μ L of dephosphorylated vector, 3 μ L of the insert, 5 μ L of H₂O, 10 μ L of 2 x Quick ligation reaction buffer and 1 μ L of Quick T4 DNA ligase.

Butanol precipitation [Thomas 1994]

This method was carried out to get rid of the salt from ligation buffer and to increase transformation efficiency and eliminate arching of cuvettes. Adjusted the volume of ligation reaction up to 50 μ L by adding sterile water. To this added 10 x amount of butanol, mixed well and spinned at 13,000 rpm for 5 min at 4 °C. The supernatant was discarded and the pellet was dried in the speedvac for 5 min. The dried pellet was resuspended in 10 μ L of sterile water and followed with transformation into *E. coli* BL21 DE3.

Transformation and expression

Prior to electroporation [Sugar et al., 1984], added 1 – 2 μ L of ligation sample to 40 μ L aliquot of *E. coli* BL21 DE3 competent cells and kept on ice for 2 min. The mixture was then added to the electroporation cuvette (0.1 cm, BioRad, Munich) and kept in the electroporator (Micro™ Pulser, BioRad, Munich). Electroporation was started with a pulse of 1.8 Kvolts. Now the contents of the cuvette was transferred to an eppendorf tube and added 500 μ L of LB medium and incubated at 37 °C for 1 hour. After incubation 50 μ L was plated on LB agar plate and left overnight at 37 °C (Incubator Model 200, Memmert, Schwabach).

Pre-inoculum was prepared by picking up the transformants from the plate and inoculating in to 5 mL LB containing the respective antibiotic at 37 °C (Incubator shaker Certomat HK/R, B.Braun Biotech International, Melsungen). After incubation for 6-8 h, final inoculum was

developed by inoculating 1 mL of this pre-inoculum with 10 mL of LB medium containing respective antibiotic and incubating overnight at 37 °C. The Nun CTD expression was carried out by inoculating the overnight culture in to the fresh media with a start OD₆₀₀ ~0.1, and grown at 37 °C, 170 rpm (C25KC Incubator shaker, New Brunswick Scientific, Edison, NJ, USA) till the OD₆₀₀ reaches ~0.7 to 0.8, and then induced with 1 mM IPTG (GERBU, Gaiberg) and allowed to grow further for 3.5 hours. Cells were pelleted at 6000 rpm (5000 g) (Centrikon T-124, Rotor A 6.9, Kontron, Eching), at 4 °C for 30 min and stored at -80 °C.

2.6.2 Cell lysis and purification of Nun CTD

Preparation of cell extract

Frozen cell pellets were thawed on ice and resuspended in 10 mL of lysis buffer per 100 mL cell culture volume.

Lysis buffer : 20 mM Tris-HCl, pH 8.0, 1 mM PMSF, 1 mM β-mercaptoethanol (Fluka, Ulm), 150 mM NaCl, DNase I, EDTA free-Protease Inhibitor tablets (Roche) and 10 mM imidazole.

The cell suspension after resuspending in the lysis buffer was stirred on ice for 30 min. The cells were then freeze and thaw three times. After freeze and thaw, the suspension was sonified (Sonifier Labsonic U, B.Braun Biotech International, Melsungen) for 3 times with 0.5 pulse and 75 % amplitude with five minutes pause in between each steps. The cell lysate was then centrifuged (Biofuge Stratus, Rotor 3334, Heraeus) at 4 °C, 13,000 rpm (19,000 g) for 30 min. After centrifugation the supernatant was filtered via a syringe filter of pore size 0.45 μm (Minisart Sterilfilter, 0.45 μm, Sartorius, Goettingen).

Purification

Wash buffer 1 :

20 mM Tris/HCl, pH 8.0, 10 mM imidazole, 150 mM NaCl, 1 mM β-mercaptoethanol

Wash buffer 2 :

20 mM Tris/HCl, pH 8.0, 10 mM imidazole, 1 M NaCl, 1 mM β-mercaptoethanol

Elution buffer :

20 mM Tris/HCl, pH 8.0, 500 mM imidazole, 1 M NaCl, 1 mM β-mercaptoethanol

The recombinant protein was purified by nickel affinity chromatography (HisTrap chelating column, Amersham Biosciences, Freiburg). The supernatant was loaded on to a Ni-ion affinity

column by using a P1-pump (Amersham Biosciences, Freiburg) with a flow rate of 1 mL/min and the flow through was reloaded on to the column again. The column was washed with 10 CV of wash buffer 1 and 10 CV of wash buffer 2 and then eluted by applying a imidazole step gradient at the ÄKTA *purifier* 10-system (Amersham Biosciences, Freiburg). The eluted fractions were analyzed by SDS gel.

Peak fractions containing Nun CTD were pooled and dialyzed using Spectra/Por dialysis tubing, [(MWCO-10000), Flatwidth-45mm, Diameter-29mm, 6.4mL/cm, Roth, Karlsruhe)] against 5 L of 20 mM Tris/HCl, pH 8.0, 100 mM NaCl and 2 mM β -mercaptoethanol in the cold room overnight. After dialysis, the N-terminal histidine tag from the Nun CTD was cleaved off using TEV protease by incubating the enzyme with the dialyzed fractions in the cold room overnight with gentle shaking. The cleaved sample was then loaded onto the Histrap and collected the flow through containing target protein. The fractions were analyzed by Schagger-Jagow gel and the pure fractions were pooled and dialyzed against water and set to lyophilize (ABM Greiffenberger, Marktredwitz) for further use.

2.7 NusA full length

2.7.1 Expression of NusA

The recombinant NusA (1-495) construct was cloned in pTKK 19k (by Dr. Sabine Schwarz, Department of Biopolymers, University Bayreuth). A starter culture was set by inoculating the glycerol stock in 200 mL of LB medium containing kanamycin antibiotic. The culture was grown overnight at 37 °C (Incubator shaker Certomat HK/R, B.Braun Biotech International, Melsungen). Expression was carried out by inoculating 1.2 L of LB medium with kanamycin antibiotic with the overnight culture, to have an initial concentration of cells corresponding to OD₆₀₀ of about 0.1.

The culture was then incubated with shaking (170 rpm) at 37 °C (C25KC Incubator shaker, New Brunswick Scientific, Edison, NJ, USA), until the culture has reached an OD₆₀₀ of 0.7. The expression was induced by addition of 2 mM IPTG. The cells were then allowed to grow for further five hours and harvested by centrifuging at 6000 rpm (5000 g), 4 °C for 20 min (Centrikon T-124, Rotor A 6.9, Kontron, Eching). The cell pellets were stored at -80 °C until further use.

2.7.2 Cell lysis and purification of NusA

Preparation of cell extract

Thawed cell pellets were resuspended in 100 mL of 1 x binding buffer with 1 mM β -mercaptoethanol, 0.2 μ g/mL DNase I and protease inhibitor tablet (Roche), stirred on ice for 30 min. The cell suspension was sonified 3 x 1 minute with 0.5 pulse and 75 % amplitude (Sonifier Labsonic U, B.Braun Biotech International, Melsungen). Sonicated sample was centrifuged at 13,000 rpm (19,000 g) for 30 min at 4 °C (Biofuge Stratus, Rotor 3334, Heraeus). The supernatant was then loaded onto a histrap column (Amersham Biosciences, Freiburg) for purification.

Purification

1 x Binding buffer : 20 mM Tris-HCl, pH 8.0, 500 mM NaCl, 5 mM imidazole

1 x Elution buffer : 20 mM Tris-HCl, pH 8.0, 500 mM NaCl, 500 mM imidazole

Washed the column with 10 CV of binding buffer and then eluted with a step gradient using ÄKTA *purifier* 10-system (Amersham Biosciences, Freiburg). The eluted fractions were loaded on 10 % SDS gel and stained with Coomassie Brilliant Blue after electrophoresis.

Fractions containing NusA were then pooled and dialyzed (Spectra/Por dialysis tubing (MWCO-10000), Flatwidth-45mm, Diameter-29mm, 6.4mL/cm, Roth, Karlsruhe) against 20 mM Tris/HCl, pH 8.0, 150 mM NaCl overnight. N-terminal his-tag was cleaved from the fusion protein by 1 unit of PreScission protease enzyme for 100 μ g of protein by incubating the sample with the enzyme in the cold room overnight by gentle stirring.

The cleaved sample was further loaded onto the Histrap and the flow through was collected. The flow through containing NusA was analyzed by running 10 % SDS gel. The fractions corresponding to NusA were pooled and dialyzed against water, which was later subjected to lyophilization and stored in the cold room till further use.

2.8 NusA acidic repeat 1 (NusA ar1)

2.8.1 Expression of NusA ar1

NusA ar1 (cloned by Dr. Stefan Prasch, Department of Biopolymers, University of Bayreuth) was expressed in *E. coli* BL21(DE3). The expression of ¹⁵N labeled NusA ar1 was carried out as spar-preparation [Marley et al., 2001]. The cells were grown in 2 x 2 L LB_{amp} medium until OD₆₀₀ reached 0.8. Cells were then centrifuged at 16 °C, 6000 rpm (5000 g) for 15 min (Centrikon T-124, Rotor A 6.9, Kontron, Eching) and the cell pellet was washed with 500 mL of wash buffer (6.4 g Na₂HPO₄, 1.5 g KH₂PO₄ and 0.25 g NaCl). After washing, it was subjected to centrifugation at 16 °C, 6000 rpm (5000 g) for 15 min and the cell pellets were resuspended in 1 L of 1 x M9 minimal medium containing ¹⁵NH₄Cl and all the additives. Later the cells were incubated at 37 °C for 1 h with shaking at 170 rpm. After 1 h, the cells were induced by 1 mM IPTG. After 4 h of induction the cells were harvested by centrifuging at 6000 rpm, 4 °C for 15 min. Cell pellets were resuspended in binding buffer (20 mM sodium phosphate, pH 7.4, 500 mM NaCl and 1 mM DTT) and stored at -80 °C until further use.

2.8.2 Cell lysis and purification of NusA ar1

Preparation of cell extract

The cell pellets were freeze / thawed three times. After freeze and thaw, added 0.2 µg/mL DNase I and 0.2 µg/mL lysozyme and stirred on ice for 45 min. The cell suspension was sonified (Sonifier Labsonic U, B. Braun Biotech International, Melsungen) 2 x 1 minute with 0.9 pulse and 90 % amplitude and 1 x 10 min with 0.5 pulse and 60 % amplitude with 1 min pause on ice and stirring in between each step. The cell lysate was then centrifuged at 4 °C, 13000 rpm (19,000 g) for 30 min to clear the cellular debris. After centrifugation the cleared lysate was filtered by using a syringe filter of pore size 0.45 µM.

Purification

Binding buffer for Histrap : 20 mM sodium phosphate, pH 8.0, 500 mM NaCl, 1 mM DTT
Elution buffer for Histrap : 20 mM sodium phosphate, pH 8.0, 500 mM NaCl, 1 mM DTT,
1 M imidazole

The recombinant protein was purified by nickel affinity chromatography (5 mL HisTrap chelating column, Amersham Biosciences, Freiburg). The supernatant was loaded onto a Ni-ion affinity column with a flow rate of 1 mL/min and eluted by applying a imidazole step gradient. The eluted fractions were analyzed by Schagger-Jagow gel. Peak fractions containing NusA ar1 were pooled and dialyzed (Spectra/Por dialysis tubing (MWCO-3500), 6.4mL/cm, Roth, Karlsruhe) against 50 mM Tris/HCl, pH 8.0, 150 mM NaCl and 1 mM DTT overnight. After dialysis, N-terminal deca-Histidine tag was cleaved off using PreScission protease (1 unit for 100 µg) by incubating the enzyme with the dialyzed fractions overnight with gentle shaking. The cleaved sample was then dialyzed against 50 mM Tris/HCl, pH 7.4, 1 mM DTT overnight. Further purification was carried out by loading the sample onto a 5 mL HiTrap™ QXL column (Amersham Biosciences, Freiburg).

Binding buffer for QXL : 50 mM Tris/HCl, pH 7.4, 1 mM DTT

Elution buffer for QXL : 50 mM Tris/HCl, pH 7.4, 1 mM DTT, 1 M NaCl

The sample was eluted from the QXL column by a NaCl step gradient and the eluted fractions were pooled and concentrated using Vivaspin concentrators of MWCO–5000 (Vivascience AG, Hannover, Germany). Concentration was carried out by centrifuging the concentrators at 4 °C, 5000 rpm (Universal 320R, Rotor 1494, Hettich).

2.9 NusA acidic repeat 2 (NusA ar2)

2.9.1 Expression of NusA ar2

The recombinant NusA ar2 construct was cloned in pET 19b (by Dr. Stefan Prasch, Department of Biopolymers, University of Bayreuth) and expressed in *E. coli* BL21 (DE3). Pre-day culture was grown in 50 mL LB with 0.1 mg/mL ampicillin for 8 h at 37 °C, 180 rpm (Incubator shaker Certomat HK/R, B. Braun Biotech International, Melsungen) and then transferred into 500 mL of LB and grown overnight at 37 °C. 5 L of LB_{amp} was inoculated with the overnight culture to a start OD₆₀₀ of ~ 0.2, and grown at 37 °C until the OD₆₀₀ reached 0.8. The gene expression was induced for 4 h with 1 mM IPTG. 4 h after induction, cells were harvested by centrifugation at 4 °C, 6000 rpm (5000 g) for 10 min (Centrikon T-124, Rotor A 6.9, Kontron, Eching) and pellet was resuspended in binding buffer and stored at -80 °C for further use. Gene expression was monitored by Schagger-Jagow gel analysis.

2.9.2 Cell lysis and purification of NusA ar2

Cell lysis and purification

Cells were resuspended in (4 mL/g) 20 mM sodium phosphate buffer, pH 7.4, 500 mM NaCl, 1 mM DTT. After three freeze–thaw cycles, lysozyme, DNase I and one protease inhibitor cocktail tablet were added and the suspension was stirred on ice for 30 min. The cell suspension was sonified (Sonifier Labsonic U, B.Braun Biotech International, Melsungen) for 3 x 1 min, pulse 0.9, 100 % amplitude. After sonication the cell extract was centrifuged (Biofuge Stratus, Rotor 3334, Heraeus) for 45 min at 4 °C, 13,000 rpm (19,000 g) to separate the cell debris. The supernatant was filtered using Minisart Sterilfilter, 0.45 µm. Purification of the fusion protein was performed by a step gradient using nickel ion affinity (10 mL Histrap column, Amersham Biosciences, Freiburg, Germany) chromatography by a ÄKTA purifier system.

Binding buffer : 20 mM sodium phosphate, pH 7.4, 500 mM NaCl, 1 mM DTT

Elution buffer : 20 mM sodium phosphate, pH 7.4, 500 mM imidazole, 1 mM DTT

The filtered supernatant was loaded onto the column with a flow rate of 1 mL/min. After loading the sample, the column was washed with 5 CV of binding buffer and eluted with a imidazole step gradient. Fractions containing NusA ar2 were pooled, dialyzed against 50 mM Tris/HCl, pH 8.0, 500 mM NaCl, 1 mM DTT and cleaved with PreScission protease (1 unit for 100 µg protein) overnight at 4 °C. To further purify the protein and to remove the NaCl, the cleaved sample was dialyzed against 50 mM Tris/HCl, pH 7.4, 1 mM DTT overnight in the cold room. The dialyzed sample was loaded onto a 5 mL Q XL column connected in tandem to a 5 mL GST column. After loading, the column was washed with 5 CV of binding buffer and eluted by a NaCl gradient.

Binding buffer: 50 mM Tris/HCl, pH 7.4, 1mM DTT

Elution buffer : 50 mM Tris/HCl, pH 7.4, 1mM DTT, 1 M NaCl

Fractions eluted were analyzed by Schagger-Jagow gel. The pure fractions were then dialyzed against NMR buffer (50 mM sodium phosphate buffer, pH 7.6, 100 mM NaCl, 10 mM β-mercaptoethanol, 1 % glycerol (v/v), 0.5 mM EDTA) overnight and concentrated in Vivaspin concentrators with a MWCO-5000 (Vivascience, Stonehouse, UK).

2.10 S1+KH1+KH2 domain of NusA (SKK)

2.10.1 Expression of SKK

The NusA RNA binding domains containing amino acid 132–348 referred as SKK was cloned via BamHI and NdeI restriction sites in to the *E. coli* expression vector pET11a (Dr. Stefan Prasch, Department of Biopolymers, University of Bayreuth). The N-terminal His₆ tagged SKK domain was expressed and purified according to the published protocols [Mah et al., 1999; Mah et al., 2000] with minor changes. To study the interaction of NusA with *nut* RNA, construct (SKK) lacking the N-terminal domain and the two acidic repeat domains was used, since these regions are not directly involved in RNA binding [Mah et al., 2000]. For the expression of unlabeled SKK, plasmid pET 11a-SKK domain was transformed into *E. coli* BL21(DE3) host cells. The pre-inoculum was set by growing the strain harboring the recombinant plasmid in 100 mL LB medium at 37 °C containing ampicillin antibiotic. The pre-culture was used to inoculate 2 L of LB medium to an OD₆₀₀ of 0.2. Cells were then incubated at 37 °C, 170 rpm (C25KC Incubator shaker, New Brunswick Scientific, Edison, NJ, USA). At an OD₆₀₀ of 0.6, expression was induced by addition of 0.1 mM IPTG followed by incubation at 37 °C for 4 h. To monitor the expression, aliquots containing equal amounts of cells were taken from the culture every hour and applied to a Schagger-Jagow gel. Cells were harvested 4 h after induction by centrifugation (Centrikon T-124, Rotor A 6.9, Kontron, Eching) at 6000 rpm (5000 g) for 15 min at 4 °C and the cell pellet was stored at –80 °C.

Increasing molecular size of the protein leads to crowded spectra, increase in number of resonances, faster relaxation, broad lines, low intensity, overlapping signals, and long experiment time. So in order to improve the resolution and sensitivity of the NMR spectra in triple resonance experiments and in the binding studies, deuterated SKK was prepared.

Deuterated ¹⁵N-labeled samples were expressed in cells grown on M9 minimal media containing 1.5 g/L ¹⁵N (NH₄)₂SO₄ and 2 g/L glucose. While preparing ¹³C,¹⁵N labeled samples the unlabeled glucose was replaced by 2 g/L of ¹³C glucose. In order, to optimize the expression of SKK domain, cell cultures were adapted to grow in D₂O by increasing the amount of D₂O in the M9 minimal medium from 0 to 100 % with each growth cycle. To produce deuterated SKK, five pre-cultures with increasing D₂O content were grown prior to inoculation of the main culture. At first, an overnight pre-culture was set at 37 °C in LB

containing ampicillin antibiotic (100 $\mu\text{g}/\text{mL}$). Further cultures were grown at 37 °C in M9 minimal medium with increasing D₂O concentration of 25 %, 50 %, 75 %, and 100 % respectively. All of the cultures were maintained at subsaturating cell densities, with A₆₀₀ typically below 0.6. Each pre-culture was inoculated to an OD₆₀₀ of 0.2 with the respective amount of the preceding pre-culture. The final cell culture containing ~ 100 % D₂O was induced with 0.1 mM IPTG when the culture density reached A₆₀₀ = 0.6 and grown for further 4 h before harvesting. Cells were pelleted at 6000 rpm (5000 g) (Centrikon T-124, Rotor A 6.9, Kontron, Eching) for 10 min and stored at -80 °C overnight.

2.10.2 Cell lysis and purification of SKK domain

Cell lysis and purification

Frozen cell pellets were thawed and resuspended in cell lysis buffer (20 mM Tris/HCl, pH 7.9, 500 mM NaCl, 10 % glycerol (v/v), 5 mM β -mercaptoethanol; 10 mL/g pellet). The cell suspension was shock frozen three times. Lysozyme, DNase I (0.1 mg/mL) and one protease inhibitor cocktail tablet (Complete, EDTA free, Roche) were added to the suspension. Now the cell suspension was stirred on ice for 30 min and followed by sonication (Sonifier Labsonic U, B. Braun Biotech International, Melsungen) on ice. The cells were sonicated 4 x 1 minute with 1.0 pulse and 100 % amplitude with 1 minute pause on ice and stirring in between each step. The lysate was then centrifuged (Biofuge Stratus, Rotor 3334, Heraeus) at 13,000 rpm (19,000 g) for 45 min at 4 °C.

The supernatant was filtered by using a syringe filter of pore size 0.45 μM . Purification of the soluble his-tagged protein was performed by nickel ion affinity chromatography on an ÄKTA *purifier* 10-FPLC system. The supernatant was loaded on a 5 ml HisTrap column (Amersham Biosciences, Freiburg, Germany) at 1 mL/min which was pre-equilibrated with cell lysis buffer.

Binding buffer : 10 mM HEPES, pH 7.6, 100 mM NaCl, 10 % glycerol (v/v), 5 mM β -mercaptoethanol

Elution buffer : 10 mM HEPES, pH 7.9, 100 mM NaCl, 10 % glycerol (v/v), 5 mM β -mercaptoethanol, 300 mM imidazole

The column was first washed with 10 CV of cell lysis buffer followed by 10 CV of binding buffer. After the washing step, the bound protein was eluted by applying a imidazole step gradient consisting of 5, 10, 20, 50, 80, and 100 % elution buffer. The fractions containing protein were dialyzed (Spectra/Por, MWCO 3500, ROTH, Karlsruhe, Germany) against 5 L of 50 mM sodium phosphate buffer, pH 7.6, 100 mM NaCl, 10 mM β -mercaptoethanol, 0.5 mM EDTA, and 1 % glycerol (v/v).

The sample was dialyzed 3 times against 5 L of the above mentioned buffer. The first two dialysis step was done for 5–6 hours and the final dialysis step was done overnight. All the dialysis steps were carried out at 4 °C. After the dialysis, the protein sample was concentrated using vivaspin concentrators of MWCO–5000 Da (Vivascience AG, Hannover, Germany). The concentration was carried out by centrifuging the concentrators at 4 °C, 5000 rpm (Universal 320R, Rotor 1494, Hettich) till a concentration of 400 – 500 μ M was reached. The concentrated sample was stored at -80 °C till further use.

2.11 NusG

2.11.1 Expression of NusG

BL21 (DE3) cells were transformed with pET 11a/NusG in LB medium containing ampicillin antibiotic and grown with shaking at 170 rpm (Incubator shaker Certomat HK/R, B. Braun Biotech International, Melsungen) overnight. Expression was carried out by inoculating 1.2 L of LB medium containing ampicillin antibiotic with the overnight culture so as to have an initial concentration of cells corresponding to OD₆₀₀ of about 0.1.

The culture was then incubated with shaking (170 rpm) at 37 °C (C25KC Incubator shaker, New Brunswick Scientific, Edison, NJ, USA), until the culture has reached the mid-log phase of the growth (OD₆₀₀ ~0.7). At this point the expression was induced by the addition of 1 mM IPTG. The cells were then allowed to grow till the stationary phase (~4 h) and were then harvested by centrifuging at 6000 rpm, 4 °C for 30 min (Centrikon T-124, Rotor A 6.9, Kontron, Eching). The cell pellets were stored at –80 °C until further use. Expression analysis was carried out by 19 % SDS page.

2.11.2 Cell lysis and purification of NusG

Preparation of cell extract and purification [Pasman et al., 2000]

The cell pellets were resuspended at 10 mL/g of lysis buffer (20 mM Tris-HCl, pH 7.8, 3 mM EDTA, 1 mM DTT, 100 mM NaCl, 1mM PMSF, lysozyme, DNase I, protease inhibitor cocktail tablet (Complete, EDTA-free, Roche Diagnostics GmbH, Mannheim). After resuspending the cells were incubated for 10 min at 22 °C (Thermomixer 5436, Eppendorf), and then 20 min on ice. Sodium deoxycholate was added to a final concentration of 0.06 % (w/v), and the resulting solution was incubated for 20 min on ice. After incubation, NaCl was added to a final concentration of 0.3 M from a stock solution of 4 M and stirred for 10 min on ice. The lysate was sonicated 4 times for 30 s at room temperature by applying 0.7 pulse and 70 % amplitude with 3 min on ice between each sonication treatments. Polymin P was added drop wise (from a stock solution of 10 % at pH 7.8) with stirring to a final concentration of 0.6 %, after which the lysate was incubated for 30 min on ice and the lysate was sonicated once again as previously mentioned. After sonication the sample was centrifuged at 13,000 rpm (19,000 g) for 30 min (Biofuge Stratus, Rotor 3334, Heraeus).

Saturated ammonium sulphate solution at pH 8.0 was prepared (515,3g/L at 4 °C). Ammonium sulfate at 50 % of saturation was added to the supernatant drop wise, with stirring. After this, the sample was incubated for 30 min on ice. The lysate was centrifuged for 30 min at 13,000 rpm (Biofuge Stratus, Rotor 3334, Heraeus) at 4 °C and the pellet was resuspended in 30 mL (for 1 L culture volume) of buffer Q (10 mM Tris-HCl, pH 7.8, 1 mM EDTA, 1 mM DTT, 5 % glycerol (v/v)). The lysate was then dialyzed against 4 x 1 L of buffer Q for a total of 16 h and after dialysis the solution was spun at 13,000 rpm (Biofuge Stratus, Rotor 3334, Heraeus) for 30 min and the supernatant was filtered using a syringe filter of pore size 0.45 µm and applied onto HiTrap™ QXL column (Amersham Biosciences, Freiburg) at a flow rate of 0.5 mL/min.

Binding buffer : 10 mM Tris-HCl, pH 7.8, 1 mM EDTA, 1 mM DTT, 5 % glycerol (v/v)

Elution buffer : 10 mM Tris-HCl, pH 7.8, 1 mM EDTA, 1 mM DTT, 5 % glycerol (v/v),
200 mM NaCl

Washed the column with 10 CV of binding buffer and then eluted with a linear of 0-200 mM NaCl gradient. Pure fractions were pooled and concentrated with Vivaspin concentrators.

2.12 NusB

2.12.1 Expression of NusB

To produce NusB, auto-induction of recombinant protein expression, a method introduced by Studier [Studier 2004; Grabski et al., 2003] was carried out.

Expression plasmids were transformed into *E. coli* BL21 DE3, by electroporation (Micro™ Pulser, BioRad, Munich). The transformed cells were then plated on a LB agar plate with kanamycin antibiotic and grown overnight at 37 °C. In the morning, a single colony was picked from the transformation plate and transferred into a culture tube containing 20 mL of P-5052 medium (2.1.3). The culture was grown at 37 °C with shaking at 170 rpm overnight (Incubator shaker Certomat HK/R, B. Braun Biotech International, Melsungen). On the next day, 10 mL of the overnight culture was used to inoculate 1 L of 1 x P-5052 medium. Cells were grown at 37 °C till it reaches an OD₆₀₀ of 0.5 and then transferred to 20 °C and allowed to grow overnight. From the overnight culture the cells were harvested by centrifugation at 6000 rpm for 15 min and the cell pellets were resuspended in the lysis buffer (20 mM of 50 mM Tris/HCl, pH 7.5, 150 mM NaCl) and stored at -80 °C.

2.12.2 Cell lysis and purification of NusB

Preparation of cell extract

Frozen cell pellets were freeze/thawed three times. After adding one protease inhibitor tablet, 0.2 mg/mL DNase I and 0.2 mg/mL lysozyme, the cell lysate was stirred on ice for 45 min. The sample subjected to sonication (0.5 pulse and 100 % amplitude) for 5 x 1 minute. The crude cell lysate was centrifuged at 4 °C with 13,000 rpm for 30 min. After sonication, to the supernatant added 20 mM imidazole as an end concentration and filtered the sample via syringe filters of pore size 0.45 μM.

Purification

Binding buffer (A1) : 50 mM Tris/HCl, pH 7.5, 150 mM NaCl and 20 mM imidazole

Binding buffer (A2) : 50 mM Tris/HCl, pH 7.5, 150 mM NaCl and 50 mM imidazole

Elution buffer (B) : 50 mM Tris/HCl, pH 7.5, 150 mM NaCl and 500 mM imidazole

Histrap chelating column, was equilibrated with 10 CV of buffer A1. The cleared lysate was loaded onto the column with a very slow flow rate of 0.2 mL/min. The column was washed with buffer A2 for about 10 CV and eluted with a step gradient by using the buffer B. The peak fractions containing NusB were pooled and dialyzed against 50 mM Tris/HCl, pH 7.5, 10 mM NaCl, 2 mM DTT, and TEV protease to cleave the N-terminal histidine tag. The cleaved sample was further loaded onto the histrap column and the flow through was collected. The flow through was pooled and dialyzed against 50 mM Tris/HCl, pH 7.5. The dialyzed sample was loaded onto the QXL column for further purification.

Binding buffer for QXL : 50 mM Tris/HCl, pH 7.5

Elution buffer for QXL : 50 mM Tris/HCl, pH 7.5, 1 M NaCl

The loaded protein was then eluted with a step gradient ranging from 0 to 1 M NaCl. The corresponding fractions containing pure protein were pooled, dialyzed against water and subjected to lyophilization.

2.13 RNA oligonucleotide

In vitro transcription and RNA preparation

All RNAs were prepared by *in vitro* transcription from synthetic DNA templates (IBA GmbH, Göttingen, Germany) using a single polypeptide chain enzyme-T7 polymerase. The sequence of the *nutL* DNA template was (GCC CTT CTT CAG GGC TTA ATT TTT AAG AGC GCT ATA GTG AGT CGT ATT A) and the *nutR* was (GCC CTT TTT CAG GGC TGG AAT GTG TAA GAG CGC TAT AGT GAG TCG TAT TA) respectively.

The transcription cocktail contains 100 μ M DNA template, 20 mM of each NTP's, 0.5 M of $MgCl_2$, 40 % PEG 8000, 200 μ M T7 promoter, 0.1 mg/mL of T7 RNA polymerase enzyme and 10 x transcription buffer which contains (40 mM Tris-HCl pH 8.1, 50 mM DTT, 10 mM spermidine, 0.1 % (v/v) Triton X-100) was set to a total volume of 5 ml and incubated for 4 h at 37 °C (Incubator Model 200, Memmert, Schwabach). The reaction was quenched by adding 0.2 M EDTA, pH 8.0. The reaction mixture was precipitated by adding 3 M sodium acetate, pH 5.3 and 100 % chilled ethanol and incubated overnight at -80 °C. Later, the sample was centrifuged for 30 min at 5000 rpm. The pellet was subjected to speedvac (ABM Greiffenberger Antriebstechnik, Marktredwitz) for drying.

The dried sample was dissolved by heating with 8 M urea at 95 °C (Block thermostat BT100, Kleinfeld labortechnik, Gehrden) for 5 min. The RNA was purified on denaturing 20 % polyacrylamide gels containing 8 M urea by an overnight run. 600–1000V constant by a Multidrive XL (Amersham Biosciences, Freiburg) was used for running the RNA gel. Product bands were cut from the gel using UV shadowing. The RNA was electroeluted using a Schleicher and Schuell electroelution apparatus and subsequent ethanol precipitation. To remove multivalent ions and other low molecular weight impurities the RNA sample was first dialyzed against 10 mM potassium phosphate, pH 6.4, 100 mM NaCl, 5 mM EDTA and then finally dialyzed against water.

RNA was quantified by UV absorption at 260 nm wavelength. The extinction-coefficient of the RNAs were obtained from http://www.ambion.com/techlib/misc/oligo_calculator.html. The RNA sample was lyophilized and stored at -20 °C till further use. Approximately five reactions on a 5 mL scale yielded around 6–7 mg of purified RNA.

2.14 NMR spectroscopy

2.14.1 NMR sample preparation

NMR samples were prepared by dialyzing the purified protein against the respective NMR buffer. After dialysis, the samples were either concentrated with Vivaspin concentrators (Vivascience AG, Hannover, Germany).

For the interaction study of Nun with various Nus factors (NusA ar1, NusB, NusG) either buffer containing 10 mM KPO₄, pH 6.4, 50 mM NaCl or buffer with 50 mM NaPO₄, pH 7.0, 50 mM NaCl was used. NMR samples of SKK domain were prepared in 50 mM sodium phosphate, pH 7.6, 100 mM NaCl, 10 mM β-mercaptoethanol, 0.5 mM EDTA and 1 % glycerol (v/v).

All the NMR samples contained 0.04 % sodium azide as an antimicrobial agent, 2 x complete protease inhibitor tablet from a stock of 25 fold (prepared by dissolving one tablet in 2 ml of sterile water) and 10 % (v/v) D₂O for the field frequency lock. The total volume was adjusted to 550 μL or 600 μL and the solution was then transferred to a 5 mm ultra precision NMR tubes (Norell, Landsville, NJ, USA).

2.14.2 NMR spectrometers and measurements

All NMR experiments were performed at 298 K on Bruker DRX 600 MHz, Bruker Avance 700 MHz and Bruker Avance 800 MHz spectrometers in standard configuration by using triple-resonance probes equipped with self-shielded gradient coils. The water suppression was performed by either a binomial 3-9-19 WATERGATE sequence [Sklenar et al., 1993; Piotto et al., 1992] or using *Water Flip-Back pulses* [Grzesiek et al., 1993a; Grzesiek et al., 1993b] or a coherence selection with pulsed field gradients [Schleucher et al., 1994]. To acquire data in the time domain, phase sensitive quadrature detection was used. The quadrature detection in the indirectly detected dimensions was obtained by the popular methods known as States or States-TPPI (*time proportional phase incrementation*) method [States et al., 1982; Marion et al., 1989]. When the coherence selection with gradients was employed during the three dimensional NMR experiments, the echo/anti-echo method was used [Kay et al., 1992; Schleucher et al., 1993].

The proton chemical shifts were referenced to external standard DSS (2,2-dimethyl-2-silapentane-5-sulfonic acid). The chemical shifts of ^{13}C and ^{15}N resonances were referenced indirectly using Ξ ratios of the zero-point frequencies of the proton ($\omega_{\text{C}}/\omega_{\text{H}} = 0.251449527$ and $\omega_{\text{N}}/\omega_{\text{H}} = 0.101329051$) [Live et al., 1984].

For backbone resonance assignments, we had implemented transverse relaxation optimized spectroscopy (TROSY) in to triple resonance experiments for increased sensitivity. The standard TROSY-type NMR experiments like [^{15}N , ^1H]-TROSY, tr-HNCO, tr-HN(CA)CO, tr-HNCA, tr-HN(CO)CA, tr-HNCACB and tr-HN(CO)CACB were performed [Salzmann et al., 1998]. Additionally backbone assignments were completed using a combination of NNH-NOESY and ^{15}N -HSQC-NOESY spectra to trace the sequential connectivities.

2.14.3 NMR data processing and analysis

The NMR data were processed using *in-house* written software [Schweimer 2000] and analyzed with the program package NMRview [Johnson et al., 1994]. The digital resolution of the recorded spectra was enhanced by zero filling [Cavanagh 1996] to at least twice the number of acquired points, and Lorentzian-to-Gaussian or sine-bell or squared sine-bell window functions applied. Data processing consists typically of SVD-Linear Prediction

[Barkhuijsen et al., 1985] with root reflection [Press et al., 1992] in one heteronuclear dimension (normally the ^{15}N dimension of triple resonance experiments or the X-dimension in the X-edited spectra) and Fourier transformation. Baseline correction in the acquisition dimension was performed using a model free algorithm [Friedrich et al., 1995]. For the three dimensional experiments the Program process_f1f3 and process_f2 [Schweimer 2000] was used.

2.14.4 Experiments necessary for Biomolecular NMR

2.14.4.1 The HSQC experiment

The HSQC (Heteronuclear Single Quantum Coherence) [Müller 1979; Bodenhausen et al., 1980; William et al., 1994] experiment is a routinely used experiment in biomolecular NMR spectroscopy. It was invented by Geoffrey Bodenhausen and D. J. Ruben in 1980.

The resulting spectrum is two-dimensional with one axis for ^1H and other for a heteronuclei most often ^{15}N . The natural abundance of ^{15}N is very low and their gyromagnetic ratio is markedly lower than that of protons. Therefore, two strategies are used for increasing the signal intensity of these nuclei: Isotopic enrichment of these nuclei in proteins and enhancement of the signal to noise ratio by the use of inverse NMR experiments in which the magnetization is transferred from protons to the heteronucleus.

The ^{15}N -HSQC is often referred to as the fingerprint of a protein because each protein has a unique pattern of signal positions [Cavanagh et al., 1996; Ernst et al., 1992]. Thus in the ^{15}N -HSQC one signal is expected for each amino acid residue with the exception of proline which has no amide-hydrogen due to the cyclic nature of its backbone. The HSQC also contains signals from the NH_2 groups of the side chains of asparagine and glutamine and of the aromatic H^{N} protons of tryptophan and histidine. It correlates the chemical shift of the proton with that of its attached heavy atom (^1H - ^{15}N pair). This information can be very useful, in particular for recognizing whether a protein is folded and intact. It also forms the basis for nearly all multi nuclear 3D spectra.

2.14.4.2 TROSY experiment

For large molecular systems, the conventional NMR methods have two main problems [Fernandez et al., 2003]. First, the large number of resonances causes signal overlap, which can make analysis of the spectra very difficult. Second, transverse spin-relaxation *via* dipole-dipole coupling and chemical shift anisotropy leads to an overall increase in signal line width, and a corresponding decrease in spectral resolution (Fig 2.1).

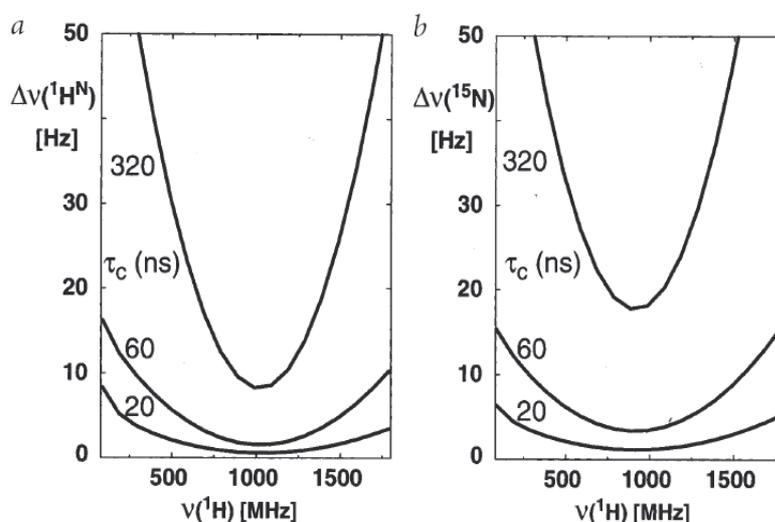


Figure 2.1 Frequency dependence from 100–1800 MHz of the full resonance line width at half height for amide groups in TROSY experiments calculated for three correlation times of $\tau_c = 20, 60,$ and 320 ns, which represent spherical proteins with molecular weights of 50, 150, and 800 kDa. (A) $^1\text{H}^{\text{N}}$ linewidth. (B) ^{15}N linewidth. The calculation uses axial symmetric CSA tensor of $^{15}\text{N} = 155$ ppm, and $^1\text{H}^{\text{N}} = 15$ ppm, and the angle between the principal tensor axis and the N–H bond was assumed to be 15° for ^{15}N and 10° for $^1\text{H}^{\text{N}}$; $l(\text{N–H}) = 0.104$ nm; effects of long-range dipole-dipole couplings with spins outside of the ^{15}N - ^1H moiety were not considered. Figure reproduced from [Wüthrich 1998].

In principle, the overlap of signals in the NMR spectra can be overcome by reducing the number of resonance lines by a proper choice of isotopic labeling schemes. Major sources of relaxation are the omnipresent hydrogen atoms. Their replacement by deuterons substantially reduces transverse relaxation, resulting in increased resolution and significant sensitivity gains. It has been recognized that at very high magnetic field strengths, dipole-dipole (DD) and chemical shift anisotropy (CSA) interactions in a ^{15}N - ^1H pair can be utilized to obtain sharp line widths for very large proteins or protein complexes. An important pulse sequence called TROSY has been developed [Pervushin et al., 1997; Pervushin 2000]. TROSY takes

advantage of mutual cancellation of CSA and DD relaxation effects at high fields. TROSY is basically a heteronuclear correlation experiment (particularly for ^{15}N - ^1H spin pairs) in which the proton magnetization is first transferred to ^{15}N , then evolves during t_1 under differential relaxation mechanisms of the ^{15}N doublet due to CSA (^{15}N) and dipole-dipole interaction (^{15}N - ^1H).

Magnetization is then transferred back to the proton prior to detection under differential line broadening of the proton doublet due to CSA (^1H) and dipole-dipole interaction (^{15}N - ^1H). In TROSY experiments, decoupling is not used, and J-coupled peaks are resolved. When monitoring the ^{15}N decoupled spectra, two peaks would be seen as there are two possible orientations of the bound hydrogen (*spin up* or *spin down* states). When the hydrogen nucleus is in the *spin up* state, the dipole-dipole coupling between the ^{15}N and ^1H will lead to a local ^1H field which always has the same directionality as the CSA contribution.

Conversely, when ^1H is in the *spin down* state, the local ^1H field always has directionality opposite to that of CSA contribution. This means that in the *spin down* state, the DD coupling of the system effectively reduces the chemical shift anisotropy. Since the chemical shift anisotropy is directly proportional to the square of the external magnetic field, it is possible to adjust the external field to a level at which the DD coupling and CSA exactly cancel each other (which occurs at 1.1 GHz).

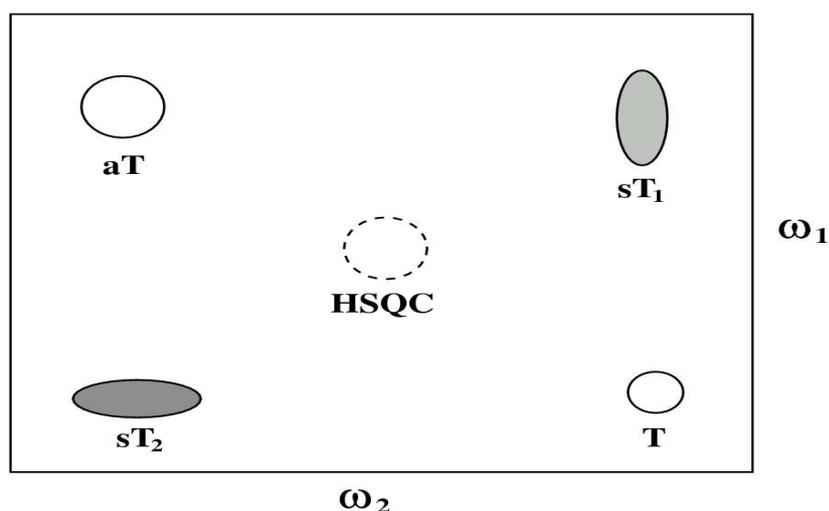


Figure 2.2 Representation of the ^{15}N - ^1H TROSY multiplet pattern. The slowest relaxing component, the TROSY peak (marked with a T), the two semi-TROSY peaks in ω_1 (sT_1) and ω_2 (sT_2), anti-TROSY peak labeled as (aT) are shown in the figure. In a decoupled HSQC, the central peak (labeled as HSQC) appears as a superposition of fast and slow relaxing components, thus being prone to rather fast relaxation.

The resulting cross peak is a multiplet of four peaks, each having different width and relaxation rate in the ω_1 and ω_2 dimensions. In contrast, these four multiplets (arising from two different line widths for each N and H) for each amide proton are superimposed in the HSQC spectra due to decoupling in F_1 and F_2 .

In TROSY spectra, among these four multiplets, only the one which is not affected by line broadening due to DD and CSA is selected by application of appropriate phase cycling. These four multiplets are shown in Fig 2.2. Several methodologies for obtaining TROSY spectra free of errors and artifacts have been developed in recent years [Kojima et al., 2000; Herbrüggen et al., 2000]. Use of TROSY elements in pulse programs for sequence specific assignment has now become part of routine NMR.

2.14.4.3 Sequence specific assignments

Sequence specific assignments using three dimensional experiments [Ikura et al., 1990; Sattler et al., 1999] is basically an extension of Kurt Wüthrich's strategy which exclusively relies on the homonuclear ^1H NMR experiments and ^{13}C , ^{15}N experiments for additional couplings [Fesik et al., 1988]. Combination of TROSY with deuteration and standard ^1H , ^{15}N triple resonance experiments provided an avenue for successful backbone assignment of NusA-RNA binding domain.

Table 2.3 Triple resonance experiments for the backbone assignments.

Experiment	Nuclei observed
HNCO	$\text{H}(i), \text{N}(i), \text{CO}(i - 1)$
HNCA	$\text{H}(i), \text{N}(i), \text{C}_\alpha(i), \text{C}_\alpha(i - 1)$
HN(CO)CA	$\text{H}(i), \text{N}(i), \text{C}_\alpha(i - 1)$
HN(CA)CO	$\text{H}(i), \text{N}(i), \text{C}'(i)$
HN(CO)CACB	$\text{H}(i), \text{N}(i), \text{C}_\alpha(i - 1), \text{C}_\beta(i - 1)$
HNCACB	$\text{H}(i), \text{N}(i), \text{C}_\alpha(i), \text{C}_\beta(i), \text{C}_\alpha(i - 1), \text{C}_\beta(i - 1)$

The nomenclature for these triple resonance experiments reflects the magnetization transfer pathway of the experiments. Nuclei that are involved in magnetization transfers form the name of an experiment. Spins, whose chemical shifts are not evolved are put in parentheses.

For an out-and-back type experiment [Archer et al., 1991], where magnetization of a spin is transferred to a remote spin and then brought back the same way, only the first half of the magnetization transfer is used for its name. For example, the out-and-back experiment that transfers magnetization from the amide proton (H^N) via the amide nitrogen (N) to the carbonyl C' (CO) of the previous residue is called HNCO. If another magnetization transfer step to the C_α (CA) of the previous residue is included and the corresponding C_α chemical shift is recorded, the experiment is referred to as H(N)COCA. The parentheses indicate that magnetization is only transferred via the nitrogen spin, without chemical shift evolution taking place. Table 2.3 and table 2.4 provides a summary of triple resonance experiments for the backbone assignments and all the parameter sets for the recorded NMR experiments.

Further developments involved including the chemical shifts of side-chain carbon and proton spins to achieve the sequential assignment. In these experiments chemical shifts of side chain resonances are correlated to the amide proton by combinations of the 3D experiments HNCACB and HN(CO)CACB [Wittekind et al., 1993]. Informations about the chemical shifts of the C_α and C_β carbons are especially valuable for the assignment process, since they are characteristic of the different types of amino acids and can therefore help to position a sequentially connected stretch of amino acids within the known primary sequence of the protein [Grzesiek et al., 1993a].

However, for proteins > 15 kDa the HNCACB experiment become less sensitive. Then the more sensitive experiments HNCA and HN(CO)CA can be used in addition to establish the sequential connectivities. If C_α and C_β chemical shifts obtained from these four experiments still leave some ambiguities, the pair of HNCO and HN(CA)CO can be used to resolve the overlap [Clubb et al., 1992; Engelke et al., 1995].

Since, the HN(CA)CO experiment is quite insensitive, this approach will be useful only in combination with a deuterated protein. The set of six backbone resonance assignment experiments shown in Table 2.3 allowed the unambiguous assignment for SKK. In case of ambiguity, NOESY experiments were used additionally. Backbone assignments of SKK domain was completed using a combination of 3D- ^{15}N -HSQC-NOESY and 3D- ^{15}N -NNH-NOESY spectra to trace the sequential connectivities between amide protons.

Table 2.4 Summary and parameters of all NMR experiments recorded for the interaction studies and for the backbone assignment of SKK domain.

Dimension	Resonance	NS	SW (Hz)	TD	SF0 (MHz)	Reference
¹H¹⁵N-HSQC						
F1	¹⁵ N		1155.54	192		
F2	¹ H	16/32	7788.16	1024	600	[1, 2]
¹H¹⁵N-HSQC						
F1	¹⁵ N		1905.49	256		
F2	¹ H	16/32	10416.67	1024	800	[1, 2]
TROSY						
F1	¹⁵ N		1864.98	256		
F2	¹ H	8	11160.71	1024	800	[3]
HNCO						
F1	¹³ C		2817.19	80		
F2	¹⁵ N		1864.90	64		
F3	¹ H	8	10416.67	1024	800	[4]
HNCA						
F1	¹³ C		6036.12	96		
F2	¹⁵ N		1864.98	64		
F3	¹ H	16	10416.67	1024	800	[4]
HNCACB						
F1	¹³ C		13078.08	128		
F2	¹⁵ N		1864.98	64		
F3	¹ H	16	11160.71	1024	800	[5]

Dimension	Resonance	NS	SW (Hz)	TD	SF0 (MHz)	Reference
HN(CO)CACB						
F1	¹³ C		13078.08	128		
F2	¹⁵ N		1864.98	64		
F3	¹ H	16	11160.71	1024	800	[4]
HN(CA)CO						
F1	¹³ C		2615.96	64		
F2	¹⁵ N		2108.24	64		
F3	¹ H	16	10416.67	1024	800	[6]
HN(CO)CA						
F1	¹³ C		6036.12	72		
F2	¹⁵ N		1864.98	64		
F3	¹ H	8	11160.71	1024	800	[7]
¹H¹⁵N¹H-NOESY						
F1	¹ H		10401.74	256		
F2	¹⁵ N		1864.98	64		
F3	¹ H	8	10416.67	1024	700	[8]
¹⁵N¹⁵N¹H-NOESY						
F1	¹⁵ N		1864.98	64		
F2	¹⁵ N		1864.98	64		
F3	¹ H	32	10416.67	1024	700	[9]

SW = spectral width (Hz) in the observe dimension; TD = total number of data points being acquired; SF0 = spectrometer frequency used; NS = number of scans. (1) [Mori et al., 1995]; (2) [Vuister et al., 1992]; (3) [Kojima et al., 2000]; (4) [Grzesiek et al., 1992b]; (5) [Wittekind et al., 1993]; (6) [Clubb et al., 1992]; (7) [Bax et al., 1991]; (8) [Sekhar et al., 1996]; (9) [Ikura et al., 1990].

2.14.5 Protein-Protein, Protein-RNA interaction studies

NMR is very well suited to the study of protein-protein and protein-RNA interactions. To study the interactions we carried out titrations, because this allows, in addition to the mapping of the interface, a good estimation of the affinity, stoichiometry, and specificity of binding as well as the kinetics of binding [Zuiderweg 2002].

In a nutshell, the ^{15}N - ^1H HSQC or ^{15}N - ^1H TROSY spectrum of one protein is monitored when the unlabeled interaction partner is titrated in, and the perturbations of the chemical shifts are recorded. The interaction causes environmental changes on the protein interfaces and, hence, affect the chemical shifts of the nuclei in this area [Pellacchia et al., 2000; Steven et al., 2001].

The chemical shifts of the labeled protein change during the titration is determined by the kinetics of the interaction. If the complex dissociation is very fast, then the resonances of the nuclei at the interface move in a continuous fashion during the titration. This regime is referred to as “fast chemical exchange” and is often observed for weaker interactions [Hall et al., 2001]. The trajectories of the shifting resonances in fast exchange are informative. If all two-dimensional trajectories are linear and occur at the same rate, a single binding event is indicated. If the trajectories for different resonances occur at a different rate, and/or if they are curved, more than one binding site is implicated.

If the complex dissociation is very slow, we observe one set of resonances for the free protein and one set for the bound protein. During the titration, the “free set” will disappear and will be replaced by the bound set. Most of the resonances of the two sets will overlap with each other, but the differences will mark the interaction interface. This regime is referred to as “slow chemical exchange”. In slow exchange one does not automatically know to which new location the resonance has moved, unless one carries out an independent assignment procedure for the bound state. Consequently, we cannot easily quantitate the degree of change. One approach for this problem is to assume that the new resonance which appears closest to the “free” resonance corresponds to its bound state [Williamson et al., 1997; Muskett et al., 1998]. In the slow exchange case, the binding constant can still be quantitated by measuring the intensities of the disappearing and/or appearing peaks as a function of the titration progression [Van nulan et al., 1993].

In the “intermediate chemical exchange” the frequencies of the changing resonances become poorly defined, and extensive kinetic broadening sets in [Zuiderweg et al., 1981]. If the lines become broad enough, the resonances may disappear from the NMR spectrum. Here, the interaction interface become delineated by progressively disappearing resonances.

Depending on the nature of the exchange process, the different environments undergoing exchange will be characterized by different values of the NMR parameters: chemical shift, coupling constant and relaxation rates. If the measured variable is chemical shift, with exchange occurring between two environments characterized by shifts δ_A and δ_B , then the three exchange regimes are defined by [Jeremy 1995].

Slow exchange	$K \ll \delta_A - \delta_B $
Intermediate exchange	$K \approx \delta_A - \delta_B $
Fast exchange	$K \gg \delta_A - \delta_B $

Generally a rule of thumb is that the interaction with $K_d < 10 \mu\text{M}$ are in slow exchange and intermediate/fast exchange otherwise. However, there are always many exceptions.

During this study ^1H - ^{15}N labeled samples of Nun-full length, Nun N-terminal domain, Nun C-terminal domain, NusA ar1, NusG, NusB, and ^2H - ^{15}N labeled sample of SKK domain were used to study the interaction with their respective interactive partners. The most common experiment carried out with these labeled proteins was ^1H - ^{15}N -heteronuclear single quantum correlation experiment (2.14.4.1).

To investigate the interaction of Nun with other Nus factors, a series of [^{15}N , ^1H]-HSQC spectra was recorded upon gradual addition of the respective interacting partner to a 3 to 4 fold molar excess. The ^1H - ^{15}N resonance shift, disappearing of signals and line width changes were monitored by analyzing the spectra. The titrations were carried out until no further changes could be observed in the spectra or in other words till the point of saturation.

On regard of SKK domain, to increase the spectral resolution and to decrease the relaxation rates of many of the nuclei during triple resonance experiments, (^2H , ^{13}C , ^{15}N) labeled SKK domain samples were prepared. For the backbone assignment and to study the interaction with RNA and other proteins for SKK domain, the methods were combined with TROSY.

2.14.6 Chemical shift mapping

Chemical shift mapping is used to identify putative sites of interaction on a protein surface by detecting chemical shift perturbations in simple $^1\text{H},^{15}\text{N}$ -HSQC NMR spectra of a uniformly labeled protein as a function of added (unlabeled) target protein. Information on the backbone resonance assignment and the protein structure (or a homology based model) for a protein is a prerequisite for chemical shift mapping. The identity and location of resonances that undergo binding-dependent chemical shift perturbations are mapped onto the three-dimensional structure of the protein to yield the binding site for the partner protein [Jeremy 1995; Rajagopal et al., 1997].

Ligand binding causes change in the electronic environment of the residues which are in the vicinity of the ligand. This changed environment induces change in the chemical shift for these residues in the $^1\text{H},^{15}\text{N}$ -HSQC or $^1\text{H},^{15}\text{N}$ -TROSY experiments. When the assignments of the free protein were available, they can be readily transferred to the complex by tracking the changes that occur during the titration.

Normalized chemical-shift changes are expressed as the weighted geometric average of $^1\text{H}^{\text{N}}$ and ^{15}N chemical shift changes for each residue.

$$\Delta\delta_{norm} = \sqrt{(\Delta\delta_{1H})^2 + 0.1(\Delta\delta_{15N})^2} \quad [2.2]$$

$\Delta\delta(\text{X})$ represents the chemical shift difference of spin X between free and bound states. The analysis of perturbation spectra is usually focused on the difference between the perturbed (X) and the non-perturbed (Y) spectra. The cross peaks were picked from the Y spectra and used to define the integration areas of the X spectra. For each X spectra, the integration of data points from these areas are compared with the corresponding ones from the Y spectra to calculate the similarity (correlation coefficient) between the X and the Y spectra. If the shift changes are mapped on to the protein structure, a clear surface patch of affected residues is generally observed, and this indicates the location of the binding site.

Due to the high sensitivity of $^1\text{H},^{15}\text{N}$ -HSQC and $^1\text{H},^{15}\text{N}$ -TROSY experiment, these investigations are frequently used in drug discovery-related ligand binding studies like for example, Structure Activity Relationships (SAR) by NMR [Shuker et al., 1996].

2.14.7 Dissociation constant

The dissociation constant K_D , in the simplest case of a protein with a single binding site is defined as follows

$$K_D = [P][L] / [PL] \quad [2.3]$$

where $[P]$, $[L]$ and $[PL]$ are the equilibrium concentrations of protein, ligand and complexed state, respectively. A value of K_D in the mM range implies an approximately 1:1000 ratio of free to bound states in an equimolar mixture of P and L and a K_D in the μM range implies an approximately 1:10,00,000 ratio of these states, i.e., a much more stable complex with less of the ‘free’ species present [Fielding et al., 2007].

To measure K_D by means of NMR experiments implies quantitative analysis of solutions that are potentially μM in the observed nucleus. The significance of K_D is that ligands of weaker affinity have larger K_D and thus require the addition of more ligand to saturate the receptor binding site [Christopher et al., 2004].

The dissociation constant K_D is determined from the changes in chemical shifts of ^{15}N -labeled SKK domain in ^1H - ^{15}N TROSY after gradual addition of the corresponding unlabeled binding partner. In the fast exchange region the observed chemical shift represents the population average of the chemical shift of the free and bound state. Therefore, the observed chemical shift is described by the equation 2.4 and can be used for K_D determination [Morton et al., 1996].

$$\delta_{obs} = \delta_P + (\delta_{PL} - \delta_P) \left[\frac{\{K_D + (1+r)[P]_0\}}{2[P]_0} - \frac{\sqrt{(K_D + (1+r)[P]_0)^2 - 4[P]_0^2 r}}{2[P]_0} \right] \quad [2.4]$$

where δ_{obs} , δ_P , and δ_{PL} are the chemical shifts for the actual mixture, the free protein, and the completely bound protein, respectively. $[P]_0$ is the total concentration of SKK domain, and r describes the SKK/RNA or SKK/NusA ar2 ratio. The curves were fitted using the program MATLAB (6.0.0.88 Release 12, The Math Works Inc., Natick, Massachusetts, USA) by using the script `kd_fit_func2.m` (Appendix 9.1) (in-house written script by Dr. Kristian Schweimer, Department of Biopolymers, University of Bayreuth).

3 Experiments and Results

3.1 Expression and purification of Nun constructs

3.1.1 Expression and purification of Nun (1-112)

The nucleotide sequence and the physical/chemical parameters of Nun encoding 1-112 amino acids are represented in Appendix 9.2. The overexpression and purification of recombinant Nun was performed as described in section 2.5.1/2.5.2. Overexpression of Nun full length was analyzed by 19 % SDS-PAGE (Fig 3.1). From the gel we can observe that the expression rate for Nun was efficient.

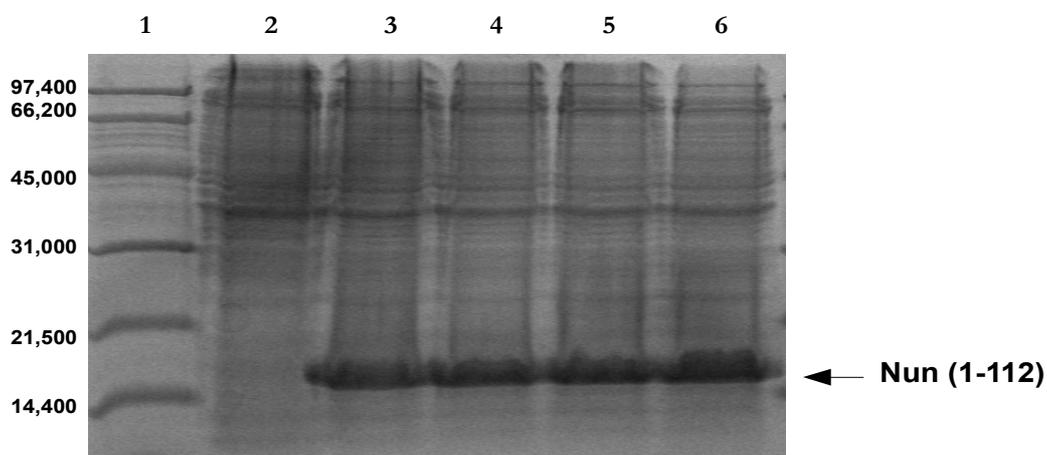
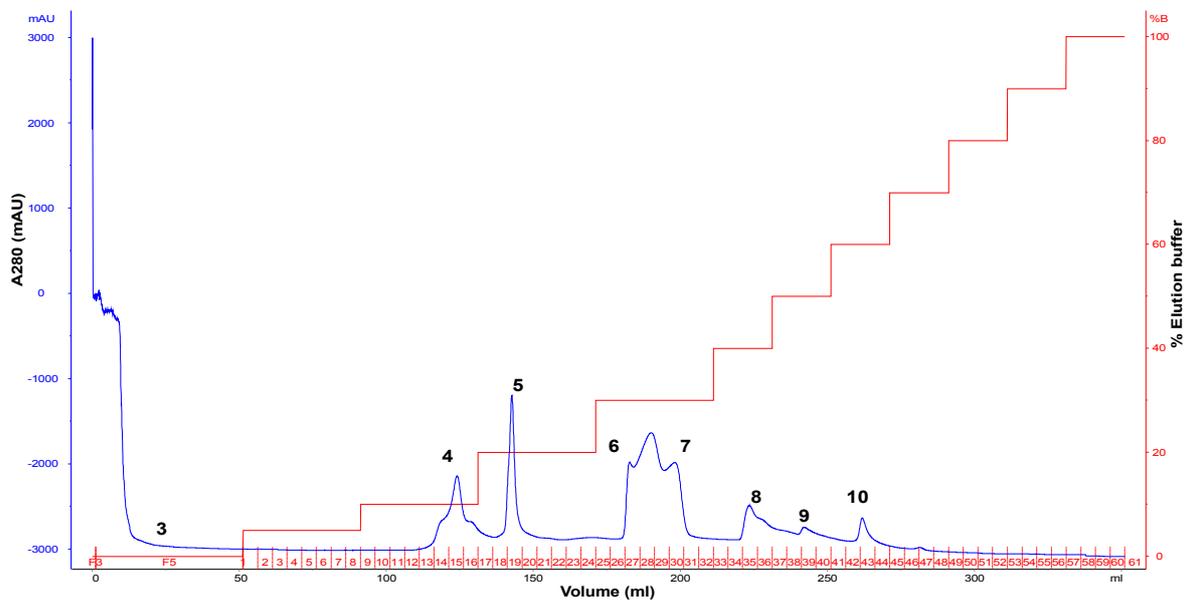


Figure 3.1 Expression of HK022-Nun (1-112) in *E. coli* BL21 (DE3). The overexpression of recombinant HK022-Nun protein was induced by 1 mM IPTG. The whole cell extract were subjected to SDS-PAGE and stained with Coomassie Brilliant Blue. As molecular weight standard, low range molecular weight marker from Biorad was laid on the gel. Lane 1, protein marker; Lane 2, uninduced state; Lane 3, 1 hour after induction; Lane 4, 2 hours after induction; Lane 5, 3 hours after induction; Lane 6, 4 hours after induction.

Following the expression, the cell lysis and purification were performed as described in section 2.5.2. Nun was purified by using cation exchange chromatography applying a step gradient elution (Fig 3.2 (A)). Fractions containing Nun were identified on a 19 % SDS-PAGE (Fig 3.2 (B)). We had observed that the Nun protein elute from 20 % till 60 % gradient step. Among all the fractions, the eluate at 30 % gradient was more pure compared to the other fractions.

(A)



(B)

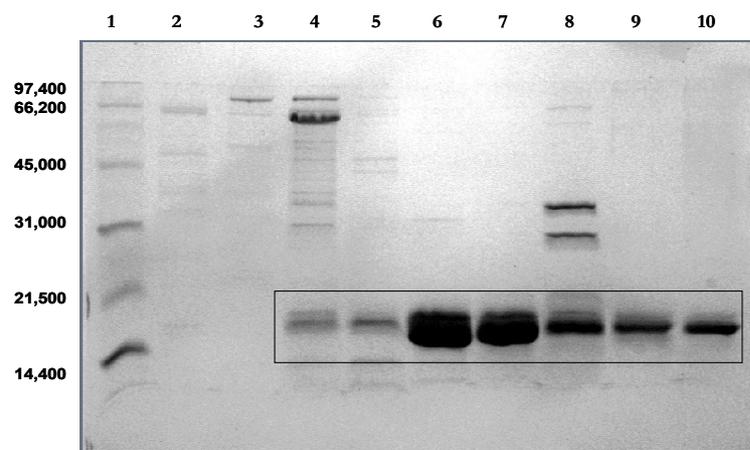
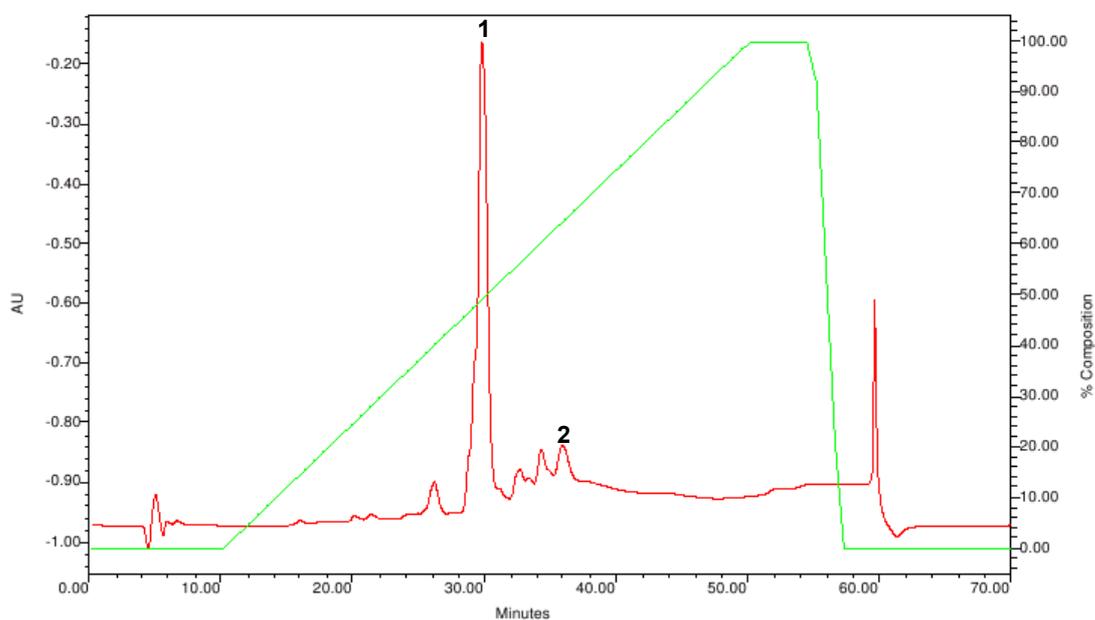


Figure 3.2 Purification of Nun (1-112) using cation exchange chromatography. (A)-Chromatogram of the purification monitored at A_{280} . The numbers in the chromatogram corresponds to the lane in the SDS-PAGE. (B)-Lane 1, protein marker; Lane 2, flow through; Lane 3, wash; Lane 5-10 represents the eluates.

The Nun fractions which were not pure after the cation exchange chromatography were pooled and further subjected to HPLC for purification. The HPLC run was carried out as mentioned in the section 2.5.2. The elution of Nun protein from the HPLC column was performed by applying a slow linear gradient (Fig 3.3 (A)). The peak fractions eluted were analyzed by 19 % SDS-PAGE (Fig 3.3 (B)). The eluted fractions from HPLC column was sufficiently pure for further studies.

(A)



(B)

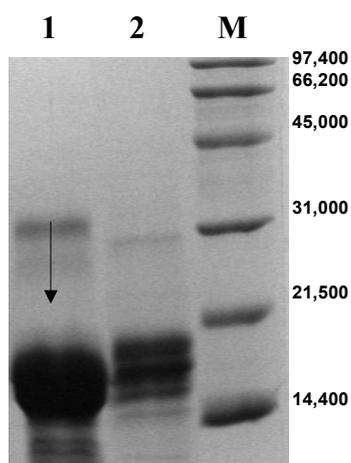


Figure 3.3 (A) Chromatogram of the HPLC monitored at A_{280} . The numbers in the chromatogram correspond to the bands assigned in SDS-polyacrylamide gel. (B) The alphabet (M) stands for protein marker; Lane 1 and 2 shows the eluate from the HPLC run.

3.1.2 Expression and purification of Nun (45-112)

The nucleotide sequence of the Nun C-terminal domain (CTD) is shown in Appendix 9.3. A detailed protocol about the expression and purification of Nun CTD is mentioned in section 2.6.1/2.6.2. Nun CTD was cloned into the expression vector pET-GB1. The fusion protein carried a hexa-histidine-tag at the N-terminus and with a TEV protease cleavage site between GB1 and Nun CTD. The expression was achieved by inducing the cells with 1 mM IPTG. The efficiency of the expression of Nun CTD was visualized on a 19 % SDS-PAGE (Fig 3.4)

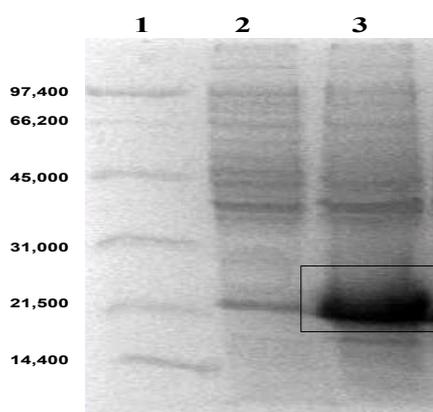
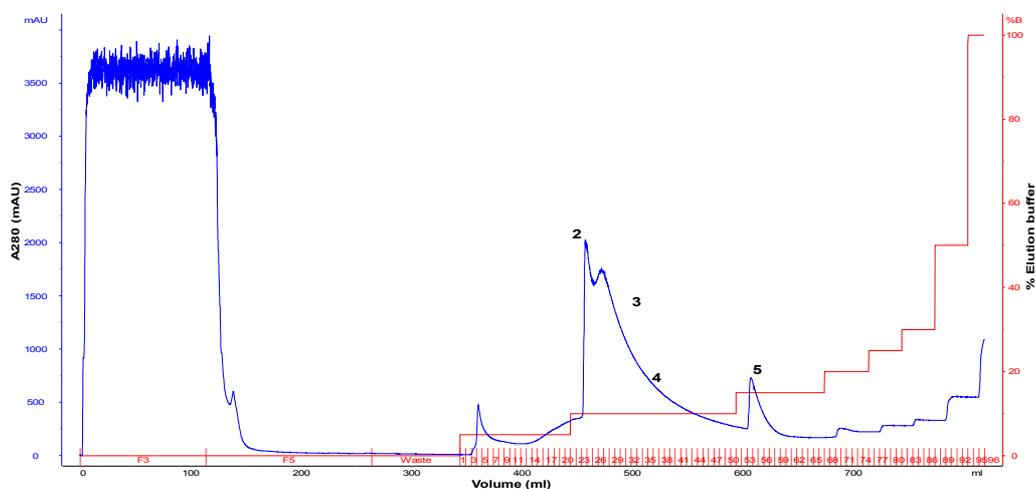


Figure 3.4 Expression of Nun CTD. Lane 1, Biorad marker; Lane 2, Uninduced Nun CTD; Lane 3, 3.5 hours after induction.

Expressed recombinant protein was purified by nickel affinity chromatography. In eluting the protein, a step gradient ranging from 0-500 mM imidazole was used (Fig 3.5 (A)). The eluted fractions were analyzed on a 19 % SDS-PAGE (Fig 3.5.(B)). The Nun CTD eluted at 10 % gradient was about 90 to 95 % pure.

(A)



(B)

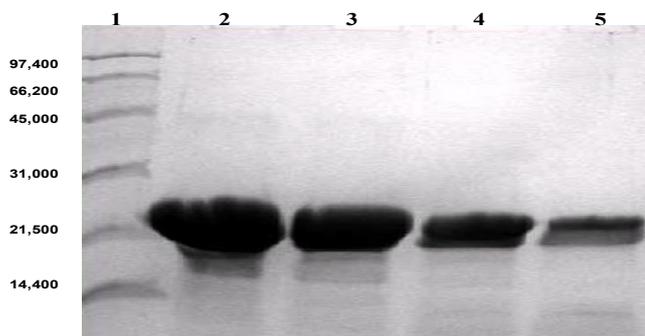


Figure 3.5 (A)-Chromatogram of the purification of Nun CTD monitored at A_{280} . The numbers in the chromatogram corresponds to the lane in the SDS-PAGE. (B)-Lane 1, protein marker; Lane 2-4 are the eluted fractions from the Histrap column.

The pure fractions containing Nun CTD were pooled and then subjected to TEV protease cleavage. After the cleavage process, the sample was further loaded onto the Histrap column to collect the pure Nun CTD in the flow through. The purity was observed by loading the flow through on a Schagger-Jagow gel. (Fig 3.6)

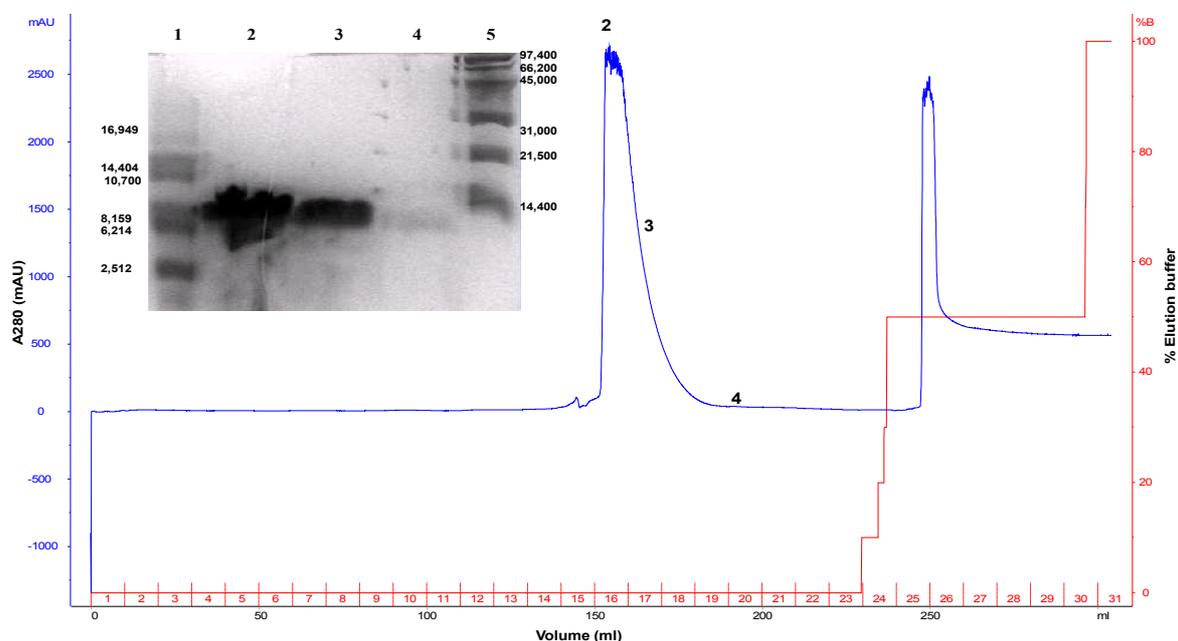


Figure 3.6 Final purification of Nun CTD after the TEV protease cleavage. The numbers in the chromatogram corresponds to the lane in the SDS-PAGE. Lane 1, Fluka molecular weight marker; Lane 2-4 are the eluted fractions from the Histrap column; Lane 5, Biorad marker.

3.2 Expression and purification of NusA and its domains

3.2.1 Overexpression and purification of NusA

E. coli BL21 (DE3)/pTKK19 NusA (1-495) (Appendix 9.4) was expressed by growing the cells in LB medium at 37 °C. The procedure for overexpression and purification is explained in detail in section 2.7.1/2.7.2. The expression state was analyzed by applying the sample of cell pellets after four hours of induction on a 10 % SDS gel (Fig 3.7). The NusA (1-495) has a histidine affinity tag for the purification, hence, the sample was purified by using a Histrap column. The eluted fractions were analyzed on 10 % SDS-PAGE (Fig 3.8). The protein eluted from the Histrap column at 20 % gradient was absolutely pure for further studies. The peak fractions were pooled and subjected for cleavage by PreScission protease. After the cleavage the sample was applied onto a Histrap column and the NusA was obtained from the flow through.

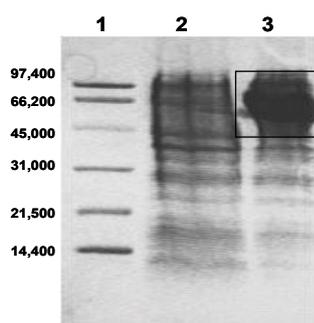


Figure 3.7 Expression of NusA (1-495). Lane 1, Biorad marker; Lane 2, before induction; Lane 3, after induction.

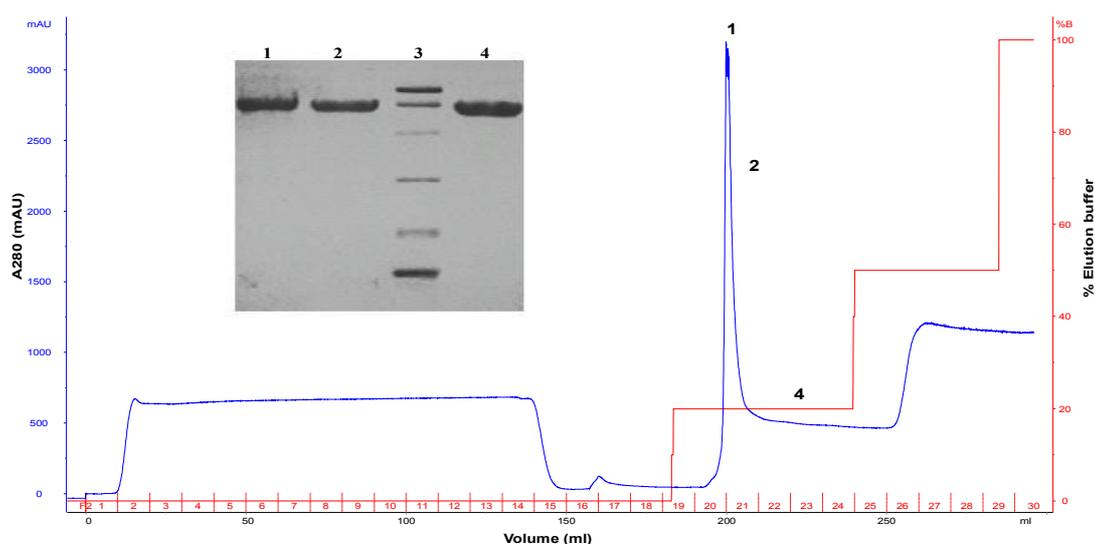


Figure 3.8 Peak fractions of NusA (1-495) after the Histrap column. The numbers in the chromatogram represents the corresponding lane in the gel. Lane 3, is loaded with Biorad marker; Lane 1, 2 and 4 are the eluted fractions.

3.2.2 Expression and purification of NusA ar1

The ar1 domain of NusA (339-426) was expressed in *E. coli* BL21 (DE3) using the pET19b expression vector. The nucleotide sequence and the physical/chemical parameters is shown in Appendix 9.5. The expression and purification was carried out as described before in the section 2.8.1/2.8.2. The efficiency of the expression was monitored by Schagger-Jagow gel (Fig 3.9). For ^{15}N labeling of the protein, cells were grown in M9 minimal medium with ^{15}N NH_4Cl as the sole nitrogen source.

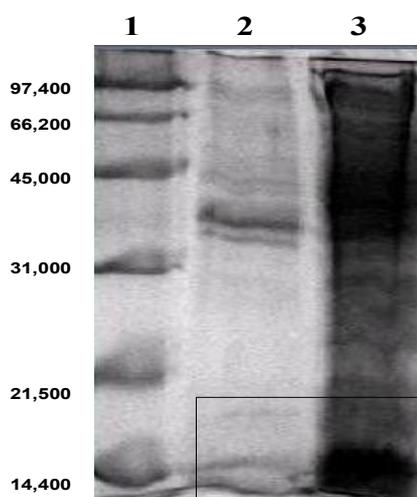


Figure 3.9 Expression of NusA ar1 in *E. coli* BL21 (DE3). Lane 1, Biorad marker; Lane 2, uninduced; Lane 3, after induction by 1 mM IPTG.

The resulting soluble recombinant protein contained a N-terminal deca-His tag. To purify NusA (339-426), the supernatant after sonication was loaded onto a Ni-ion-affinity column and eluted by applying an imidazole step gradient (Fig 3.10). All the eluted fractions were analyzed by Schagger-Jagow gel and visualized by Coomassie staining. NusA ar1 was eluted at 15 % gradient.

The eluted fractions containing NusA ar1 were pooled together and set to cleave the N-terminal deca-His-tag using PreScission protease. The cleaved sample was dialyzed against the required buffer and further purified to get rid of the cleaved tag via a QXL column by a NaCl step gradient with up to 1 M NaCl (Fig 3.11). The eluted fractions containing NusA (339-426) were pooled and concentrated with Vivaspin concentrators for NMR studies.

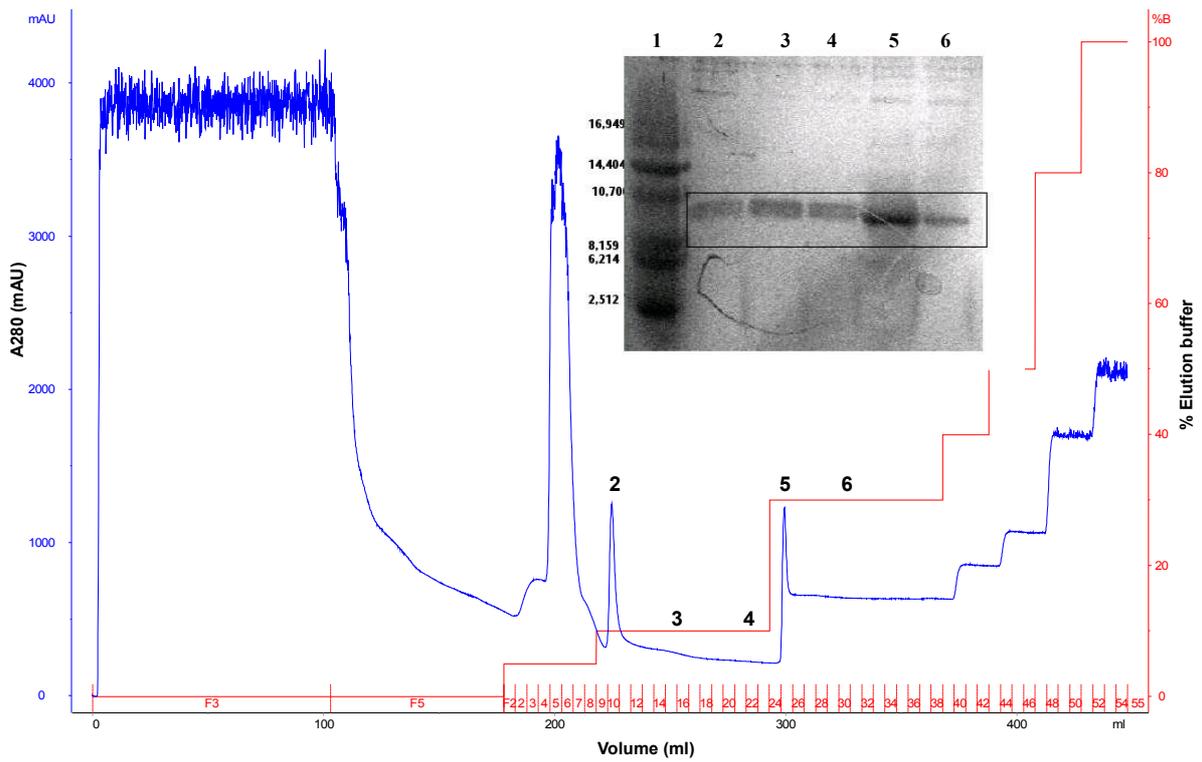


Figure 3.10 NusA ar1 (339-426) after the Histrap column. The numbers in the chromatogram represents the corresponding lane in the gel. Lane 1, Fluka marker; Lane 2-6, eluates.

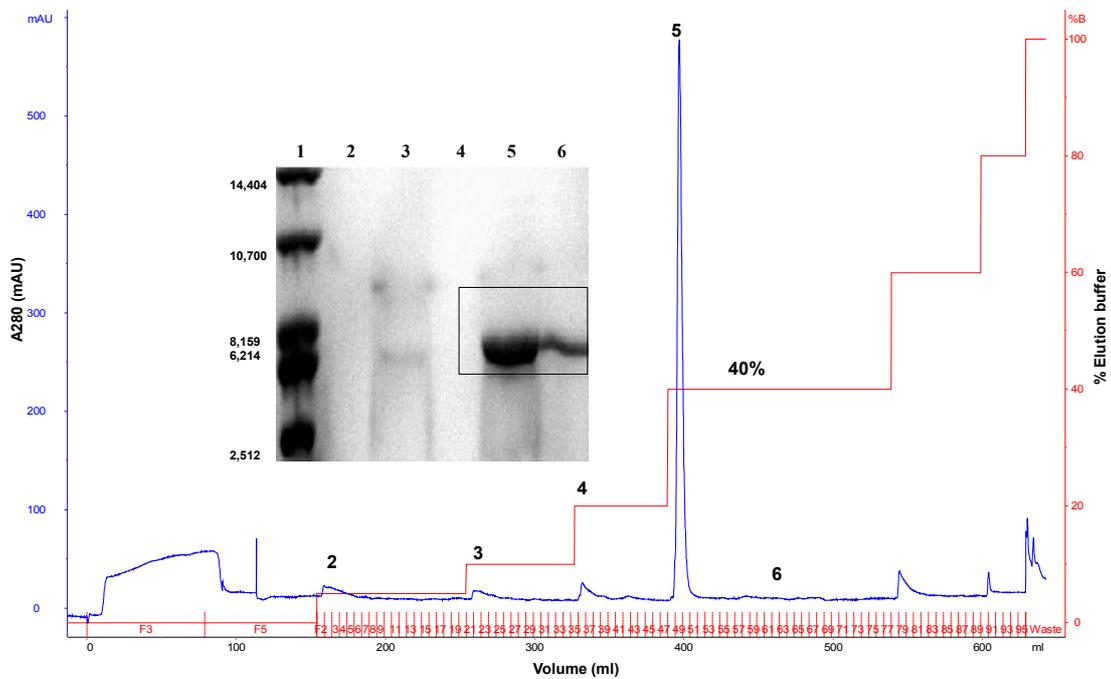


Figure 3.11 NusA ar1 (339-426) purified via QXL column after the cleavage. The numbers in the chromatogram shows the corresponding lane in the gel. Lane 1, marker; Lane 2-6, eluates.

3.2.3 Expression and purification of NusA ar2

The nucleotide sequence of acidic repeat 2 of NusA (424-495) is shown in Appendix 9.6. Expression and purification of the protein was carried out as mentioned in the section 2.9.1/2.9.2. Purification was carried out by a step gradient using nickel ion chromatography. The eluted fractions were analyzed by Schagger-Jagow gel (Fig 3.12). Subsequent cleavage by PreScission protease and a further QXL step yielded a pure protein at 50 % gradient elution (Fig 3.13).

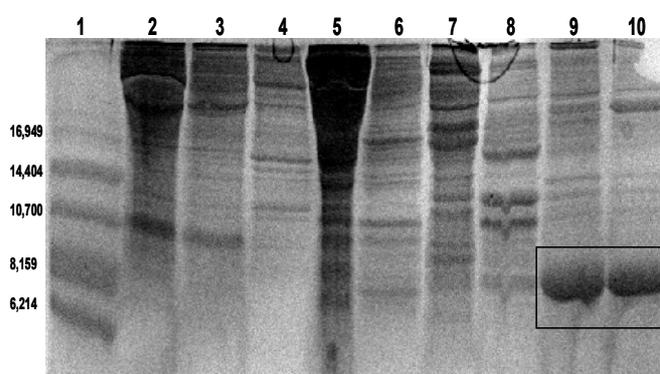


Figure 3.12 Purification of NusA ar2. Lane 1, Fluka marker; Lane 2-10 are the various fractions eluted during the nickel ion affinity chromatography step.

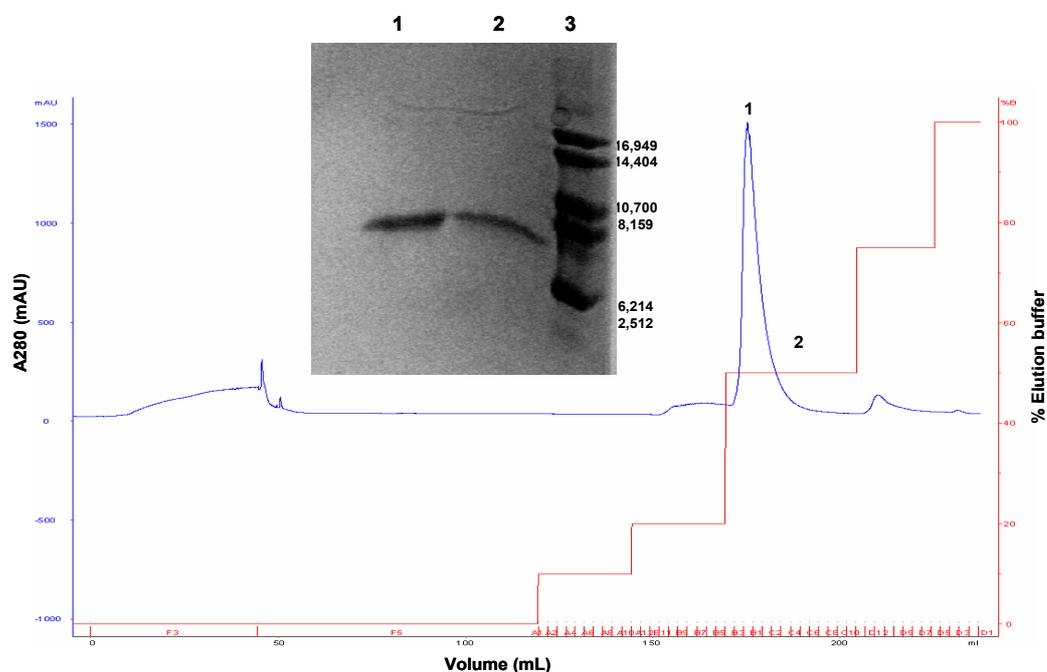


Figure 3.13 Peak fractions containing pure NusA ar2 after the QXL run. The numbers in the chromatogram corresponds to the lane in the Schagger-Jagow gel. Lane 1-2, eluted fractions from the QXL column; Lane 3, Fluka marker.

3.2.4 Expression and purification of SKK domain

The RNA binding domain of NusA consist of central S1 and two KH domains respectively. This domain is referred in short as SKK domain. The nucleotide sequence and the physical/chemical parameters of SKK domain is represented in Appendix 9.7. SKK was expressed in *E. coli* BL21 (DE3) using the pET11a expression vector. The resulting protein consisted of an amino-terminal hexa-histidine-tag followed by SKK (132-348). Unlabeled and labeled SKK protein was obtained as described in the section 2.10.1.

In case of SKK domain, the expression efficiency was increased, when transformation was carried out every time for the protein preparation and not inoculating by using a glycerol stock. In this study, deuterium labeling of SKK domain has been used to improve the resolution and sensitivity of NMR spectra in triple resonance experiments and as well as in the interaction studies.

The labeling strategy results in deuterium incorporation throughout a protein in a roughly site-independent manner (uniform or random labeling). The main problem to express a deuterated protein is the incorporation of ^2H which reduces growth rate of organisms up to 50 % and decreases protein production as a consequence of the isotopic effect. Hence, deuterium labeling requires conditions different with respect to ^{13}C and ^{15}N enrichment and requires bacteria adaptation. The expression was achieved by inducing the cells with 0.1 mM IPTG. The expression was monitored by applying the sample to a 19 % SDS gel (Fig 3.14). The expression rate was quite effective as seen from the gel.

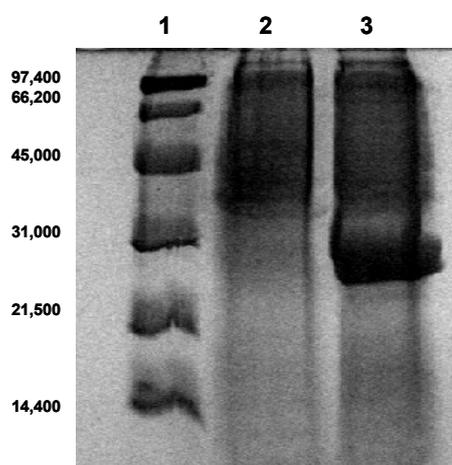


Figure 3.14 Overexpression of SKK domain in *E. coli* BL21 (DE3). Lane 1, Biorad marker; Lane 2, before induction; Lane 3, induction by 0.1 mM IPTG.

Purification of the soluble his-tagged protein from the clarified crude cell extract was performed by nickel ion affinity chromatography on an ÄKTA purifier 10-FPLC system as described in section 2.10.2. The supernatant was loaded onto a 5 ml HisTrap column and the bound protein was eluted using a stepwise gradient of imidazole. The collected fractions were analyzed for its purity on a 19 % SDS-PAGE (Fig 3.15). Pure protein was eluted at 20 % gradient. Fractions containing the protein was combined and dialyzed against the required buffer and concentrated using Vivaspin concentrators. (Fig 3.16).

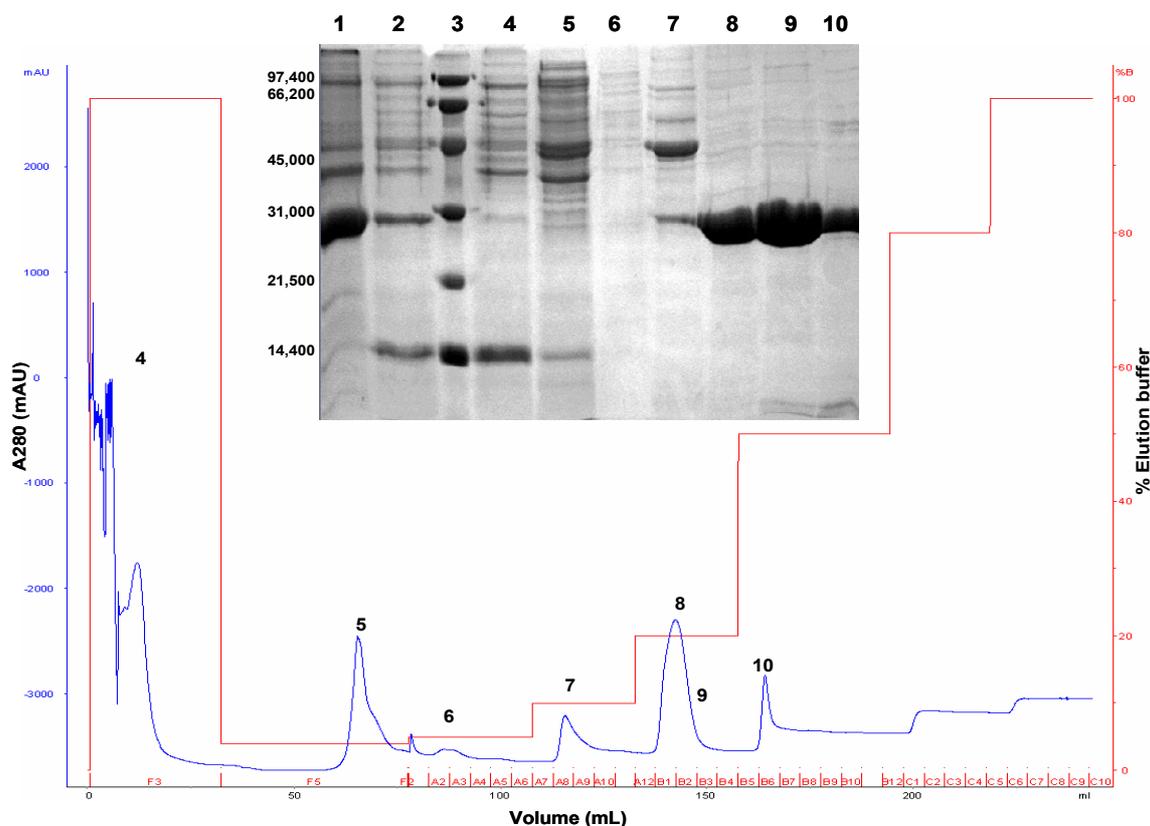


Figure 3.15 19 % SDS-PAGE analysis of the purification of SKK domain. The numbers in the chromatogram corresponds to the lane in the gel. Lane 1, supernatant after sonication; Lane 2, pellet after sonication; Lane 3, Biorad marker; Lane 4-10, eluted fractions from the HisTrap affinity column. Fractions 8-10 shows the pure SKK protein.

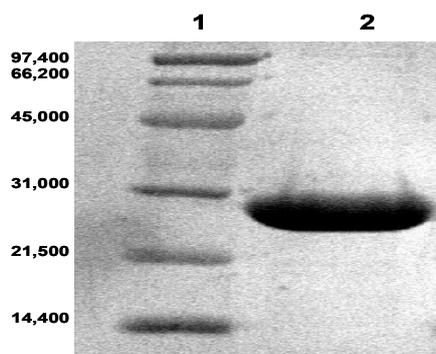


Figure 3.16 Concentrated sample of SKK protein. The purified fractions were pooled together and concentrated to a concentration of 400-500 μ M. Lane 1, Biorad marker; Lane 2, concentrated sample.

Effect of perdeuteration on SKK domain

Deuteration reduces the relaxation rates of NMR-active nuclei, in particular ^{13}C , because the gyromagnetic ratio of ^2H ($\gamma[^2\text{H}]$) is 6.5 times smaller than ^1H ($\gamma[^1\text{H}]$) and thereby it improves the resolution and sensitivity of the NMR experiments. During deuteration, replacement of protons with deuterons removes contributions to proton line widths from proton-proton dipolar relaxation and ^1H - ^1H scalar couplings. Perdeuteration of SKK was necessary to yield spectra in a suitable quality for sequential resonance assignment. Significant improvements have been noted in between non-deuterated and deuterated SKK protein sample as shown in the following figure (Fig 3.17).

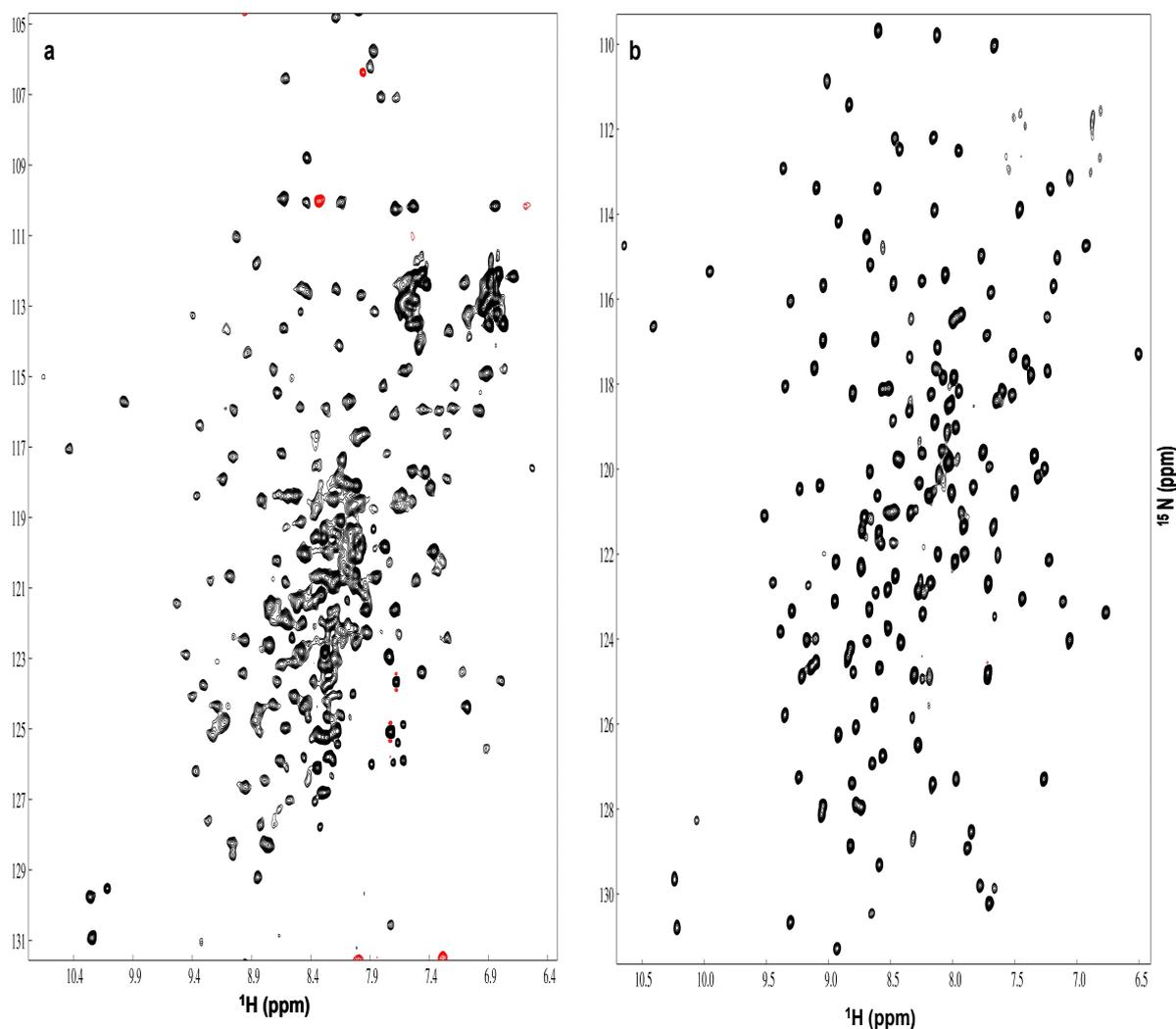


Figure 3.17 Effect of perdeuteration on ^1H - ^{15}N TROSY spectrum of the 24.4 kDa SKK domain of NusA. Both proteins are uniformly ^{15}N labeled, while the deuterated protein is approximately 80-90% uniformly deuterated. a) non-deuterated sample. b) deuterated sample.

3.3 Expression and purification of NusG

Full length NusG was expressed in *E. coli* BL21 (DE3) cells in pET 11a/NusG vector. The protein was constructed with no affinity tag. The expression, cell lysis and protein purification were performed as described in the section 2.11.1/2.11.2. The nucleotide sequence and other parameters concerning NusG construct is shown in Appendix 9.8. After cell lysis, most of the protein was in the soluble fraction. The supernatant was subjected to ammonium sulfate precipitation. The pellet obtained was resuspended and dialyzed against the respective buffer and purified using anion exchange chromatography (QXL) with a linear gradient. The eluted fractions were analyzed by 19 % SDS-PAGE for purity (Fig 3.18). The pure fractions containing NusG were pooled for concentration by Vivaspin concentrators.

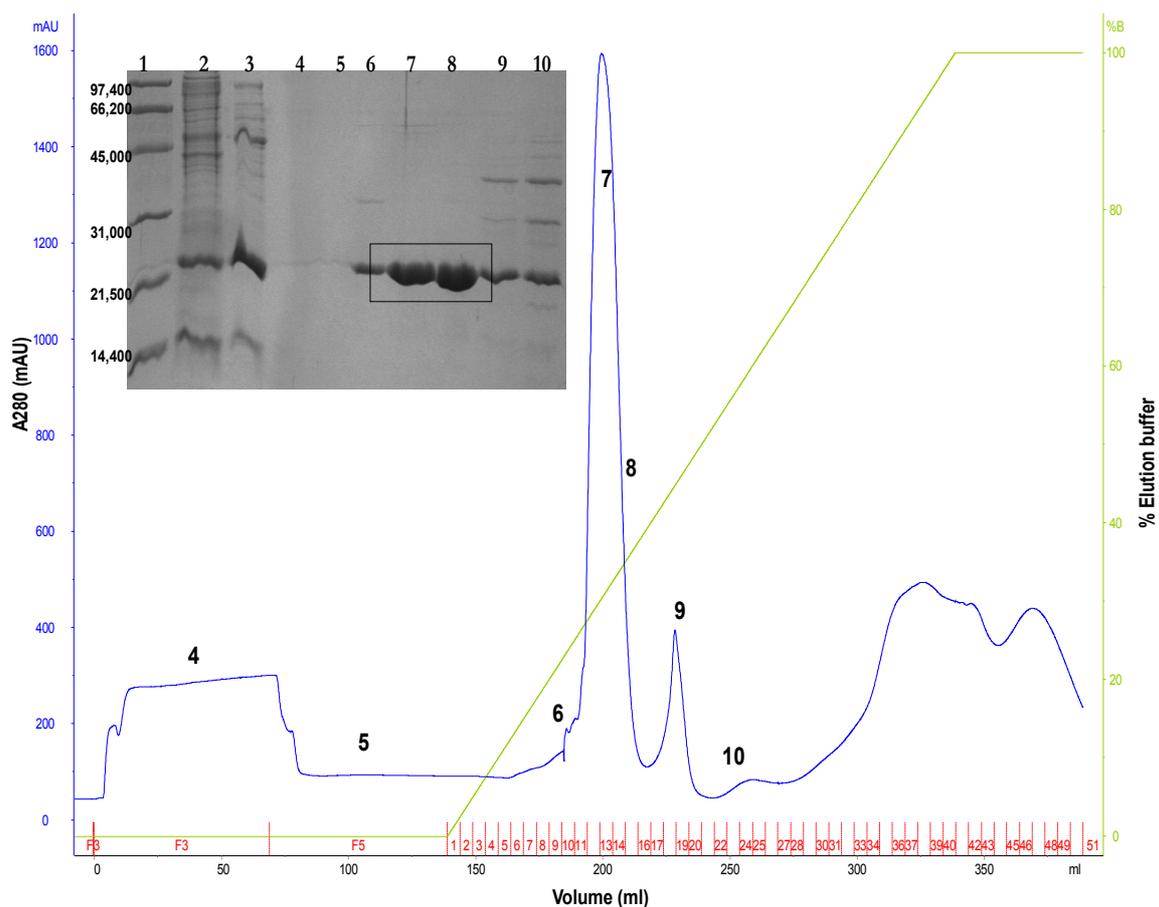


Figure 3.18 SDS-PAGE analysis of NusG purification. The numbers in the chromatogram corresponds to the lane in the SDS-PAGE. Lane 1, marker; Lane 2, pellet after cell lysis; Lane 3, supernatant after cell lysis; Lane 4, flow through; Lane 5, wash; Lane 6-10, fractions eluted from the QXL column.

3.4 Expression and purification of NusB

To produce high yield of labeled protein and to have reproducible results, one promising recent approach has been the development of auto-induction of recombinant protein expression, a method introduced by Studier. This approach is based upon the preferences of bacteria to selectively use different carbon sources during diauxic growth and upon the often-observed negative regulation of gene expression by catabolite repression.

The auto-induction approach has many advantages over traditional induction methods including avoidance of the strong induction and apparent toxicity associated with chemical inducers such as IPTG, the ability to formulate medium compositions to obtain a desired level of cell growth before expression and minimal requirements for handling of the expression culture associated with growth-dependent induction of target protein expression.

Nucleotide sequence of NusB is shown in Appendix 9.9. The expression and purification is explained in detail in section 2.12.1/2.12.2. Auto-induction of NusB was carried out as mentioned by Studier et al., 2004. The protein was purified using a HisTrap affinity column. The bound protein was eluted with a step gradient using a buffer containing imidazole. The fractions eluted were analyzed by 19 % SDS-PAGE (Fig 3.19). The eluted fractions were very pure.

The peak fractions containing NusB were pooled together and set to cleave the N-terminal histidine tag using the TEV protease. The cleaved sample was dialyzed against the required buffer for further purification.

The NusB protein was further purified to high level of purity by applying the cleaved sample to an anion exchange chromatography. The sample was loaded onto the QXL column and then eluted with a step gradient ranging from 0 to 1 M NaCl. The fractions eluted from the QXL column were analyzed by SDS-PAGE (Fig 3.20). The fractions corresponding to NusB were pooled for further use.

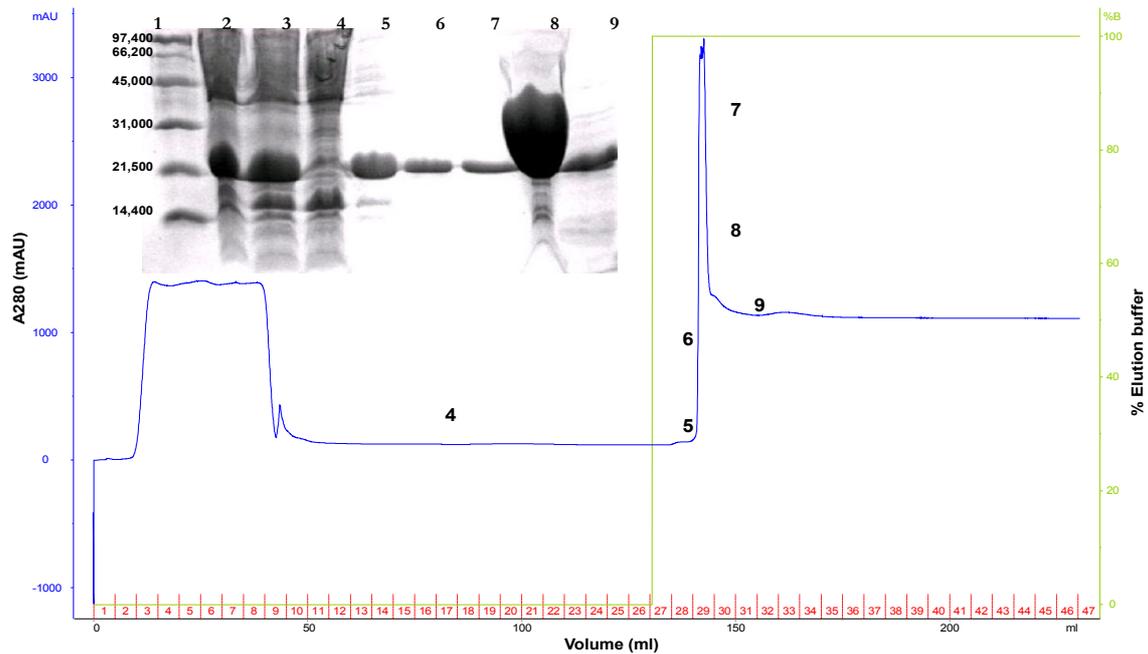


Figure 3.19 Purification analysis of NusB. The numbers in the chromatogram corresponds to the lane in the SDS gel. Lane 1, marker; Lane 2, pellet after cell lysis; Lane 3, supernatant after cell lysis; Lane 4, wash; Lane 5-9, the fractions collected from the Histrap column.

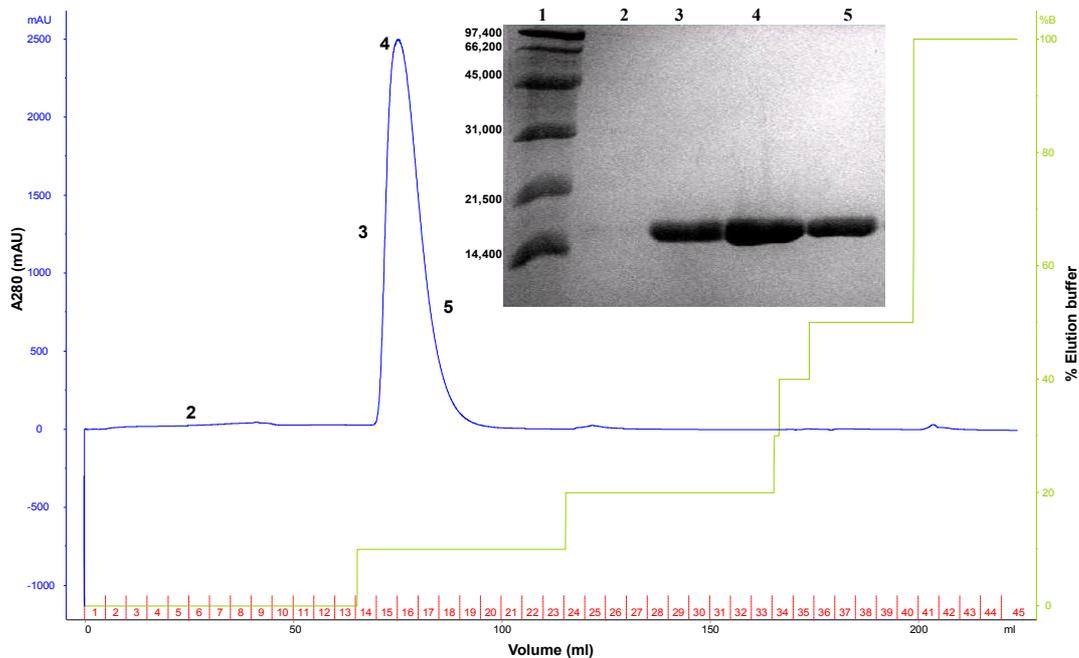


Figure 3.20 Analysis of the fractions eluted from the QXL column as the final purification step of NusB. Lane 1, marker; Lane 2-5, the fractions eluted from the QXL column.

3.5 Interaction of Nun with *E. coli* host factors

3.5.1 Interaction of Nun with NusA ar1 domain

The Nun protein of bacteriophage HK022 is a member of the arginine-rich motif family of RNA binding proteins, which also includes the phage λ N transcription antitermination protein. In contrast to λ N, which suppresses transcriptional termination, Nun terminates transcription just distal to BOXB (1.5.2)

The Nun C terminus interacts with RNA polymerase and contacts DNA template, which is 7-8 bp downstream to the RNA polymerase active center. It is likely, that Nun intercalates into template *via* W108 and this intercalation physically blocks RNA polymerase translocation.

It has been already reported [Watnick et al., 1998] that NusA, a transcription elongation protein binds to the C terminus of Nun and stimulates the Nun binding to *BoxB*. In doing so, however, NusA inhibits the interaction between Nun and RNA polymerase, which requires the C terminus of Nun. The mechanism by which NusA stimulates Nun binding appears to be novel among RNA binding proteins.

The idea, that NusA interacts with the C-terminal region of Nun consisting of amino acids **VMHRVVNHAHQRNPNKKWS** was supported based on the results from a binding assay [Watnick et al., 1998]. By interaction with Nun C-terminus, NusA exposes the RNA binding domain and allows Nun to bind BOXB.

In an other publication by Kim et al., 2006, it has been shown that Nun binding to NusA, like that of λ N, requires NusA ar1 region located between NusA residues 364 and 415. To determine this, they had expressed N-terminal hexahistidine tagged derivatives of full length NusA and four NusA C-terminal deletion mutants. Nun binding to these derivatives were determined by Ni²⁺ affinity chromatography. The results obtained indicated that NusA ar1 is crucial for Nun binding as it is for N.

Based on the above two reported results, one could suggest that NusA ar1 interacts with the C-terminal domain of Nun. On the context of this fact, interaction between Nun and NusA ar1 was studied at atomic level by NMR.

To observe the interaction between these two proteins, ^{15}N -labeled samples of NusA ar1 (2.8) was used for the titration experiments by NMR. In order to identify which amino acid residues of Nun are involved in the interaction with NusA ar1, three constructs of Nun (Nun C-terminal peptide (92-112); Nun CTD (45-112); Nun (1-112) have been used in this study (Appendix 9.2 and 9.3).

Three separate titration experiments were performed in which ^{15}N HSQC spectra of NusA ar1 were collected as a function of added Nun constructs: (1) NusA ar1 + Nun (92-112), (2) NusA ar1 + Nun (45-112), (3) NusA ar1 + Nun (1-112). The NMR titrations were optimized by adopting different buffers, pH and salt concentration. Amide ($^1\text{H}^{\text{N}}$, ^{15}N) chemical shifts are very sensitive to local structural changes. Therefore, observation of chemical shift changes on titration of a binding partner to a ^{15}N labeled protein provides a powerful method for the identification of the interaction and the binding surface.

However, no detectable signal shifting is observed during titration, up to an fourfold molar excess of Nun constructs. Consonant with the lack of chemical shift perturbations (2.14.5) observed in all the three titration experiments, we could conclude that there is a lack of an intermolecular interaction between NusA ar1 and Nun.

An overlay of $^1\text{H},^{15}\text{N}$ -HSQC spectra of three NMR titrations were shown in Fig 3.21; Fig 3.22, and Fig 3.23 respectively. All the spectra shown here clearly depicts that there is no interaction surface and no significant changes in the conformation of NusA ar1 upon titrating with Nun protein.

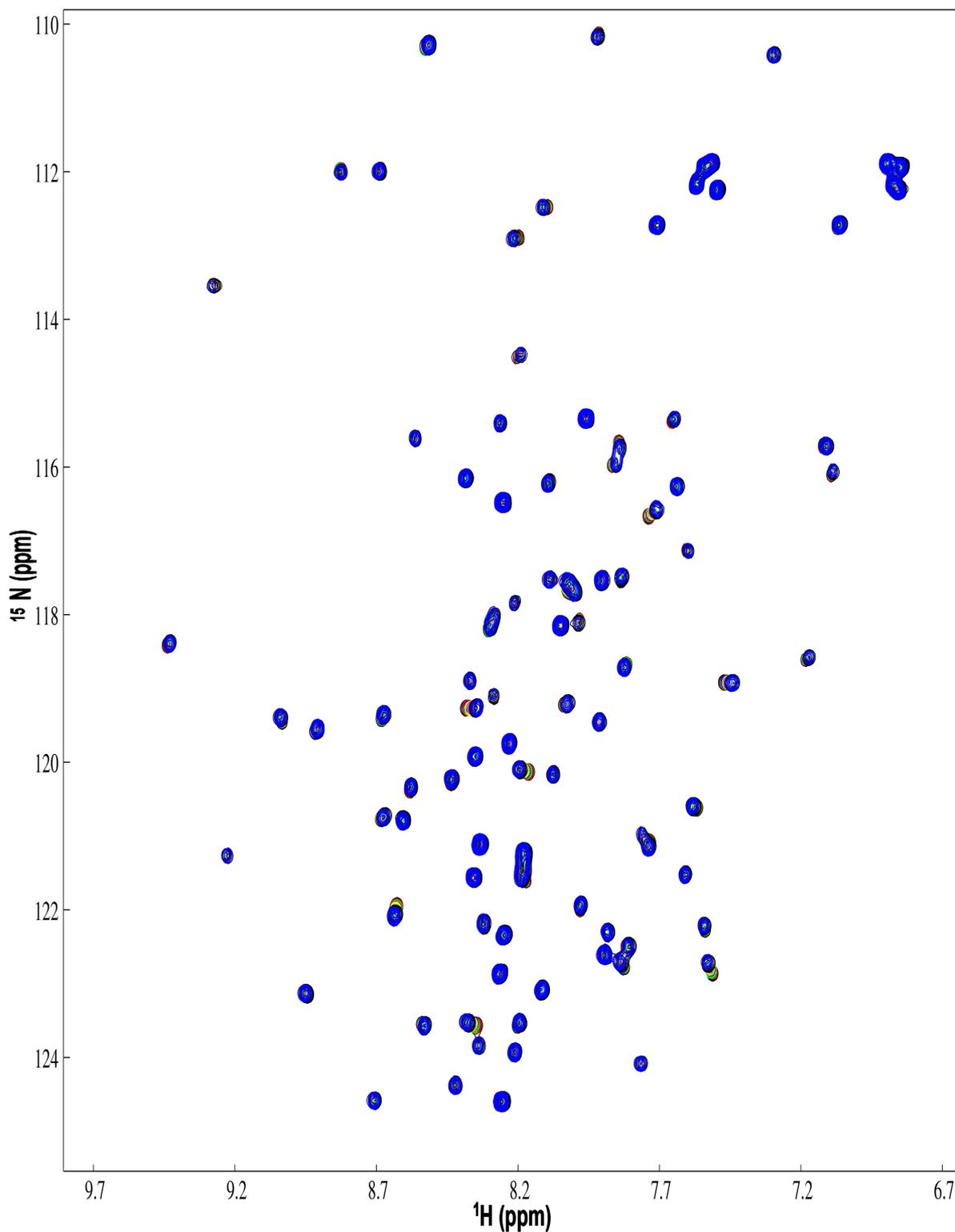


Figure 3.21 Overlay of the ^1H , ^{15}N -HSQC spectra of free NusA ar1 (black) upon titrating with Nun peptide (92-112) (blue). No significant changes are observed in NusA ar1 upon interacting with nun peptide. The appearance of the overlay did not change with other buffers and in the presence and absence of salt in the buffer.

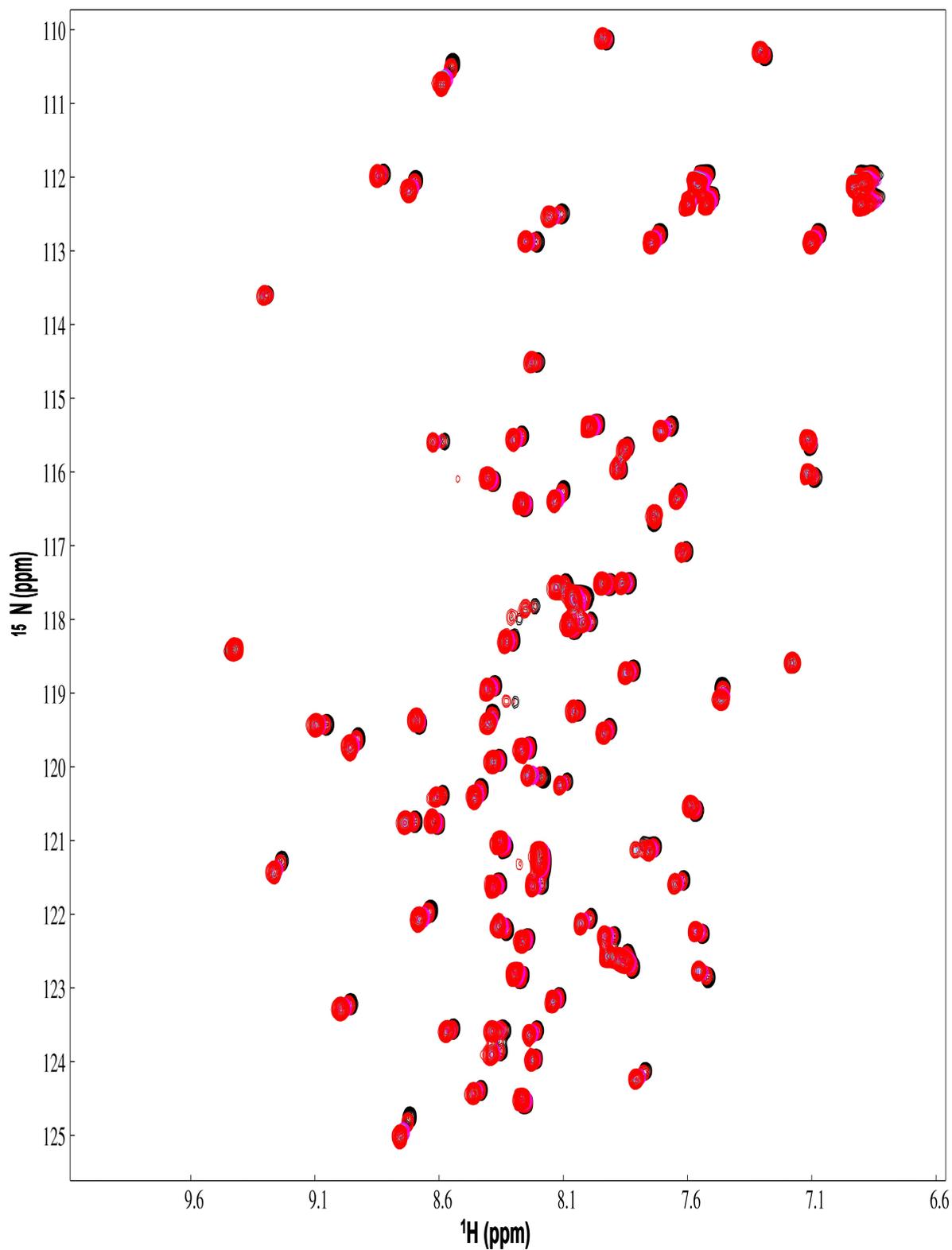


Figure 3.22 Superimposed $^1\text{H},^{15}\text{N}$ -HSQC spectra of free NusA ar1 in the absence (black) and presence (red) of Nun CTD (45-112). During the titration the black colored peaks did not shift indicating that there might be no interaction with Nun CTD.

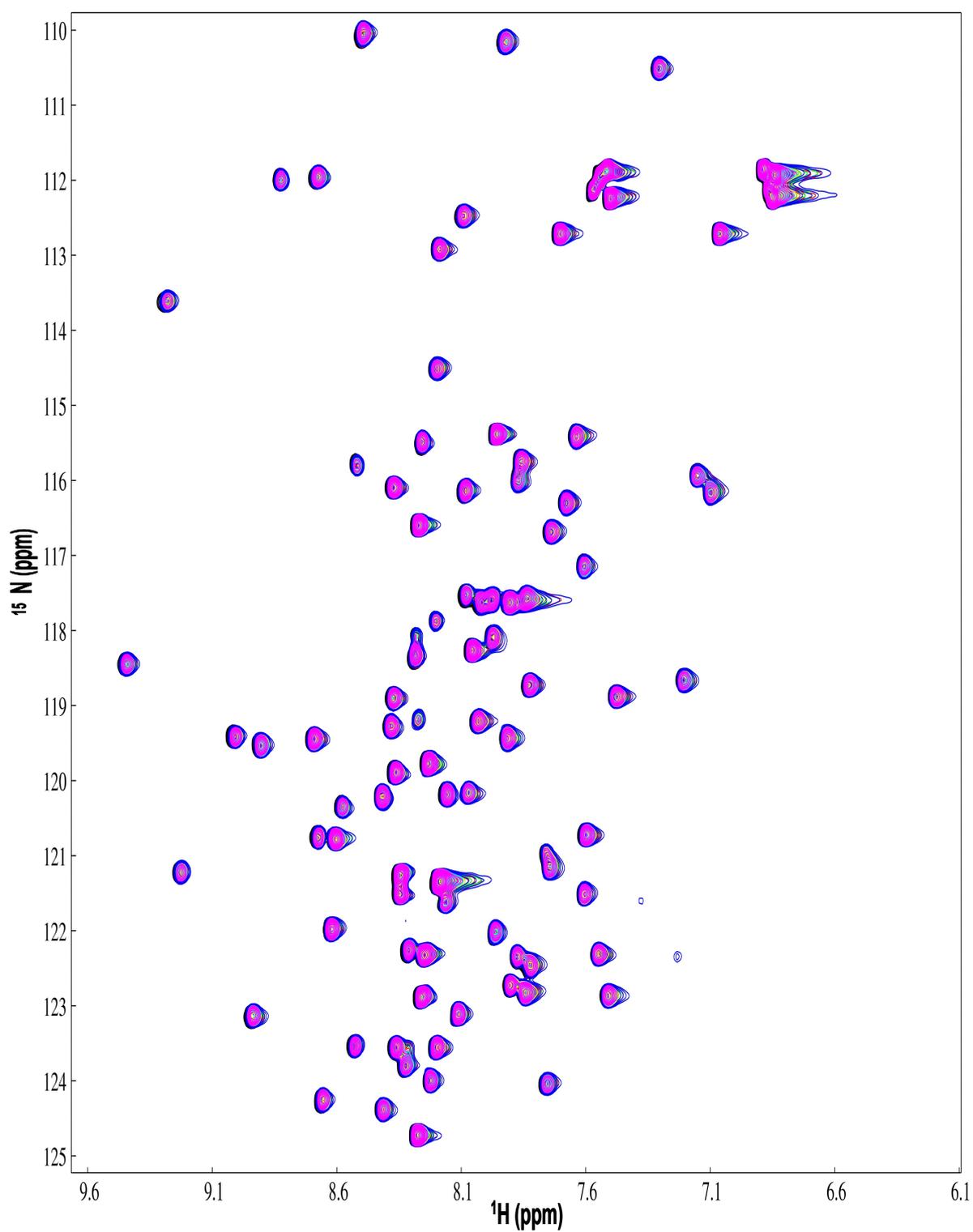


Figure 3.23 An overlay of the $^1\text{H},^{15}\text{N}$ -HSQC spectra of free NusA ar1 (black) upon titrating with Nun (1-112) (magenta). No major changes in the chemical shift were observed for any residues.

3.5.2 Interaction of Nun with NusG

NusG was first identified in the phage lambda N-mediated transcription antitermination reaction. NusG, along with other *E. coli* Nus factors is required for optimal lambda N activity in an *in vitro* transcription system.

Burova et al., 1999 had demonstrated that the *Escherichia coli* NusG gene product is required for transcription termination by phage HK022 Nun protein at the λ *nutR* and λ *nutL* site. It has been proposed that NusG stimulates termination by serving as a bridge between RNA polymerase and Rho, thus helping to recruit Rho into the termination complex [Li et al., 1993].

The requirement of NusG has been shown by depleting NusG from cells and measuring the activity of a λ pR-*nutR*-galK operon fusion [Sullivan and Gottesman., 1992]. NusG depletion prevented Nun-mediated termination but did not inhibit N antitermination. The efficiency of Nun-mediated transcription arrest *in vitro* is strongly enhanced by NusG in combination with the three other Nus factors [Hung and Gottesman., 1995].

Based on these observations, interaction studies in between Nun and NusG was performed. ^{15}N labeled NusG was used for the titration study. The titration of NusG with the Nun (1-112) was carried out to monitor the changes in the $^1\text{H},^{15}\text{N}$ -HSQC spectrum during each step of the titration. The titrations were carried out in a buffer containing 10 mM KPO_4 , 50 mM NaCl at pH 6.4. The titration experiment was also performed in the absence of salt to observe, whether is there any change in the interaction behavior due to the presence of salt or not. Titration was carried out up to a molar ratio of 1:4.

An overlay of the titration spectra is shown in the Fig 3.24. No changes in the $^1\text{H},^{15}\text{N}$ -HSQC chemical shifts or the appearance of the resonances for NusG were observed in all the stages of titration. This indicates, that no binding interaction between NusG and Nun exists. Therefore, the proposed role of NusG in Nun mediated termination might be different.

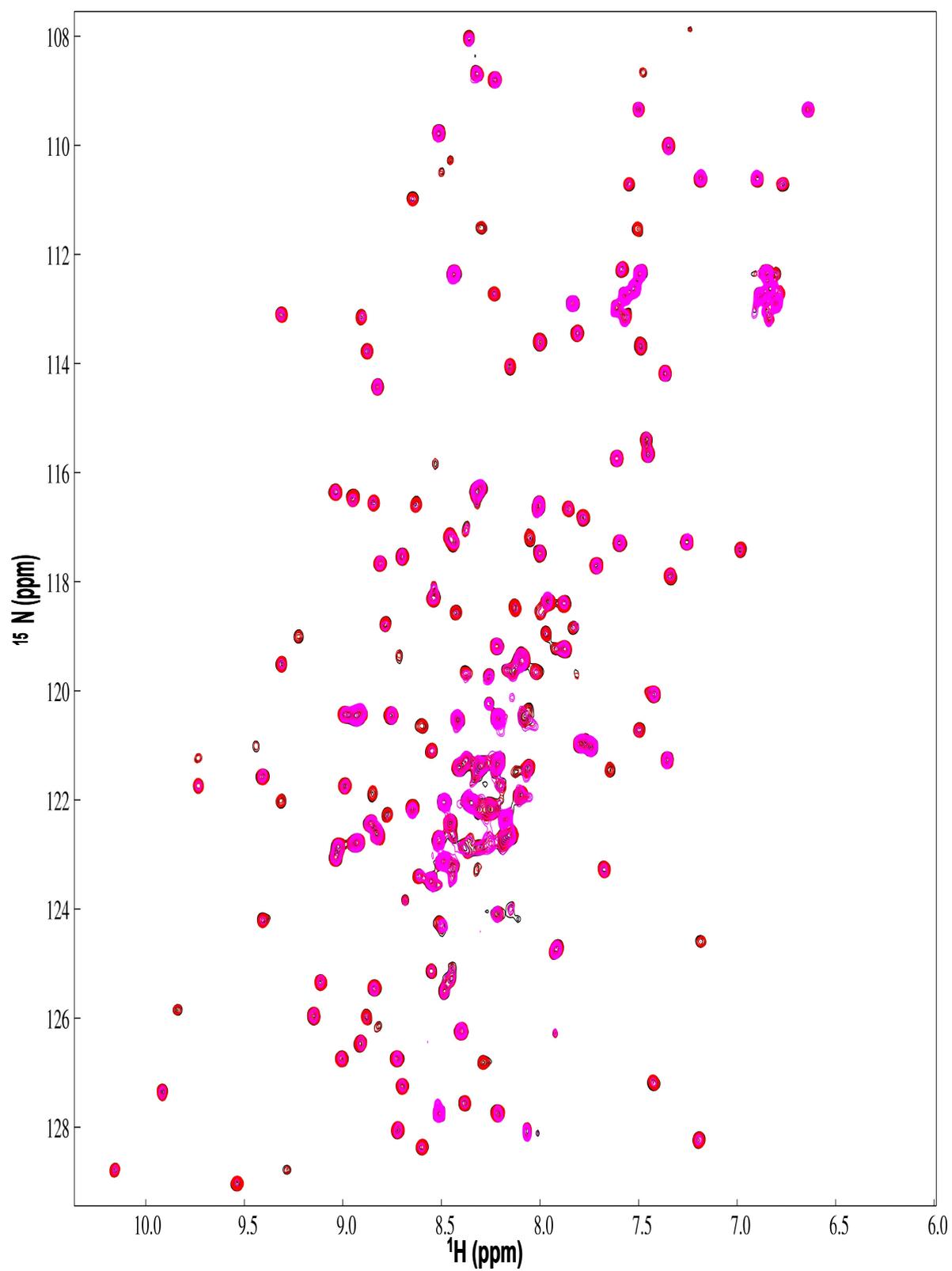


Figure 3.24 An overlay of the $^1\text{H},^{15}\text{N}$ -HSQC spectra of ^{15}N labeled NusG (black) upon titrating with Nun (1-112) (magenta).

3.5.3 Interaction of Nun with NusB

It has been reported by Washburn et al., 2006, that the Nun termination in a strain deleted for *nusB* was reduced in its activity by two fold *in vivo*. Nun termination at *nutR* is ablated by mutations in the host *nusB* genes. Till today, the exact role of NusB in Nun termination is not known. So far, it is also not known, is there any direct interaction between NusB and Nun protein. To observe the same, a series of $^1\text{H},^{15}\text{N}$ -HSQC spectra of NusB was recorded by adding gradually an increasing molar ratio of Nun (1-112). An overlay of the $^1\text{H},^{15}\text{N}$ -HSQC spectra is shown in the Fig 3.25. However, titration of NusB upon adding Nun, does not cause any detectable changes, suggesting that there might be no direct interaction.

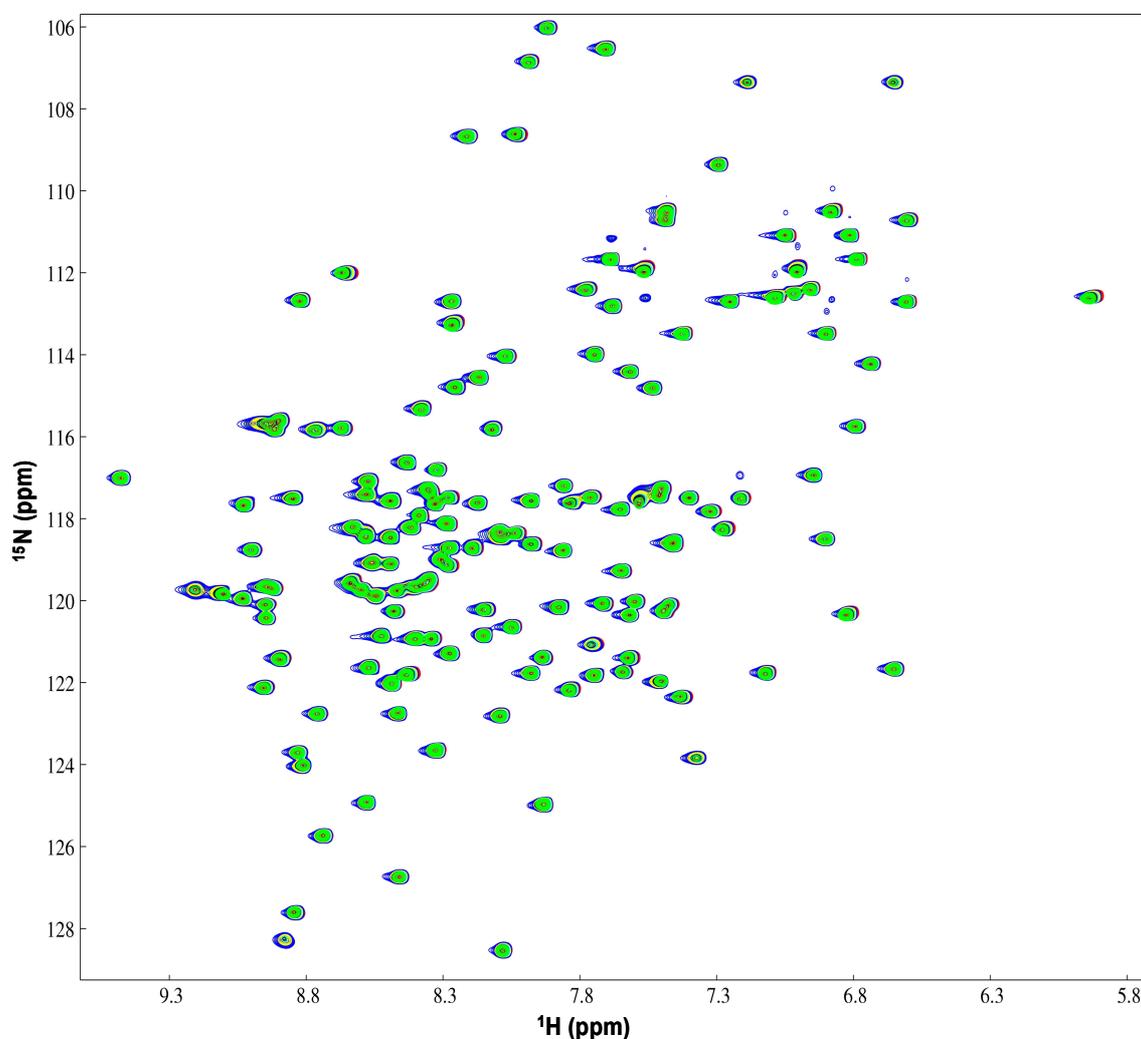


Figure 3.25 An overlay of the $^1\text{H},^{15}\text{N}$ -HSQC spectra of ^{15}N labeled NusB (black) upon titrating with Nun (1-112) (green).

3.6 Backbone resonance assignment of SKK domain of NusA

There are various evidences showing the interaction of NusA directly with the *nut* site RNA and mutations in the RNA-binding domains of NusA (S1+KH1+KH2) has an impact in the mechanism of antitermination. In the present study, we characterized the binding interaction of RNA-binding domains (SKK domain) of NusA with that of *nut* site RNA. In order to investigate the interaction between SKK domain and *nut* site RNA, and to map the binding surface, backbone resonance assignment of SKK domain is essential to proceed.

Because of the large size of the protein (24.4 kDa), chemical shift degeneracy, overlapping signals, line broadening and poor spectral sensitivity poses a challenging problem (2.14.4.2). To overcome this problem, TROSY in combination with deuterium labeling (2.10.1) has been used. The gain in spectral resolution and sensitivity is readily apparent from comparison with the corresponding conventional experiment (Fig 3.17 in section 3.2.4)

Relatively good degree of dispersion in both the ^{15}N and ^1H dimensions indicated a well-folded protein. ^1H , ^{15}N , and ^{13}C assignments have been made using TROSY based triple resonance NMR experiments on ^2H , ^{13}C , ^{15}N -labeled SKK domain (section 2.10).

The triple resonance experiments were recorded as shown in Tables 2.3 and 2.4 (section 2.14.4.3) to correlate the resonances of the peptide backbone [H^{N} (i), N^{H} (i), C^{α} (i), C^{α} (i-1), C^{β} (i), C^{β} (i-1), CO (i), and CO (i-1)]. ^1H , ^{15}N -TROSY experiment contains the ^{15}N - and H^{N} -resonances, therefore allowing the use of this pair of spins as reference and starting point for further assignment of other resonances.

The HNCA experiment, for example, correlates the H^{N} and ^{15}N chemical shift of residue (i) with the C^{α} shifts of residue (i) (via $^1\text{J}_{\text{NC}\alpha} \gg 7\text{-}12\text{ Hz}$) and residue (i-1) (via $^2\text{J}_{\text{NC}\alpha} < 8\text{ Hz}$), thereby providing sequential connectivity information. Hence, a HNCA experiment gives both the inter- and intra-residue shifts for the matching atom i, and thus will show two cross peaks for each amide group.

In order to unambiguously identify whether the peak originates from the inter- or intra-residue spin, complementary experiment is needed which generates signals from only one of the two matching atoms. For ex., the HN(CO)CA experiment complements the HNCA experiment by providing only inter-residue α -carbon shifts.

Other triple-resonance spectra would be very similar in appearance, with the significant difference being the carbon frequency axis. In the HNCO and HN(CA)CO experiments carbonyl shifts would be observed while in the HNCACB and HN(CO)CACB experiments the β -carbon shifts would be observed.

To achieve the sequential assignment of SKK domain, the recommended set of six experiments which follows are HNCO, HNCA, HN(CO)CA, HN(CA)CO, HNCACB and HN(CO)CACB. C^α and C^β are of prime importance, because their chemical shifts show a large spectral dispersion ($C^\alpha \gg 25$ ppm; $C^\beta \gg 60$ ppm), and these shifts are characteristic for the identification of the amino acids.

The most important experiment for the assignment of the backbone resonances is the HNCACB. This experiment yields the C^β shifts [in position (i) and (i-1)] in addition to those coming from the HNCA. The C^α and C^β correlations have opposite signs and can thus be distinguished. The resonances in the (i)-position can be discriminated from those in the (i-1)-position on their different intensity as explained for the HNCA experiment.

The “domino pattern” obtained for SKK domain from amino acid S295 to D299 during the sequential assignments with the triple resonance spectra is shown in Fig. 3.26. It shows the superposition of HN(CO)CACB and HNCACB spectrum of the corresponding amino acid. Pairs of consecutive residues can thus be identified using these two experiments, making it possible to “walk along” the protein backbone. The HN(CO)CACB in addition to the HNCACB makes it possible to distinguish between peaks belonging to the “in-residue” and to the preceding residue respectively.

Two dimensional “strip plots” are generated from three dimensional spectra by extracting tubes centered on each resonance in the ^{15}N -TROSY spectrum and extending across the full ^{13}C spectral width. The tubes are reduced to two dimensional strips by taking cross sections through the tubes, either in the ^1H or the ^{15}N dimensions.

The C^α and C^β peaks in the HNCACB spectrum from the double strip plot can be distinguished easily by the fact that they have opposite signs: the C^α peaks are shown in black and C^β peaks in red.

Extraction of ^{13}C chemical shifts

Carbonyl carbon “CO” chemical shifts were extracted from HNCO spectrum of the SKK domain. The HNCO experiment is the most sensitive heteronuclear experiment and correlates the amide ^1H and the ^{15}N chemical shifts of residue i with the carboxyl carbon chemical shift of residue $i-1$, by using the large one-bond $J_{\text{N-CO}}$ coupling. Since, every amide group is covalently bonded to a single “CO”, only a single cross peak per residue is observed in the HNCO spectrum. The HN(CA)CO can in principle give both inter- and intra-residue carbonyl shifts. However, the inter-residue peak is generally of low intensity and often not observable. A strip plot showing the superpositions of a HNCO spectrum (left) and a HN(CA)CO spectrum (right) is shown in Fig 3.27. The picture contains strips from the two spectra which correspond from the amino acid S295 to D299 of SKK domain. Several of these strips are placed in a row to show the sequential connectivities from each amino acid to the preceding one.

Assignment strategy

In order to link the sequentially assigned strips to stretches of the amino acid sequence, C_α and C_β shifts were compared with the shifts expected for particular amino acids. [Cavanagh et al., 1996]. Using these chemical shifts as reference, alanine residues were identified because of their high C_β shifts, while threonine and serine residues were identified due to low chemical shifts for both C_α and C_β . Glycine residues lack C_β 's and therefore has only one peak in the HN(CO)CACB, corresponding to the C_α making them very easily identifiable.

This information, taken together with the known amino acid sequence of the SKK domain, allowed the backbone sequential assignment to be nearly complete. Assignment of the ^1H - ^{15}N TROSY spectrum of SKK domain is shown in Fig. 3.28. Few residues were unassigned owing to strong overlap or missing resonances.

Leaving the tag part (7) and the proline (9) residues, 166 of the 206 residues (80.5 %) of the SKK domain could be assigned using triple resonance experiments. Among the 40 unassigned residues, no resonance signals are seen for 15 residues and for the rest of 25 residues it was not possible to get the sequential connectivities (section 4.2). The results of the backbone assignments are listed in Appendix 9.10.

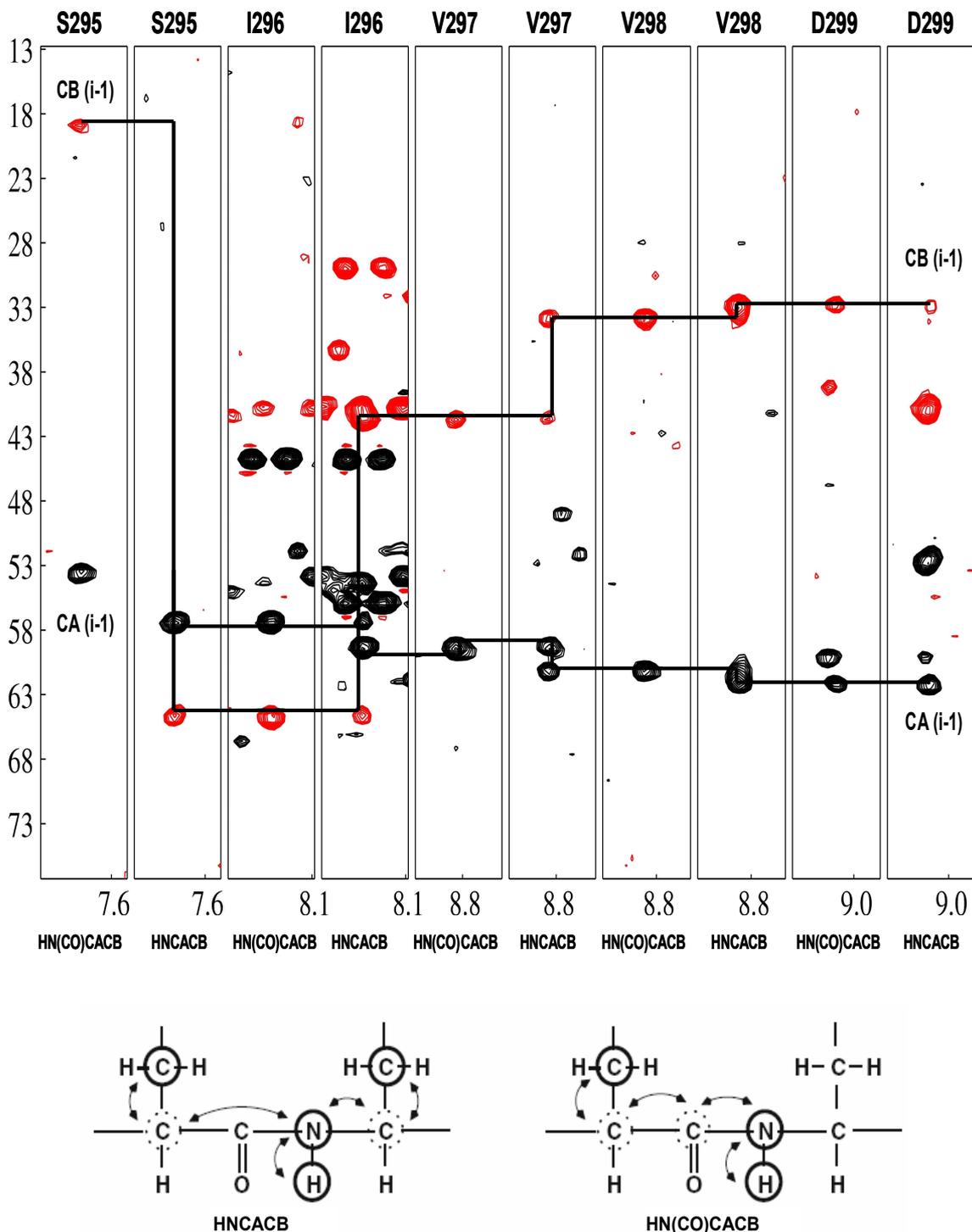


Figure 3.26 Strips of HN(CO)CACB and HNCACB spectra of ^2H , ^{13}C , ^{15}N uniformly labeled SKK domain. Strips from two spectra are shown, corresponding to a single amino acid. Several of these strips are placed in a row to show the sequential connectivities from each amino acid to the preceding one. The coherence transfer in both of these experiments for a pair of consecutive residues are shown below. The arrows indicate the magnetization transfer pathway.

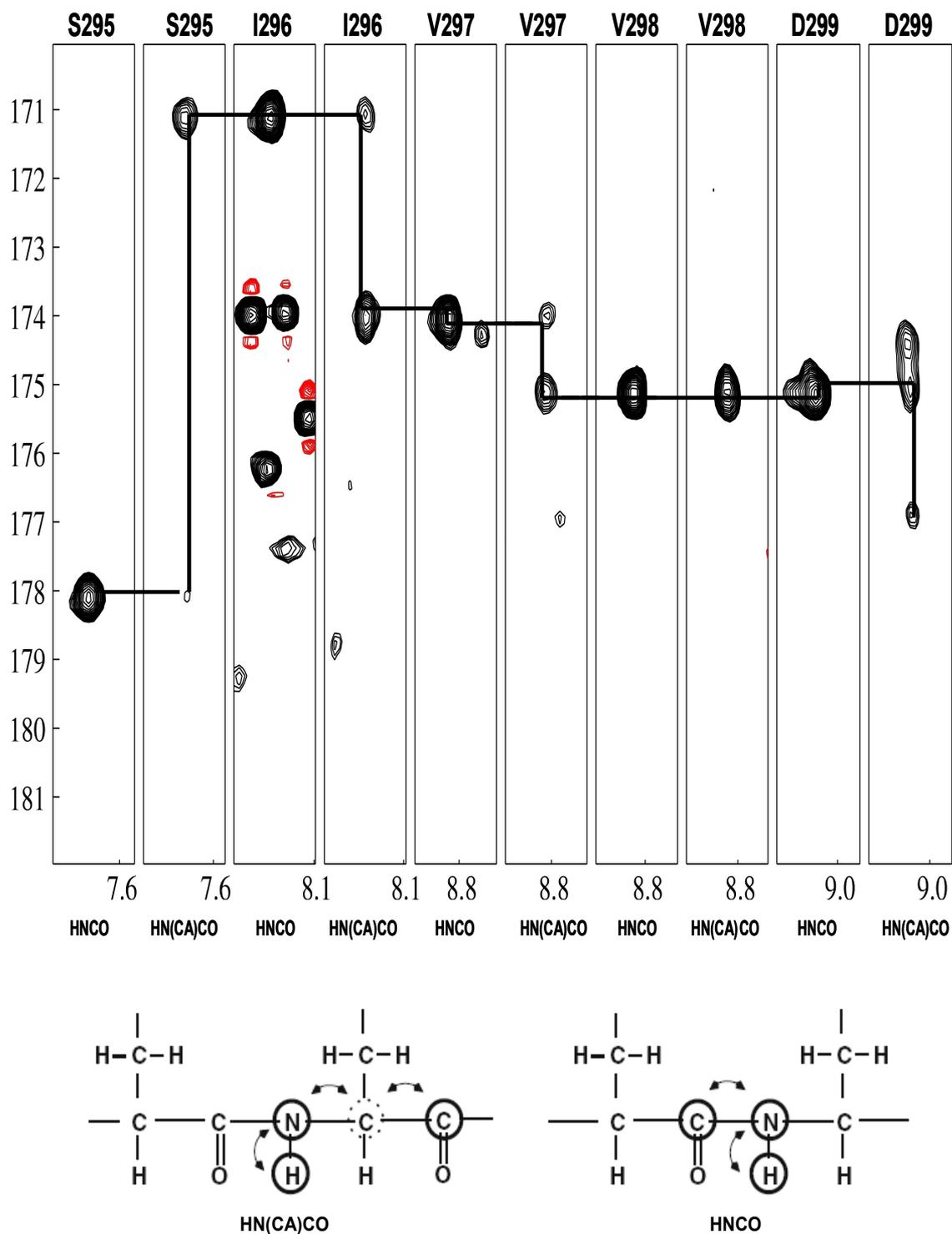


Figure 3.27 Strips showing backbone sequential connectivities of residues 295 to 299 of the SKK domain. The strips are taken from HNCO and HN(CA)CO spectra of ^2H , ^{13}C , ^{15}N uniformly labeled SKK domain and each strips from two spectra corresponds to a single amino acid. Several of these strips are placed in a row to show the sequential connectivities. The flow of magnetization is indicated by arrows.

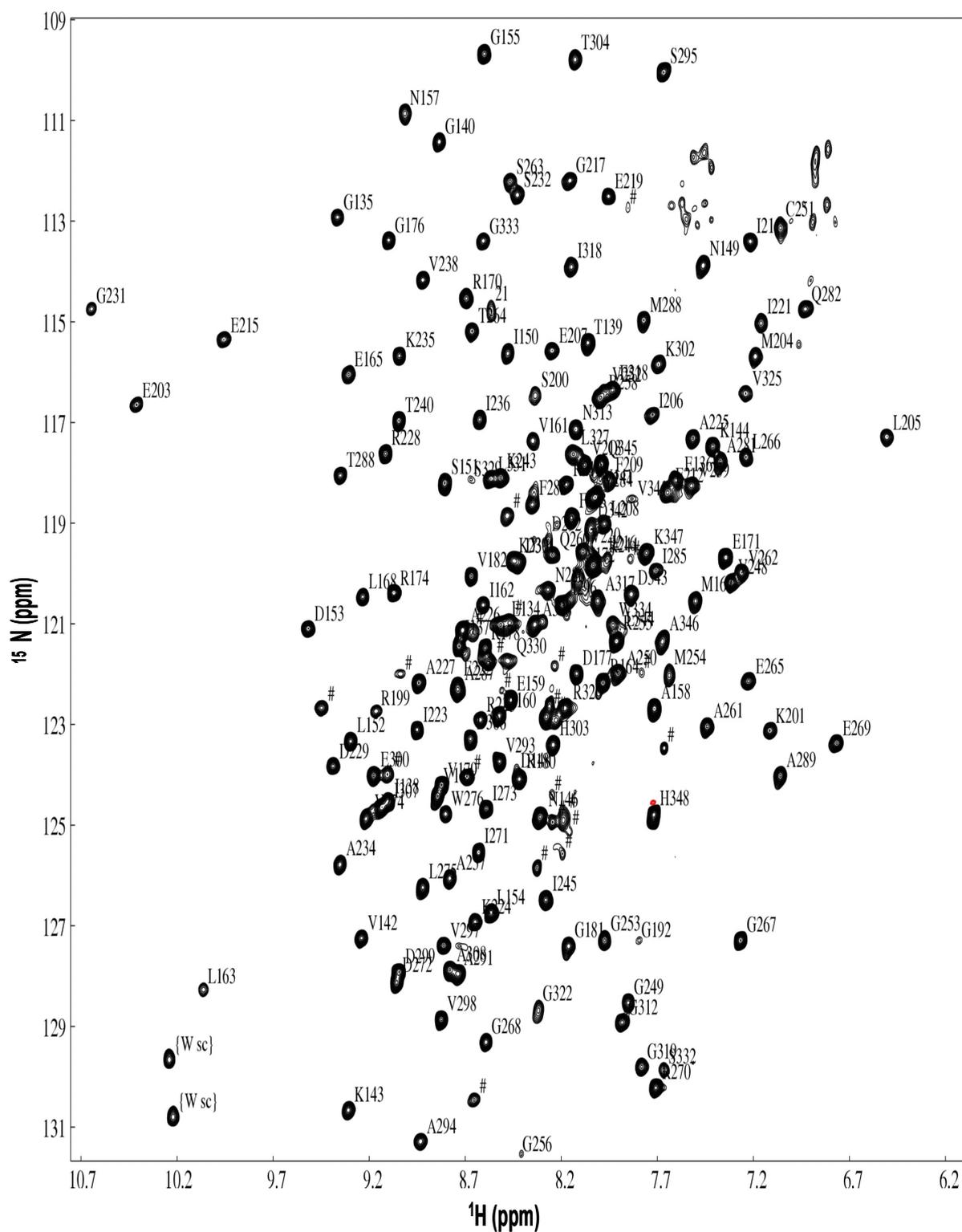


Figure 3.28 $^1\text{H},^{15}\text{N}$ -TROSY spectrum of $^2\text{H},^{13}\text{C},^{15}\text{N}$ uniformly labeled SKK domain (400 μM ; 800 MHz; 298K). Assigned backbone resonances are labeled with amino acid type in one letter code and residue number. The tryptophan NH are marked by “W sc”. The unassigned resonances are indicated by hash (#).

3.7 Interaction studies of NusA RNA binding domains (SKK)

3.7.1 Binding of SKK with *nut* RNA

As described earlier, the N protein, with a group of *E. coli* encoded proteins that, in addition to NusA, include NusB, NusG, and NusE, act at *nut* RNA sites, to modify RNA polymerase to a termination-resistant form. The *nut* RNA sequences are components of transcripts initiating at the early λ promoters P_R and P_L . The respective *nut* site in the P_R and P_L operons are *nutL* and *nutR* which lies upstream of the first terminator (section 1.4).

Sequence analyses identified three regions in NusA having homologies with sequences associated with RNA binding; one S1 and two KH domains, in the central portion of *ecoNusA*. The crystal structure of *Thermatoga maritima* NusA [Worbs et al., 2001] suggests that NusA creates an extended, mosaic RNA interaction surface by domain arraying.

Consistently, all portions of the molecule have been implicated by mutational analyses in RNA binding: the R199A mutation in the interface of S1 and KH1, and point mutation in the GXXG motifs of both KH elements, all impair binding to *nut* site RNA.

It has been shown by anisotropic fluorescence titrations [Prasch et al., 2008; in revision] that titration of SKK domain with λ *nutR* showed only weak protein-RNA interactions compared with λ *nutL* which showed slightly higher affinity towards SKK domain.

Based on these reports, the interaction of RNA binding domains of NusA with λ *nutL* RNA was investigated by NMR titration studies. To identify the interaction between SKK and λ -*nutL*, ^2H , ^{15}N labeled SKK protein was titrated with an increasing molar ratio of unlabeled λ -*nutL* RNA. The chemical shift changes have been monitored in the ^1H , ^{15}N -TROSY spectra during each step of the titration upon gradual addition of the λ *nutL* RNA. Table 3.1 shows the details of the titration experiments.

Table 3.1 Titration table for studying the interaction of SKK with λ *nutL* RNA.

Experiment. No	SKK domain (mM)	λ <i>nutL</i> RNA (mM)	Total volume (μ L)	Ratio
1	0.1000	0.000	550	1:0
2	0.0965	0.025	570	1:0.25
3	0.0932	0.050	590	1:0.50
4	0.0902	0.075	610	1:0.75
5	0.0873	0.100	630	1:1.0
6	0.0821	0.150	670	1:1.5
7	0.0775	0.200	710	1:2.0
8	0.0733	0.250	750	1:2.5
9	0.0696	0.300	790	1:3.0

An overlay of the ^1H , ^{15}N -TROSY spectra at each NMR titration step is shown in Fig 3.29. For SKK- λ *nutL* RNA complex, many peaks had significant $\Delta\delta$ (section 2.14.6) as compared to the free SKK protein, indicating a binding interface.

Few resonances moved in a continuous fashion (ex., T240, K235, V238, T264, R270, D272, S295, V297 etc.) falling in the fast exchange regime on NMR time scale. Thus, the resonances of the nuclei affected by RNA binding gradually shift their position from the resonance of the free state towards the resonance of the bound state. Apart from the differences in chemical shifts of certain residues, it was also noted that some of the resonances disappear completely (For ex., V179, G192, A234, I236, C251, G267, I271, I273, M288, I318 etc.). Upon immediate titration of λ *nutL* RNA (0.25 molar equivalence) disappearance in the resonances of these residues were observed. Absence of signals in the ^1H , ^{15}N -TROSY spectra for some of the residues, could be probably due to exchange processes on the intermediate time scale, suggesting that these residues are involved in binding. This is characteristic for affinities in the low-micromolar range.

It is important to recognize that “NMR timescale” is a relative one. For a given equilibrium, different resonances will show slow, intermediate or fast exchange behavior, depending on how much their chemical shifts differ between the free and bound states. The Fig 3.30 shows a expanded region of the overlaid ^1H , ^{15}N -TROSY spectra from SKK- λ *nutL* RNA titration.

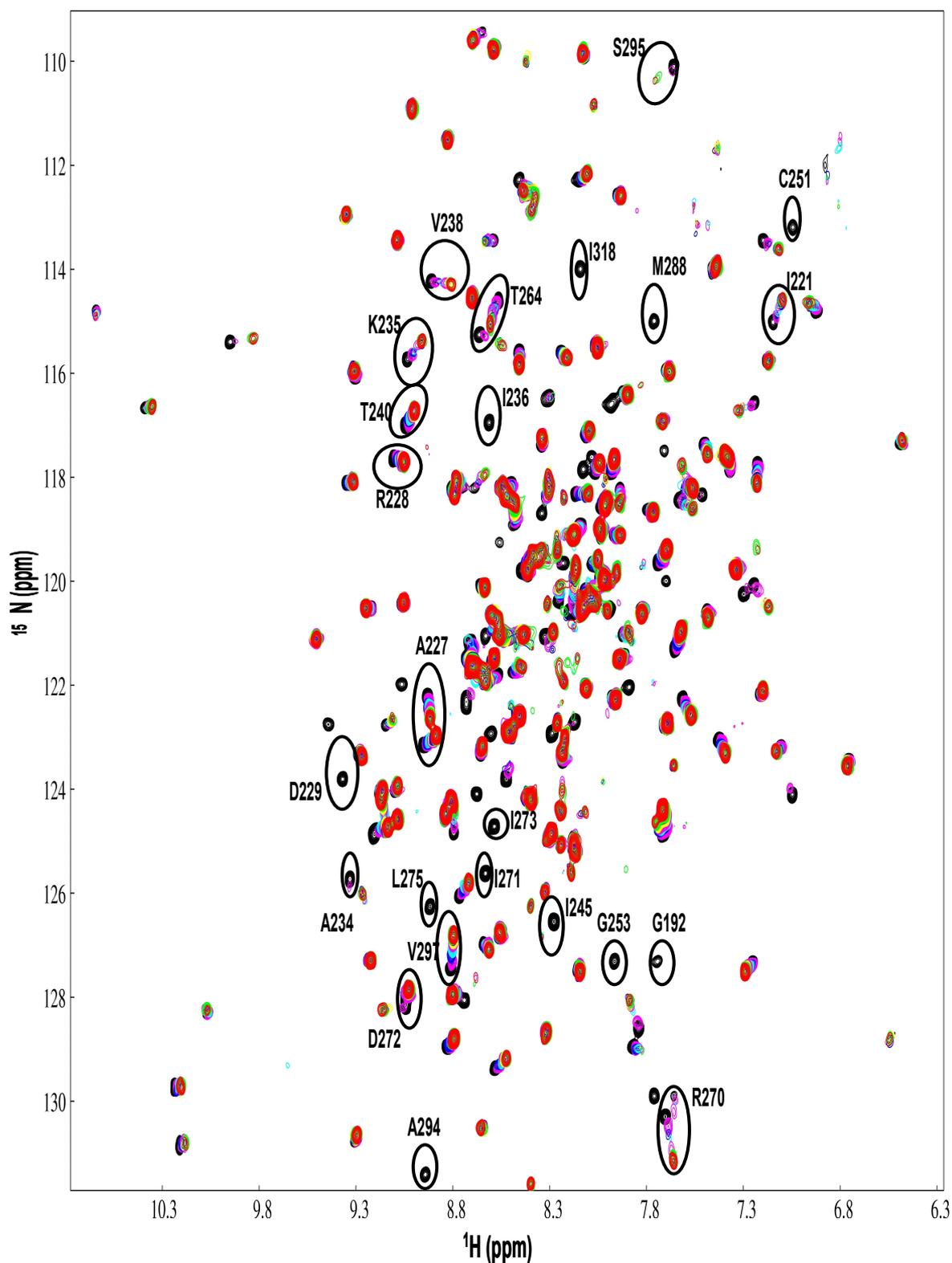


Figure 3.29 Titration of SKK domain with λ *nutL* RNA. Overlay of ^1H , ^{15}N -TROSY spectra recorded during the titration with different RNA/protein ratios. Key; black 0.0, magenta 0.25, blue 0.50, cyan 0.75, magenta 1.0, yellow 1.5, blue 2.0, green 2.5, and red 3.0. Resonance signals involved in binding are annotated.

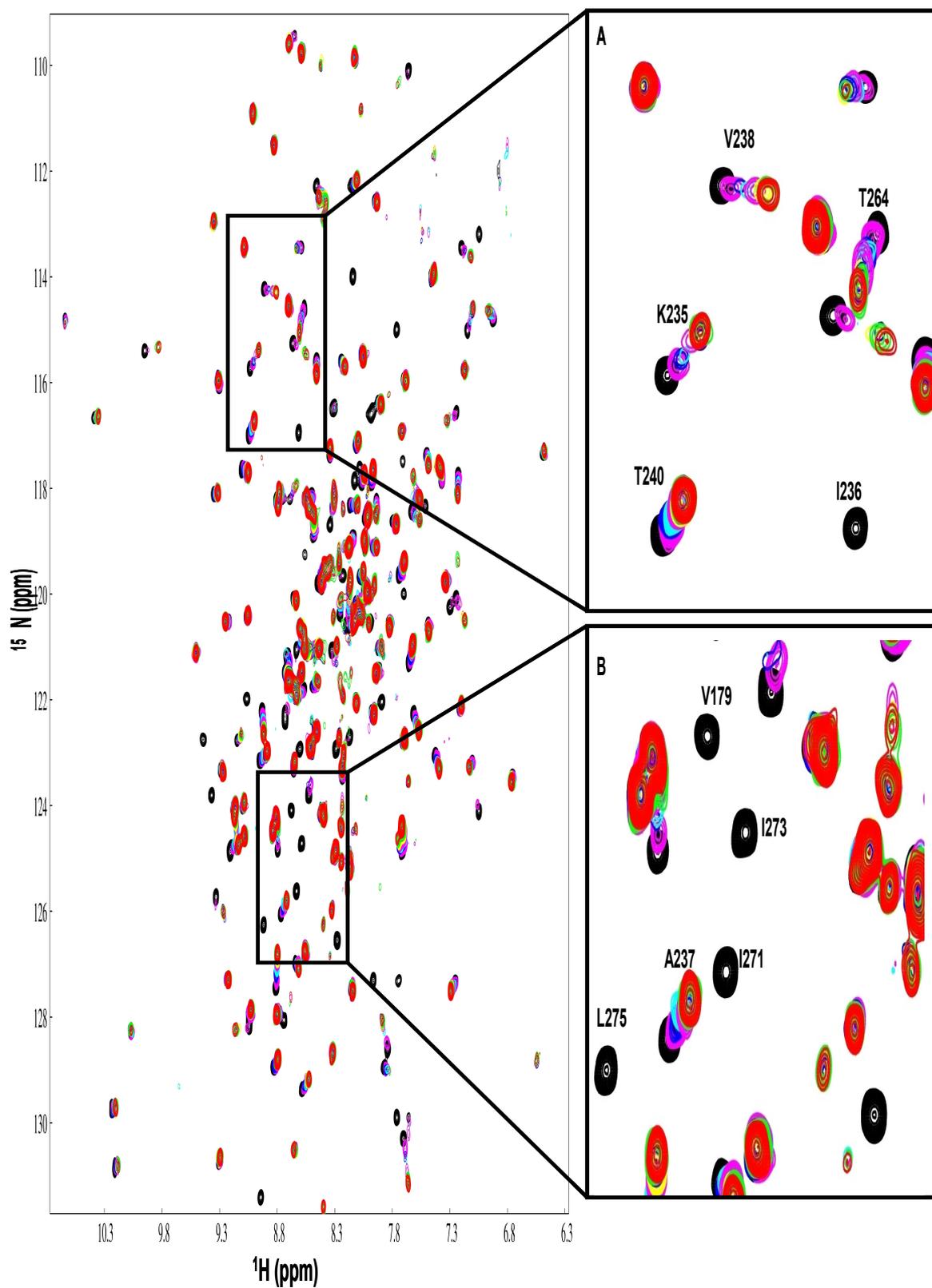


Figure 3.30 ^1H , ^{15}N -TROSY spectra of SKK upon titrating with λ *nutL* RNA. (A/B)-Residues showing different chemical environment due to λ *nutL* RNA binding

3.7.2 Normalized chemical shift changes

Perturbations of ^{15}N and $^1\text{H}^{\text{N}}$ chemical shifts of a protein upon complexation with a ligand are a qualitative tool for mapping of residues involved in binding sites and/or identifying conformational rearrangements (section 2.14.6). Hence, chemical shift mapping was used in order to identify the putative sites of interaction on SKK domain, by detecting the chemical shift perturbation in the $^1\text{H},^{15}\text{N}$ -TROSY of SKK upon titrating with λ *nutL* RNA.

In Fig 3.31 (A), changes in chemical shifts are displayed through the use of the normalized weighted chemical shift average between the free SKK and its complex with λ *nutL* RNA (equation 2.2). Normalized chemical shift changes larger than 0.04 ppm are considered to be significant [Hajduk et al., 1997] as indicated by dashed line in Fig 3.31 (A).

Nearly all $^1\text{H}^{\text{N}}$ and ^{15}N resonances in the KH domains are affected upon binding to λ *nutL* RNA, whereas, the more prominent changes occurring in the region R210-W276 which encodes for KH1 domain. Previously, it has been implicated that the KH domains are the important contributors to NusA binding to RNA [Worbs et al., 2001]. From our experimental results, it could be deduced that the KH1 domain plays a significant role in *nut* RNA binding.

The observed chemical shift changes were mapped onto the surface of the crystal structure of *Thermotoga maritima* NusA (S1+KH1+KH2) to get further insight into the binding interface. The Figure 3.31 (B) shows the surface representation of SKK domain highlighting the binding interface and the residues whose resonances are affected upon binding to λ *nutL* RNA.

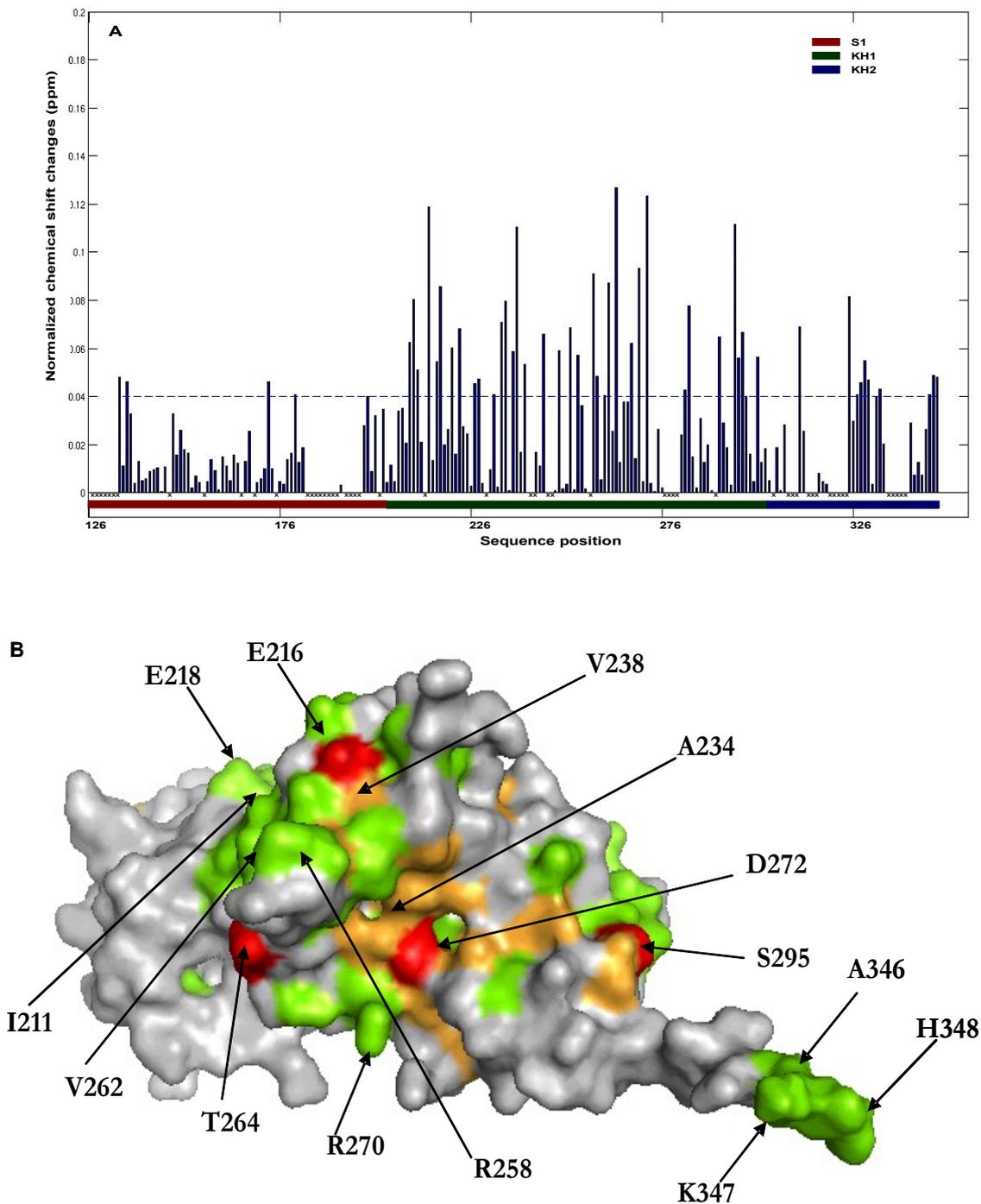


Figure 3.31 (A) Chemical shift changes of SKK upon binding to λ *nutL* RNA as a function of primary sequence; Dotted line represents the significance level of 0.04 ppm; X = residues not assigned. The colored bars represent the three RNA binding domains. (B) Surface representation of SKK highlighting the binding interface. Few of the residues with resonances showing significant chemical shift changes ($0.04 < \Delta\delta \leq 0.1$) are shown in green, those with $\Delta\delta > 0.1$ are in red. Disappearing resonances are presented in orange.

3.7.3 Dissociation constant for SKK- λ *nutL* complex

The dissociation constant K_D is determined from the changes in chemical shifts of $^2\text{H},^{15}\text{N}$ -labeled SKK in a $^1\text{H},^{15}\text{N}$ -TROSY after gradual addition of λ *nutL* RNA (section 2.14.7). Signals showing a behaviour in the limit of fast exchange on the NMR time scale were fitted to equation (2.4) for a two state model. K235, V297 and R270 of SKK domain which show fast exchange regime during the titration was chosen to determine the K_D value (Fig 3.32). It has been observed by fluorescence titration that λ *nutL* binds to SKK with an affinity of 71 μM [Prasch et al., 2008, in revision]. The difference between the K_D values obtained from NMR and fluorescence experiment could be supported by the fact that an accurate measurement of K_D requires the use of protein concentrations $\leq K_D$ and the low sensitivity of NMR in terms of concentrations required obviously sets a practical limit to the range of dissociation constants which can be measured. For the analysis of K_D , ligand concentrations greater than the protein concentration are used, which gives an error up to a factor of ten [Feeney et al., 1979].

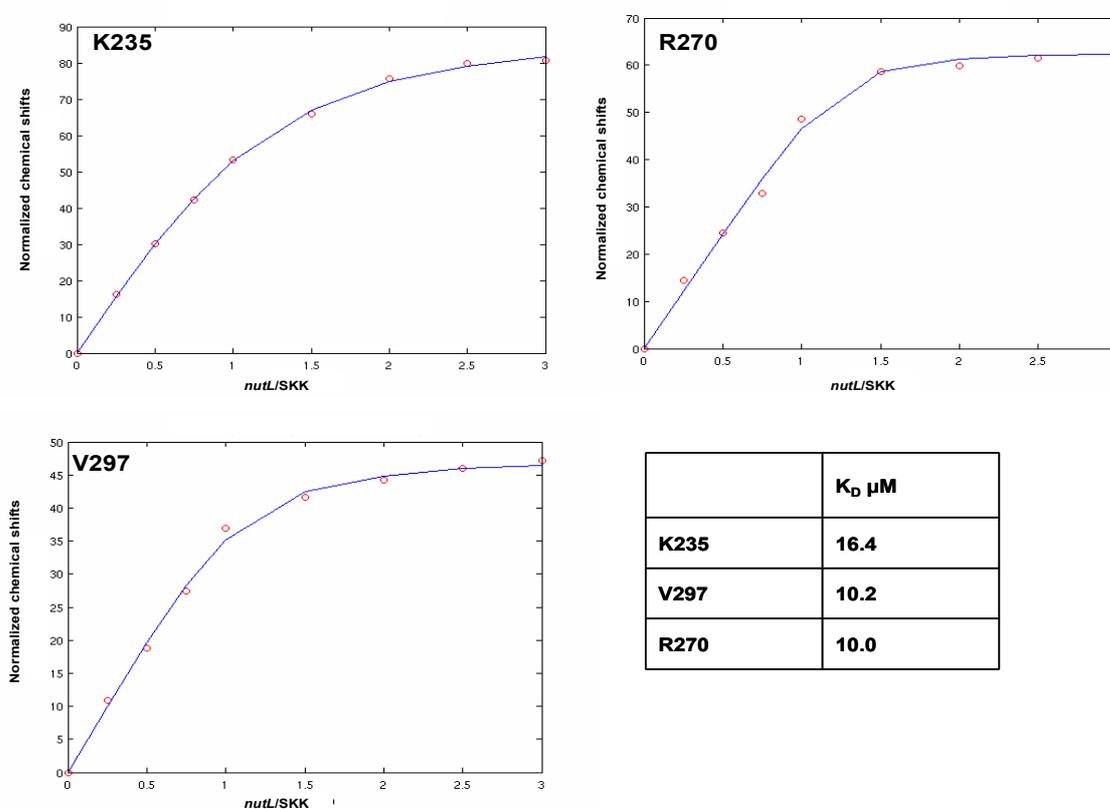


Figure 3.32 Dissociation constants for SKK- λ *nutL* complex. Fitting the curves yielded the calculated K_D values shown in the inset.

3.7.4 Interaction with λ *nutL* spacer

By anisotropic fluorescence titration experiments, it has been validated that *nutL* spacer is important for specific binding to SKK domain [Prasch et al., 2008, in revision]. The seven nucleotide long *nutL* spacer region separates *BoxA* and *BoxB*. Based on the reported results, the interactions of the SKK domain with the *nutL* spacer have been analysed by NMR. The sample conditions and the titration steps were maintained the same as for the SKK+ λ *nutL* RNA titration. ^2H , ^{15}N labeled SKK protein was titrated with *nutL* spacer with gradual increase in the concentration of *nutL* spacer. An overlay of the ^1H , ^{15}N -TROSY spectra of SKK domain upon interacting with *nutL* spacer is shown in Fig 3.33. Most of the resonances which disappeared (intermediate exchange) during the titration of SKK with λ *nutL* (3.7.1) were affected in the same manner as compared to the titration of *nutL* spacer alone. Few resonances which showed significant chemical shift changes during SKK+ λ *nutL* RNA titration were not seen with *nutL* spacer alone. This probably gives us a clue that *nutL* spacer plays an important role in binding to SKK, while some more nucleotides from *BoxA* and *BoxB* are also involved in specific binding to SKK.

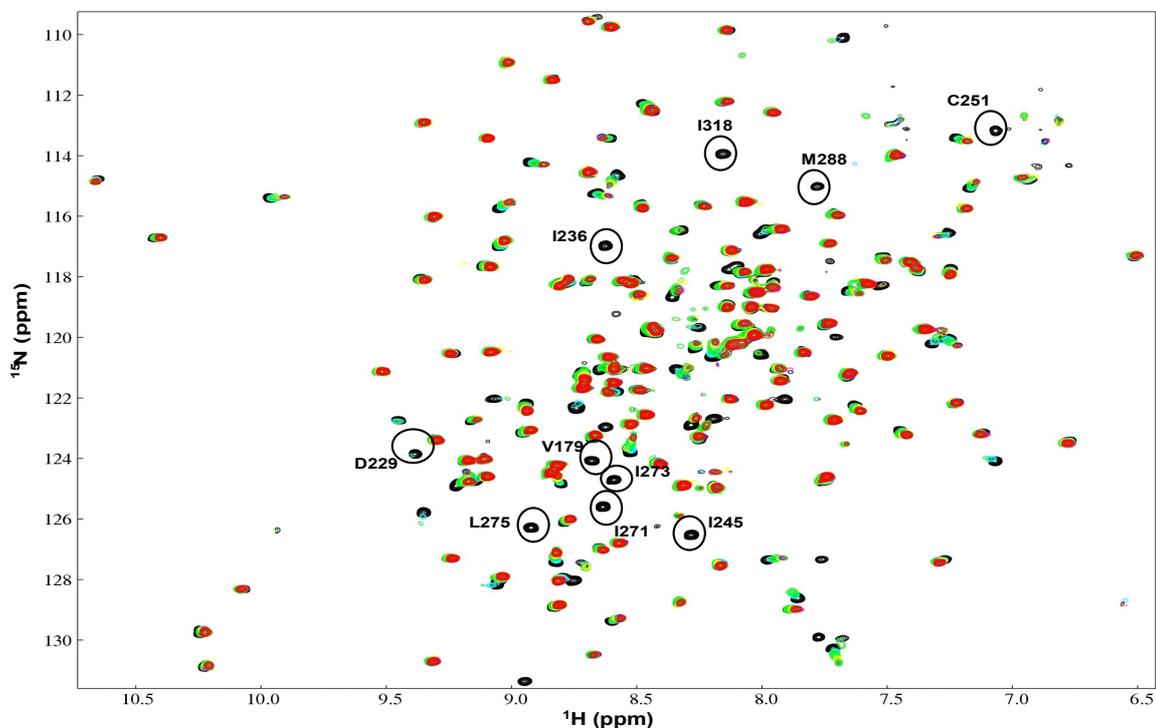


Figure 3.33 Titration of SKK domain with *nutL* Spacer. Overlay of ^1H , ^{15}N -TROSY spectra recorded during the titration with different RNA/protein ratios. Key; black 0.0, cyan 0.25, green 0.50, yellow 0.75, blue 1.0, magenta 1.5, and red 2.0. Resonances in the intermediate exchange are indicated.

3.8 Autoinhibition effect of NusA ar2

3.8.1 Interaction of NusA ar2 with SKK domain

Autoinhibition by NusA ar2 is explained in detail in section 1.6. As described earlier, none of the NusA fragments except NusA (1-416) can bind to *nut* site RNA directly in the absence of N. Recently, it was shown that the 79 carboxy-terminal amino acids of NusA hinders RNA interaction in isolation but is sequestered in transcription complexes by the C-terminal domain (CTD) of RNA polymerase subunit α (α -CTD) [Mah et al., 1999; Mah et al., 2000]. The C-terminal NusA repeats constitute a versatile protein–protein interaction region because they can form a complex with the α -CTD, λ N, and presumably, the remainder of NusA during inhibition of RNA binding.

Based on these results of autoinhibition by NusA ar2, we have performed a titration between NusA ar2 and SKK domain by NMR to characterize the interaction between these two proteins. A series of $^1\text{H}, ^{15}\text{N}$ -TROSY of $^2\text{H}, ^{15}\text{N}$ labeled SKK domain was recorded upon titrating with unlabeled NusA ar2 (Table 3.2).

Table 3.2 An overview of the titration steps for SKK with NusA ar2 titration.

Experiment. No	SKK domain (mM)	NusA ar2 (mM)	Total volume (μL)	Ratio
1	0.2300	0.000	550	1:0
2	0.2219	0.057	570	1:0.25
3	0.2144	0.115	590	1:0.50
4	0.2074	0.172	610	1:0.75
5	0.2008	0.230	630	1:1.0
6	0.1888	0.345	670	1:1.5
7	0.1782	0.460	710	1:2.0
8	0.1686	0.575	750	1:2.5
9	0.1601	0.690	790	1:3.0
10	0.1524	0.805	830	1:3.5

An overlay of the $^1\text{H},^{15}\text{N}$ -TROSY spectra at each NMR titration is shown in Fig 3.34. Addition of NusA ar2 to SKK domain showed significant chemical shift changes in the $^1\text{H},^{15}\text{N}$ -TROSY spectrum. The residues which were located in the KH1 domain (A234, V252, G249, Q260, A261, T264, S263, R270 and I271) showed remarkable chemical shift changes upon addition of NusA ar2.

The residues which were affected by binding to NusA ar2, were also affected when titrating the SKK domain with λ *nutL* RNA (3.7.1), indicating that NusA ar2 in the full-length protein may occlude the RNA binding domains.

Normalized chemical shift changes (section 2.14.6) for SKK upon binding to NusA ar2 as a function of primary sequence is shown in Fig 3.35-A. Based on the chemical shift perturbation results, the amino acids which showed significant changes were plotted on the surface of SKK domain. Fig 3.35-B shows the surface representation of SKK domain highlighting the binding surface and the residues that exhibit chemical shift changes upon binding to NusA ar2.

The dissociation constant K_D was determined from the changes of chemical shifts of $^2\text{H},^{15}\text{N}$ -labeled SKK observed in $^1\text{H},^{15}\text{N}$ -TROSY after gradual addition of the corresponding unlabeled partner (NusA ar2). Changes of chemical shifts of signals in the fast exchange limit (section 2.14.7) were fitted to the equation 2.4. Fitting the curves yielded the calculated values as shown in Fig 3.36.

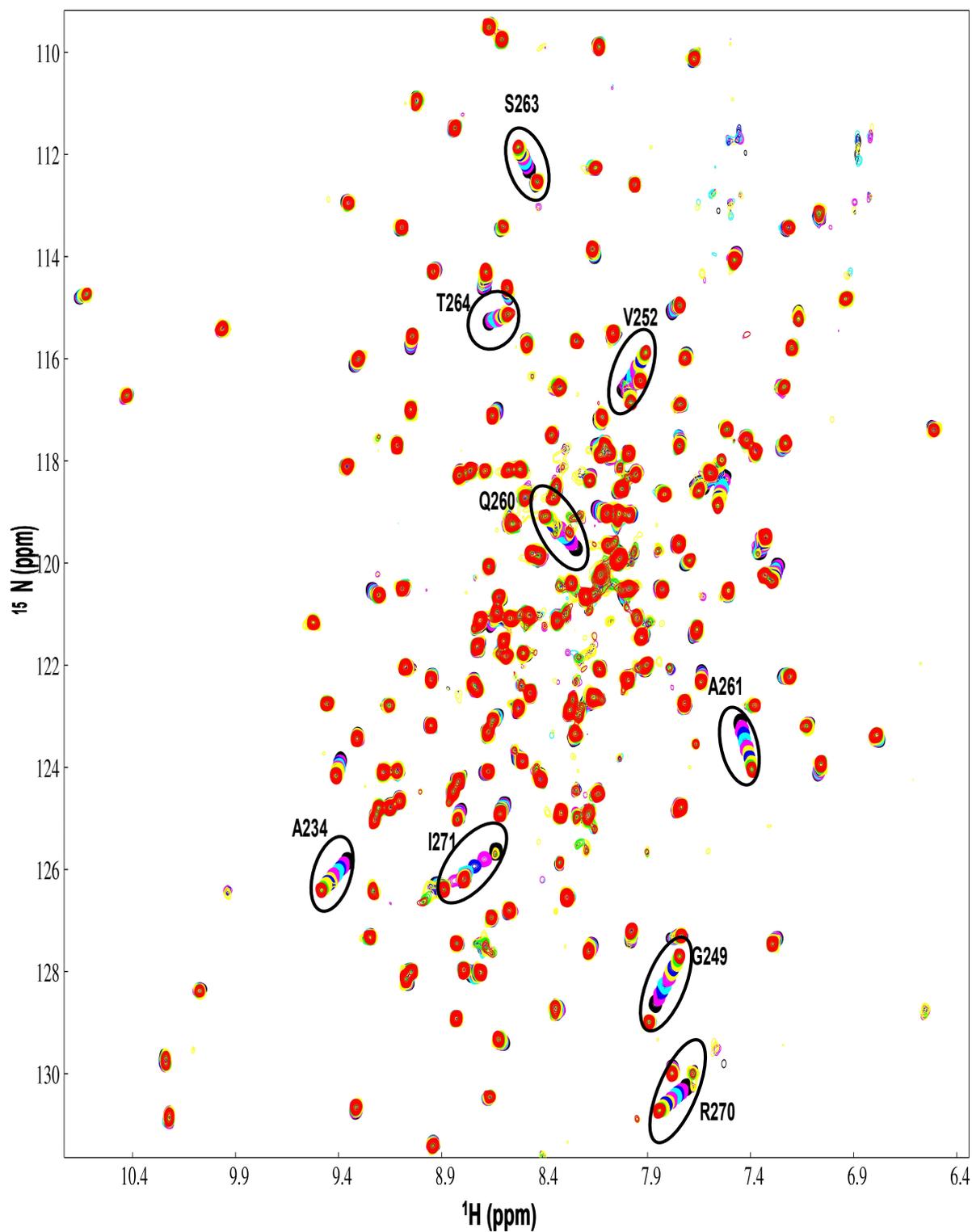


Figure 3.34 Titration of SKK domain with NusA ar2. Overlay of ^1H , ^{15}N -TROSY spectra recorded during the titration of SKK domain with increasing concentrations of NusA ar2. Increasing molar ratios; black 0.0, magenta 0.25, blue 0.50, cyan 0.75, magenta 1.0, yellow 1.5, blue 2.0, yellow 2.5, green 3.0, and red 3.5. Resonances showing significant chemical shift changes are indicated with labeling.

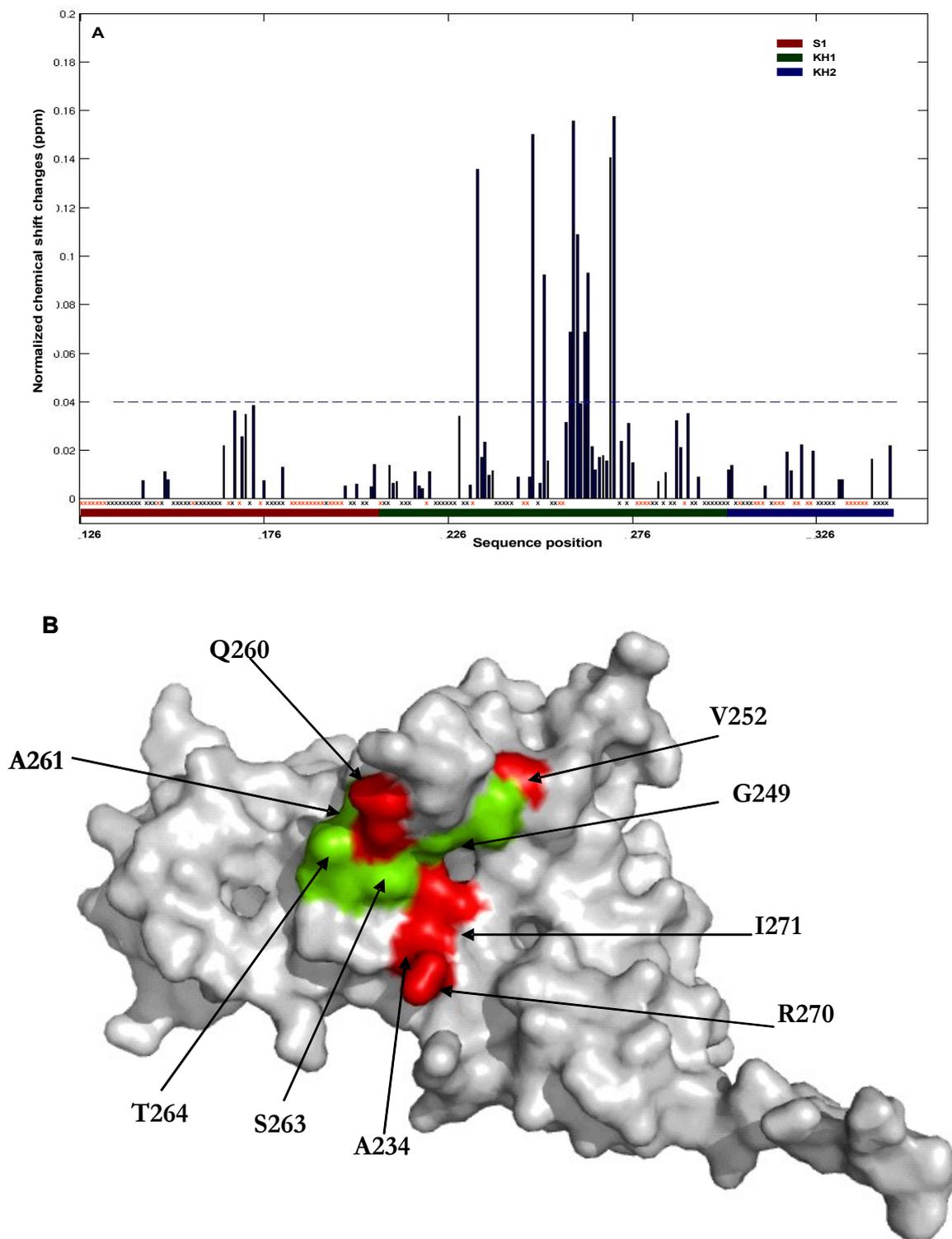


Figure 3.35 (A) Chemical shift changes of SKK upon binding to NusA ar2 as a function of primary sequence; Dotted line represents the significance level of 0.04 ppm; X (red color) = residues not assigned; X (black color) = residues for which $\Delta\delta = 0$. (B) Surface representation of SKK highlighting the binding interface. Residues with resonances showing significant chemical shift changes ($0.04 < \Delta\delta \leq 0.1$) are shown in green, those with $\Delta\delta > 0.1$ are in red.

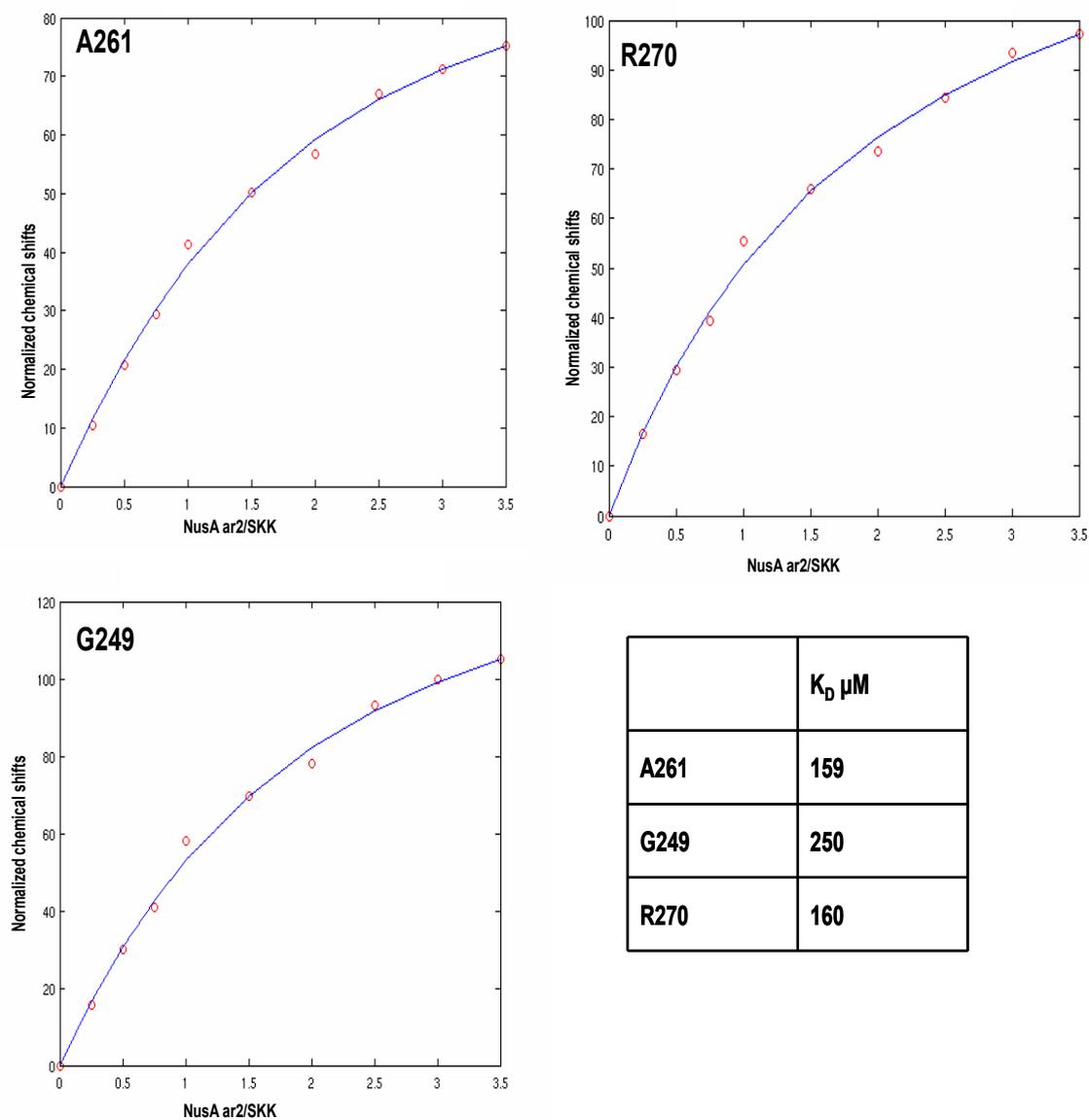


Figure 3.36 Dissociation constants for SKK-NusA ar2 complex. Fitting the curves yielded the calculated K_D values shown in the inset.

3.8.2 Displacement by α -CTD subunit of RNA polymerase

Based on the observation that NusA could bind *nut*-site RNA in the presence of α , but not in the absence, Mah and co-workers had suggested, that there might be a direct interaction between NusA and α subunit of RNA polymerase. The gel mobility shift experiments [Mah et al., 2000] had shown that the binding of NusA to *nut* site RNA is indeed inhibited by the 70 carboxy-terminal amino acids of NusA and suggests that this inhibition could be relieved by an interaction of this portion of NusA with the CTD of the RNA polymerase α subunit.

To characterize the α -dependent RNA binding by SKK domain, we titrated the complex containing SKK + NusA ar2 with α -subunit of RNA polymerase. A series of ^1H , ^{15}N -TROSY spectra of the complex (SKK + NusA ar2) was recorded by gradually adding an increased molar ratio of unlabeled α -CTD.

Our main idea through this experiment is to observe whether α -CTD displaces NusA ar2 from SKK domain and releases the autoinhibition effect of NusA ar2 or not. Overlay of the titration spectra are shown in Fig 3.37-A.

The same set of amide resonances which showed chemical shift changes during the titration of SKK with NusA ar2, were also affected on titrating the complex (SKK + NusA ar2) with α -CTD. But the direction in which the resonances have shifted could be reversed.

Fig 3.37-B, shows one such example of the residue G249 exhibiting such kind of shift. When titrating the SKK domain with NusA ar2 the residue G249 showed significant chemical shift change and thereby falling in the fast exchange regime. Upon titrating with NusA ar2, the resonances of G249 gradually shifted its position from downwards (free state) towards upwards (bound state). During the titration of the complex (SKK + NusA ar2) with α -CTD, the resonances of G249 shifted from upwards towards downwards direction, indicating that α -CTD is displacing NusA ar2 from its bound state, thereby opening the SKK domain for RNA binding.

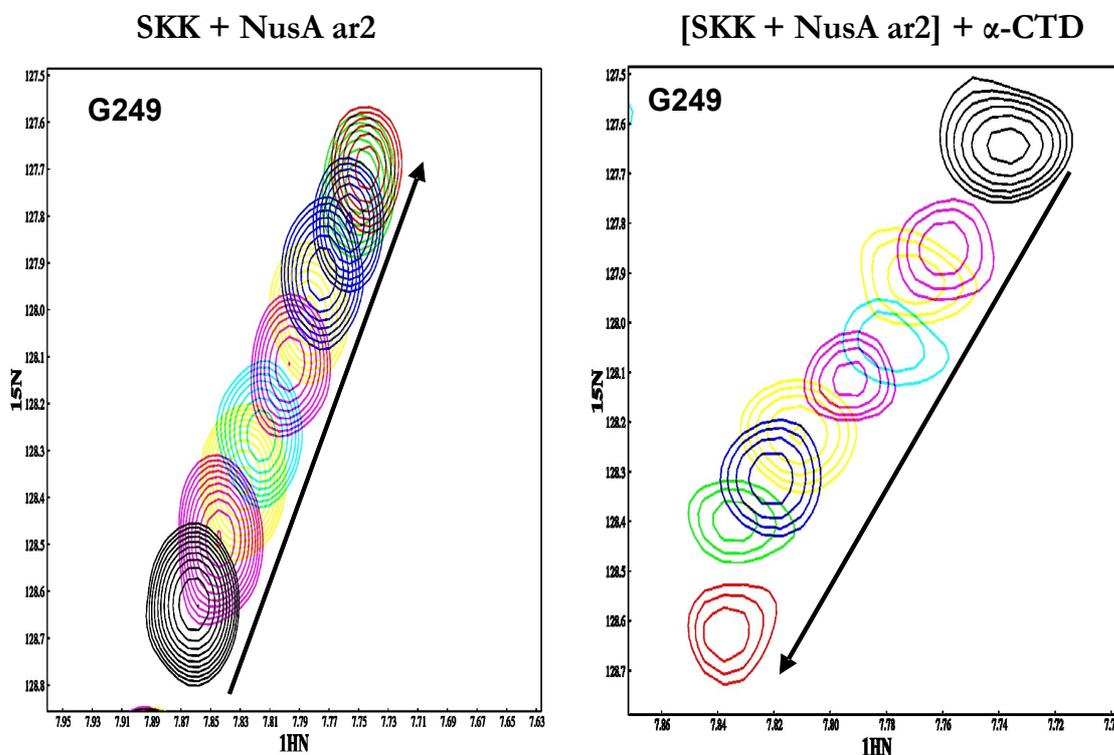
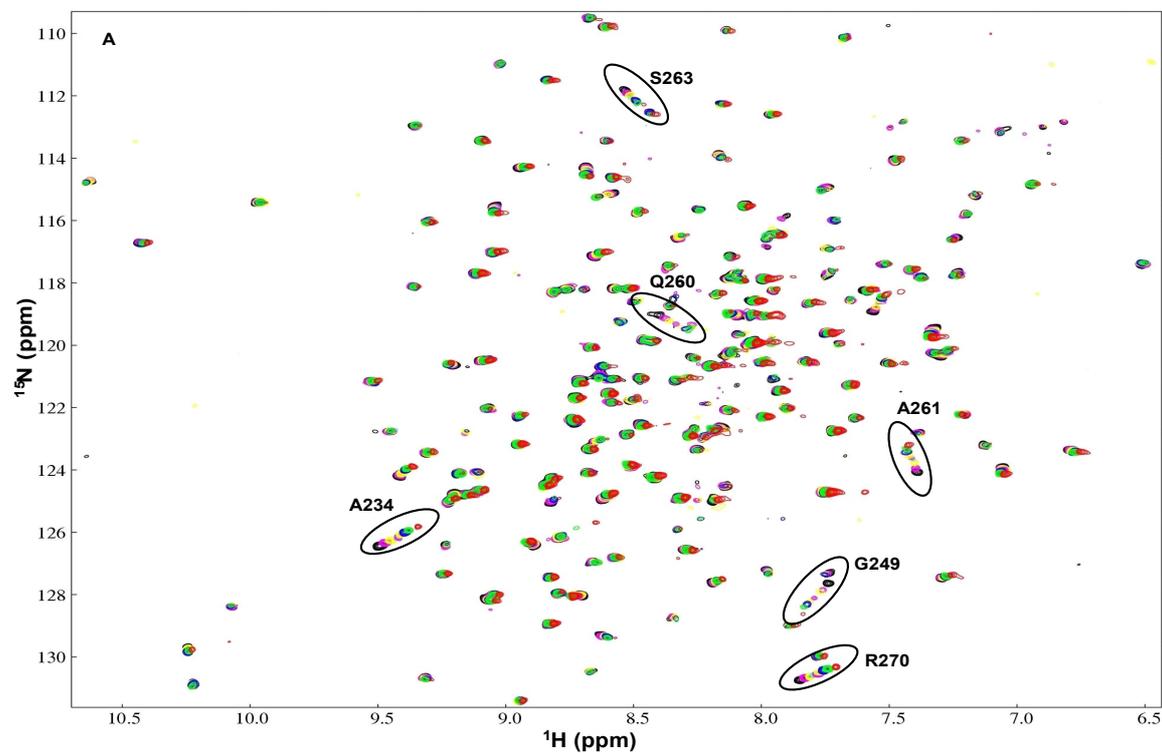


Figure 3.37 (A) Overlay of ^1H , ^{15}N -TROSY spectra recorded during the titration of complex (SKK + NusA ar2) with increasing molar ratios of α -CTD. The resonances which were affected during the titration are indicated. (B) G249 of SKK showing the direction of the shift of resonances in both the titration.

4 Discussions

4.1 Effect of Nus factors on HK022 Nun

The Nun protein of bacteriophage HK022 is a member of the arginine-rich motif family of RNA binding proteins which includes the phage λ N transcription antitermination protein and the HIV Tat and Rev proteins. In contrast to λ N, which suppresses transcription termination, Nun terminates transcription just distal to *BoxB* (1.5.2). Like λ N, action of Nun also requires the host Nus proteins. As described earlier, the *E. coli* NusA protein interacts with the C-terminal region of Nun and stimulates the binding of Nun to *BoxB*. In particular, it has been reported that NusA ar1 which is responsible for binding to phage λ N protein as well as to the C terminus of the RNA polymerase α subunit is also required for the Nun binding [Watnick et al., 1998].

In the context of the reported results, studies have been carried out to determine the interaction between Nun and NusA by NMR titration experiments. To monitor this interaction, three constructs of Nun containing 1-112, 45-112, and 92-112 were used (3.5.1). Upon titrating the Nun constructs with NusA ar1, no detectable changes in chemical shifts as well as no new signals have been observed. Based on these experimental results from NMR, it supports the idea that there might not be any direct interaction between NusA ar1 and Nun.

To optimize the experimental conditions we had used two buffer conditions with different pH (10 mM KPO_4 , pH 6.4, 50 mM NaCl / 50 mM NaPO_4 , pH 7.0, 50 mM NaCl). The results obtained with these different buffers remained the same. As the binding free energies associated with the formation of macromolecular complexes are generally extremely sensitive to ionic strength, we had performed the titration with three different salt conditions (0, 50, and 100 mM NaCl respectively) in the above mentioned buffer, to see whether there are any observable changes due to the different salt concentration or not. However, we have not observed any significant perturbations in the chemical shift during the titration.

As the C terminus of Nun includes three histidine residues that form a potential zinc binding motif [Watnick et al., 2000], all the NMR experiments have been repeated in the presence of zinc, with an aim that it might facilitate the binding of Nun with NusA. Again, no observable changes occurred in the titration even in the presence of zinc.

With the observed results, now the question arises, whether NusA ar1 is critical for Nun binding or not, whether other domains of NusA is also required for binding. With a quest to answer these questions, full length NusA (1-495) have been used for the binding studies.

As the entire complex is prohibitively large for studying by NMR spectroscopy, we replaced the Nun full length with Nun C-terminal domain containing 45-112 amino acids (3.1.2) which corresponds to the interacting region. It has been already shown that NusA binds directly to Nun C-terminal domain by affinity chromatography experiments [Watnick et al., 1998]. ¹⁵N labeled Nun C-terminal domain was titrated with gradually adding unlabeled NusA to a molar ratio of 1:3. On using the full length NusA as well, no significant chemical shift perturbations have been observed.

It has been already reported that λ N forms a complex with NusA ar1 [Prasch et al., 2006]. So, in our case to have a positive control, the titration of λ N with NusA (1-495) have been performed to observe the changes. Even with the full length NusA, one supposed to see few of those changes corresponding to the changes that was observed when the λ N forms a complex with NusA ar1. Same sample conditions were used to avoid the artifacts resulting from the non-similar sample conditions. As expected, distinct resonance changes were seen on λ N upon titrating with NusA full length, indicating clearly the interaction between λ N and NusA.

Based on our experimental results from NMR and in accordance to the lack of chemical shift perturbations from all of the performed titration experiments, it is obvious that there is probably a lack of intermolecular interaction between Nun and NusA. It might be that the interaction between Nun and NusA is also in need of other factors which could facilitate Nun binding to NusA.

Like λ N protein, apart from NusA, Nun also requires additional host factors (NusB, NusE and NusG) for efficient termination, whereas the presence of NusA alone inhibits the termination activity. With the continuing interest to know more about the interaction of HK022 Nun with Nus factors, has provided us a potent driving force to study the interaction

of Nun with various Nus factors (NusG and NusB) by NMR titration studies.

For the interaction studies of Nun with other Nus factors, ^{15}N labeled NusG and NusB have been prepared (3.3 and 3.4). The sample conditions were maintained the same as used to study the interaction of HK022 Nun with NusA.

It has been reported previously, that *E. coli* NusG gene product is required for transcription termination by phage HK022 Nun protein *in vivo* [Burova et al., 1999]. But so far, there are no experiments showing a direct interaction between Nun and NusG. With this as a subject, we proceeded to study the interaction of NusG with Nun. From the overlay of ^1H , ^{15}N -HSQC experiments (Fig 3.24), no dynamic changes were observed in all stages of titration. This indicates that no direct binding between NusG and Nun exist. Therefore, the role of NusG action in HK022 Nun mediated termination might be not directly interacting with Nun but could be in presence of other Nus factors, NusG might stimulate Nun termination.

Mutational studies had described that mutation in *nusB* genes blocks both Nun and N action *in vivo* [Friedman et al., 1976]. But up to now, it has not been shown, whether there is any direct binding of Nun to NusB. Hence, we wanted to observe the interaction between NusB and HK022 Nun by NMR. However, titration of NusB upon adding increased molar ratio of Nun, does not cause any significant detectable changes (Fig 3.25). The interaction studies was carried out with the full length and as well as with C-terminal domain of Nun alone, but the results remained the same, leading to the conclusion that there is no direct interaction between HK022 Nun and NusB.

In the termination pathway, HK022 Nun uses the host factors NusA, NusB, NusG and NusE. Nevertheless, the involvement of *E. coli* host factors in Nun termination is still not clearly explained. Based on our NMR experimental results, we were unable to observe a direct binding of any of these Nus factors to HK022 Nun.

Based on the published results it is clear that, Nun termination is facilitated by Nus host factors, but from our results we deduce that the facilitation is not by a direct binding to HK022 Nun. Therefore, the proposed role of NusA in Nun mediated termination might be different and the role of NusG and NusB still remains to be solved.

4.2 Backbone assignment of RNA binding domains of NusA (SKK)

The central S1 and KH domain regions of NusA (SKK) are involved in interactions with *nut* site RNA and are required for both transcription termination and antitermination. In order to explore the details of the binding surface on SKK domain upon *nut* RNA binding, we primarily proceeded to obtain sequence specific backbone resonance assignment.

Deuterium labeling strategy (2.10) was used for the SKK domain to increase the sensitivity and the resolution in triple resonance experiments. The labeling strategy results in deuterium incorporation throughout a protein in a roughly site-independent manner (uniform or random labeling). One of the significant advantage of deuteration is that many cross-relaxation pathways are removed, thereby reducing the overall resonance line widths and spin diffusion effects in the system.

To achieve high rate of deuteration for the interaction studies, media containing > 99.9 % D₂O was used. The main problem to express a deuterated protein is the incorporation of ²H which reduces growth rate of organisms up to 50 %. Since, expression of highly deuterated protein also requires stepwise adaptation of the bacteria to the high deuteration level, the SKK protein production with deuterated media usually resulted in significantly lower yield than that with nondeuterated media.

For the backbone resonance assignment of SKK domain, TROSY-based triple resonance experiments with ²H,¹⁵N,¹³C-labeled SKK was carried out which was superior than the conventional triple-resonance experiments. Line broadening at higher magnetic fields which is a manifestation of increased transverse relaxation rates and deterioration of the sensitivity in triple-resonance experiments, has been largely suppressed by using TROSY technique. During the course of study, TROSY combined with deuterated SKK domain was used for the backbone resonance assignments.

TROSY based triple-resonance experiments were recorded to allow sequential assignment of the backbone of SKK domain. Typically these experiments include tr-HNCO, tr-HNCA, tr-HN(CA)CO, tr-HN(CO)CA, tr-HNCACB and tr-HN(CO)CACB which are predominantly run as 3D experiments, recording the chemical shifts of ¹H^N, ¹³C and ¹⁵N.

During the course of finding sequential connectivities of the backbone resonances, we had built chemical shift “clusters” by comparing and correlating several heteronuclear 3D experiments, so that each cluster is composed of the correlated backbone chemical shifts of one amino acid residue and of its preceding or following residue. We then link these clusters to obtain sequential stretches of chemical shift sets of amino acid residues, starting from the residue *i* and looking for *i*-1 and so on. Sequential connectivities were also confirmed by the identification of NOE cross-peaks between sequential H^N groups, using a 3D-15N-HSQC-NOESY and 3D-NNH-NOESY.

For a variety of reasons, even modest increase in protein size greatly complicate the assignment process and the same was observed with the SKK domain as well. The SKK domain consist of 222 residues, 9 of which were proline residues and 7 residues in the N-terminal part belongs to the tag region, a maximum of 206 backbone amide ¹H^N-¹⁵N correlation peaks would be predicted in the TROSY spectra of SKK domain (3.6). Sequence specific resonance assignments were made for 166 out of 206 residues (80.5%).

The problems encountered during the assignment procedure is depicted in Fig 4.1, showing the strip plots from amino acid T198-K201 derived from HN(CO)CACB and HNCACB.

In the strip plot we could observe that its possible to walk along the protein backbone only by using the C^α connectivities because C^β [in position (*i*) and (*i*-1)] are missing for these residues. But in this case, its not possible to unambiguously assign these residues based on C^α resonances, since the chemical shift of C^α resonances could match to various other residues in the amino acid sequence as well.

In the process of sequential assignment, always two chemical shifts were matched, and in case of multiple possibilities, we carry out parallel searches for each possible connection until one path leads to the nearest check point.

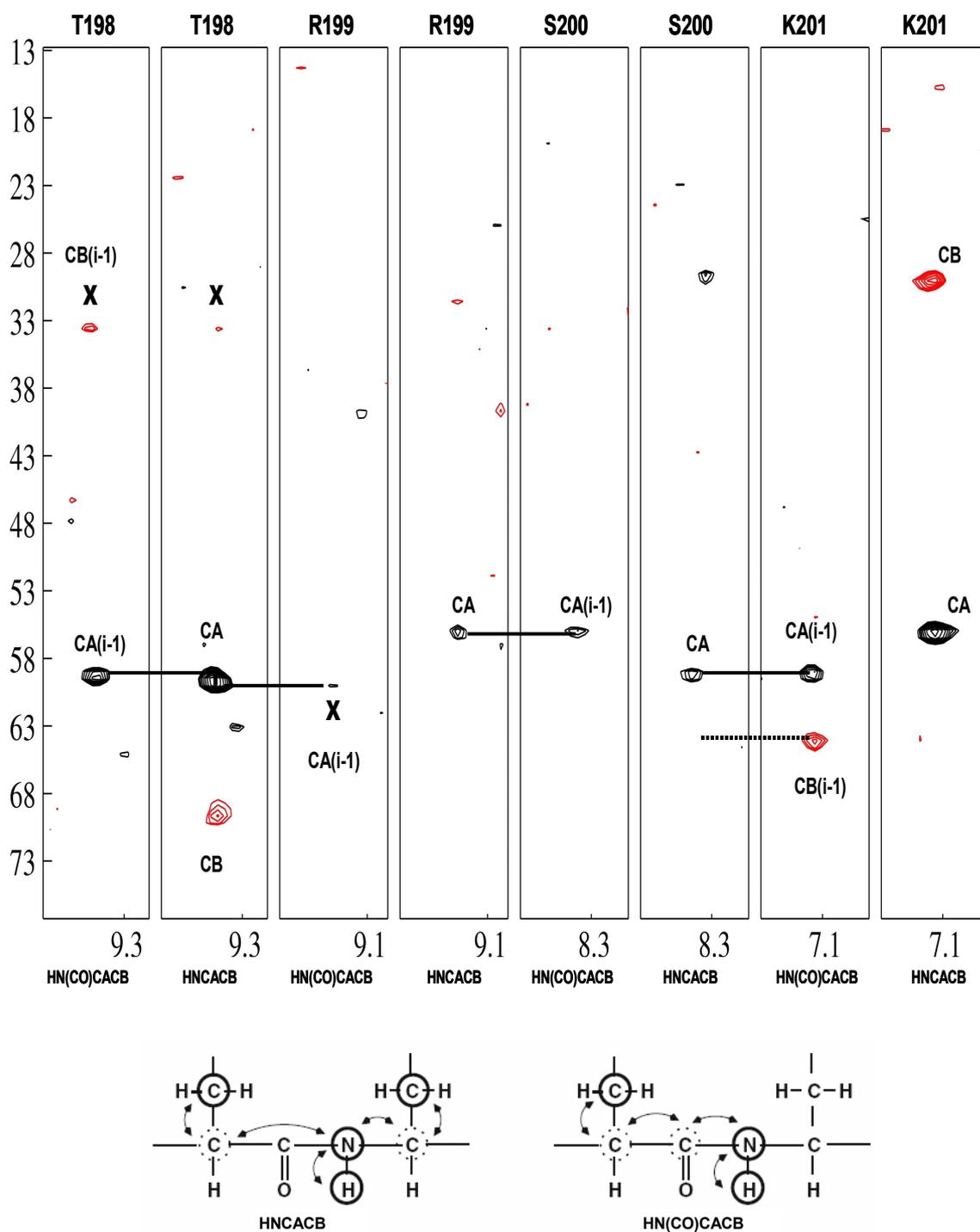


Figure 4.1 Strip plot showing the difficulties faced during resonance assignments. Strips of HN(CO)CACB and HNCACB spectra of ^2H , ^{13}C , ^{15}N uniformly labeled SKK domain (T198-K201). Strips from two spectra are shown, corresponding to a single amino acid. Several of these strips are placed in a row to show the sequential connectivities from each amino acid to the preceding one. The coherence transfer in both of these experiments for a pair of consecutive residues are shown below. The arrows indicate the magnetization transfer pathway. The dotted line represents break of further connectivities. (x) indicates the peaks which are near to noise level.

Because of the large size of the protein, chemical shift degeneracy (for ex., missing of C^β chemical shift) poses a challenging problem. To overcome this problem, we utilized the 3D-15N-HSQC-NOESY and 3D-NNH-NOESY spectra to assign those residues where the chemical shifts of C^α or C^β are missing by observing the NOE cross peaks. In a NOESY spectra, the cross peaks indicate which protons are close in space and also correlates protons which are distant in the amino acid sequence but close in space due to the tertiary structure. The intensity of the NOE is in first approximation proportional to $1/r^6$, with r being the distance between the protons. The presence of a NOE peak is direct evidence that 2 protons are within 5 Angstroms (5 Å) through space. Assignment of resonances is achieved by detecting the sequential connectivities between amide protons in the bonding network between nuclei. Supporting information for the sequential connectivities which was obtained from the NOESY spectra is represented as a strip plot in Fig 4.2.

By using the NOESY spectra, we were able to obtain the sequential connection of spin systems by observing the cross peaks from the amide protons of one residue to the amide, alpha, or beta protons of the next residue. However, to be sure whether the NOEs which were observed are of sequential connectivity or not, we refer back to HNCACB and observe the C^α chemical shift, combining both of these allow strong peak association for T198-R199-S200-K201. Though using TROSY triple-resonance experiments combined with 3D-NOESY experiments, we still were not able to assign all the resonances. The main difficulty was the missing resonances for some of the residues and degenerate resonances for some other, for which there was no NOESY cross peaks as well. The big stretch which was not assigned in SKK domain lied in the region of S1 domain from residue L183-V197 which constitutes the β4 and β5 sheet of S1 motif.

The possible reasons for the missing resonances could be cross-peak overlap, incomplete deuteron amide exchange or exchange peak broadening or with protein dynamics, where changes in mobility of protein occurs over a wide time scale. It might be possible that the residues are not locked in a more rigid conformation leading to the broadening of the resonance. It could also be possible that during the course of measurements, protein degradation occurs which leads to a reduction in the signal due to sample instability which will then affect the line shape. In some experiments, low signal to noise ratio also hindered us to assign those peaks, which were near to the noise level.

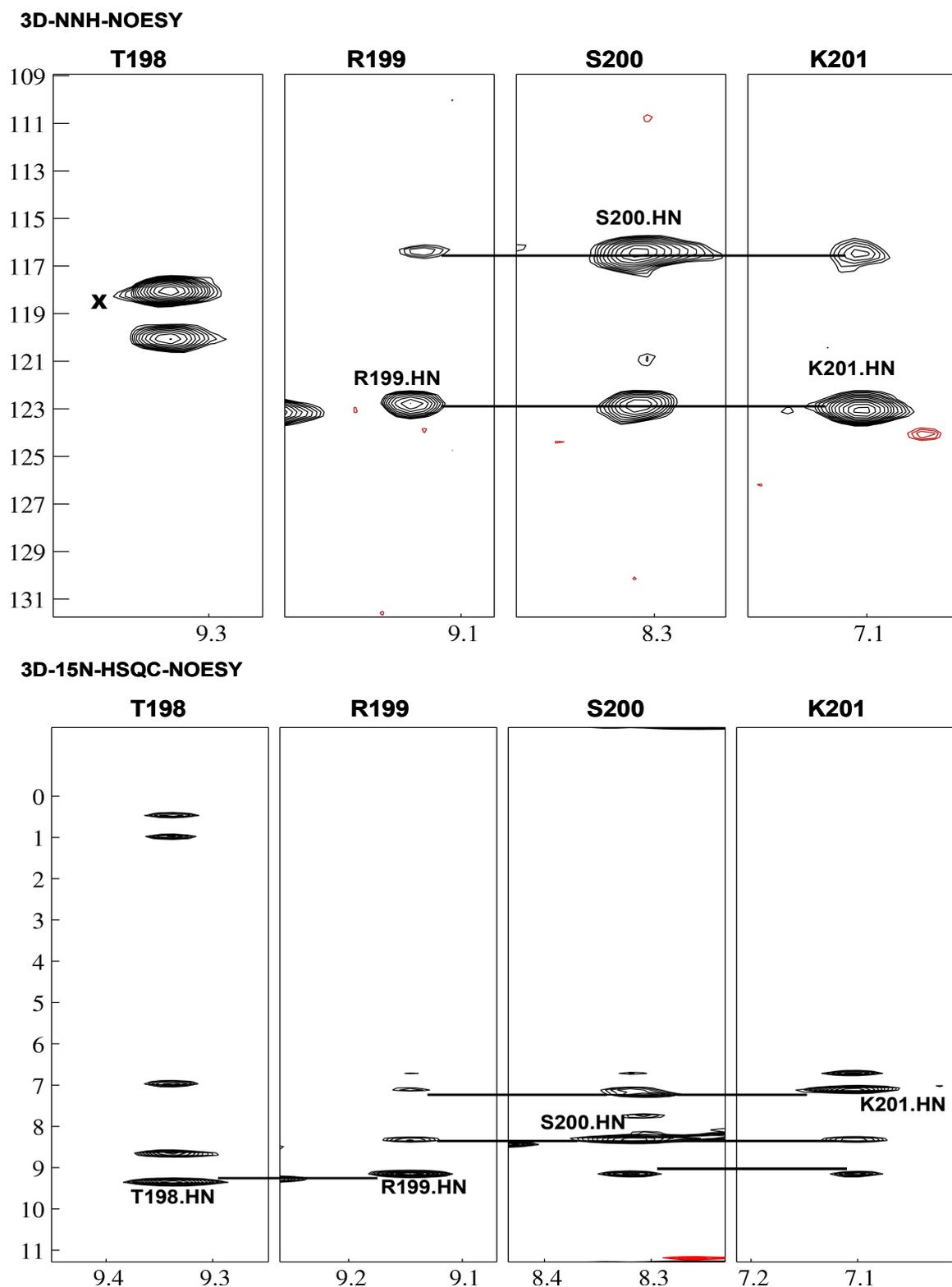


Figure 4.2 NOESY Strip plots of ^2H , ^{13}C , ^{15}N uniformly labeled SKK domain to trace the sequential connectivities. The cross peaks are indicated by connecting lines. The strips show that residues T198-K201 make a series of backward and forward NOEs, indicating that this region could be sequentially connected; (x) indicates the peaks of long range NOEs.

4.3 RNA binding studies on SKK domain

The *E. coli* NusA protein has been the subject of considerable study. NusA modulates the ability of transcription to stop at RNA hairpins which cause RNA polymerase to first pause and then, sometimes, terminate. It also forms part of a complex of proteins associated with RNA polymerase during transcription extension.

Biochemical studies [Mah et al., 1999] on the *E. coli* NusA protein have shown that NusA has RNA-binding properties consistent with the presence of a ribosomal S1 protein-like module and two KH RNA-binding motifs identified from sequence comparisons, which is referred as SKK domain during this study. Electrophoretic mobility-shift assays (EMSA) showed that the truncated *E. coli* NusA protein binds various *nut*-like RNA species with high affinity [Arnvig et al., 2004]. By anisotropic fluorescence titration experiments [Prasch et al., 2008, in revision] it has been proposed that RNA binding domain of NusA shows higher affinity towards λ *nutL* RNA when compared with λ *nutR* RNA. Though various studies imply direct NusA-RNA interaction, till now the exact binding region on SKK domain is not known.

On these basis, we drived our focus to understand the RNA recognition by SKK domain. In order to elucidate the sites of SKK domain that bind *nut*-site RNA, ^2H , ^{15}N -uniformly labeled SKK domain was used for the titration studies. The labeled SKK domain was titrated with increasing molar ratio of unlabeled λ *nutL* RNA. A series of ^1H , ^{15}N -TROSY spectra of the sample were taken after each titration. Upon titration with λ *nutL* RNA, nearly all $^1\text{H}^{\text{N}}$ and ^{15}N resonances of KH1 domain were affected (3.7.1).

There is a continuous stretch of perturbed residues between R210-W276 was seen. These residues constitute mainly on three stranded β sheets on KH1 domain and partly some of the perturbed residues are also seen on four helices on one side of the KH domain. Few residues such as I211, E215, E218, K235, V238, R132, T258, R270, D272, S295, V297 showed significant perturbation upon binding to λ *nutL* RNA, suggesting that the dissociation of the protein/RNA complex is fast on the NMR time scale. On the other hand, some of the resonances like G192, I236, C251, G267, M288, I318 etc., disappears immediately upon titration with 0.25 molar equivalence of λ *nutL* RNA indicating a intermediate exchange regime.

All together, the resonances showing comparatively large chemical shifts and from the resonances which were disappearing, we could say that all the perturbed residues mainly falls in the KH1 region of RNA binding domain of NusA.

Based on the results of chemical shift changes, we suggest the involvement of residues in KH1 domain in complex formation with λ *nutL* RNA. This result is in agreement with the evidence that the KH1 domain, possibly through its RNA binding activity, makes a significant contribution to the biological activity of NusA [Worbs et al., 2001].

The fast exchange regime is by far the most useful for the measurement of dissociation constants from changes in chemical shift. In this regime, the shift of resonances can be easily traced from the resonance of the free state towards the resonance of the bound state (2.14.5). Hence, the resonances which are in the fast exchange region upon titrating SKK domain with λ *nutL* RNA was used for K_D determination. The estimated K_D was around 10 μ M with an error up to a factor of 10 (3.7.3). Unless the titration measurements are performed with the lowest concentration as possible, it is difficult to obtain an accurate and precise values of K_D by NMR. On the other hand, during our titration studies (SKK + λ *nutL*) it was hard to find peaks which show a real fast exchange, as most of the peaks are in between intermediate and fast exchange limit. But our estimated value, gave us a hint that the λ *nutL* RNA has relatively weak binding to SKK domain with a K_D of ~ 10 μ M, which is in agreement with the previous studies.

Following the results of anisotropic fluorescence experiments from our group [Prasch et al., 2008, in revision], it has been proposed that λ *nutL* spacer region is specific for binding to SKK domain. So, we characterized the interaction of λ *nutL* spacer with SKK domain by recording a series of NMR titrations. On addition of λ *nutL* spacer on SKK domain, changes in chemical shifts were observed which could be comparable with the SKK- λ *nutL* RNA titration experiments (3.7.4). Though, not all the resonance shifts which were observed during SKK- λ *nutL* RNA titration is seen with λ *nutL* spacer, the ratio of similarity lies on the upper level. This leads us to a conclusion, that λ *nutL* spacer plays a major role in binding to SKK. But, based on the results from the titration of SKK- λ *nutL* RNA, it implies that though λ *nutL* spacer makes a major contribution in binding to SKK, some other nucleotide residues either from *BoxA* or from *BoxB* are required for specific binding to SKK domain.

4.4 Mapping of RNA binding interface

To map the SKK surface residues which were affected by λ *nutL* RNA, ^1H , ^{15}N -TROSY spectra of ^2H , ^{15}N labeled SKK domain were acquired when λ *nutL* RNA was titrated. Significant chemical shift changes and disappearance of most of the signals in the ^1H , ^{15}N -TROSY spectrum provided the first NMR evidence for an actual interaction event occurring between SKK and λ *nutL* RNA.

As mentioned previously, in cases of fast exchange associated with weak binding or a small change in chemical shift, a single peak at the weighted average chemical shift is observed. As a first step towards mapping the binding surface on SKK protein, normalized chemical shift changes were determined (2.14.6/3.7.2). It was clear from the chemical shift perturbation that a cluster of amino acids in the region of KH1 domain is affected, and the chemical shift changes for some of the residues could not be determined as the corresponding signals disappeared during the titration, indicating that these residues show intermediate exchange in contrast to all other residues whose exchange behavior is fast on the NMR time scale.

The Fig 4.3 (a) shows the ribbon representation of S1+KH1+KH2 domain and all the residues which were affected on interacting with λ *nutL* is represented in red color. This clearly shows us that KH domains plays a role in complex formation with λ *nutL* RNA and the S1 domain remains unaffected. A schematic representation is shown in Fig 4.3 (b) highlighting all the amino acids which were affected on titrating with λ *nutL*. The shift mapping indicates two sites with significant shifts, which constitutes KH1 and KH2 domain, and among the two sites the more extensive was in KH1 domain which is consistent with the argument saying that KH1 domain is a prime candidate for specific RNA binding [Worbs et al., 2001]. Mutational analyses [Mah et al., 2000] had shown that R199A mutation in *ecoNusA* abrogate the N-mediated antitermination and reduce interaction with N-*nut* site complex. It has been also suggested that R199 is completely buried in the S1-KH1 interface, a direct RNA contact through this residue can be excluded, which is consistent with our experimental results, as no significant chemical shift changes were observed with the R199 residue on interacting with λ -*nutL* RNA.

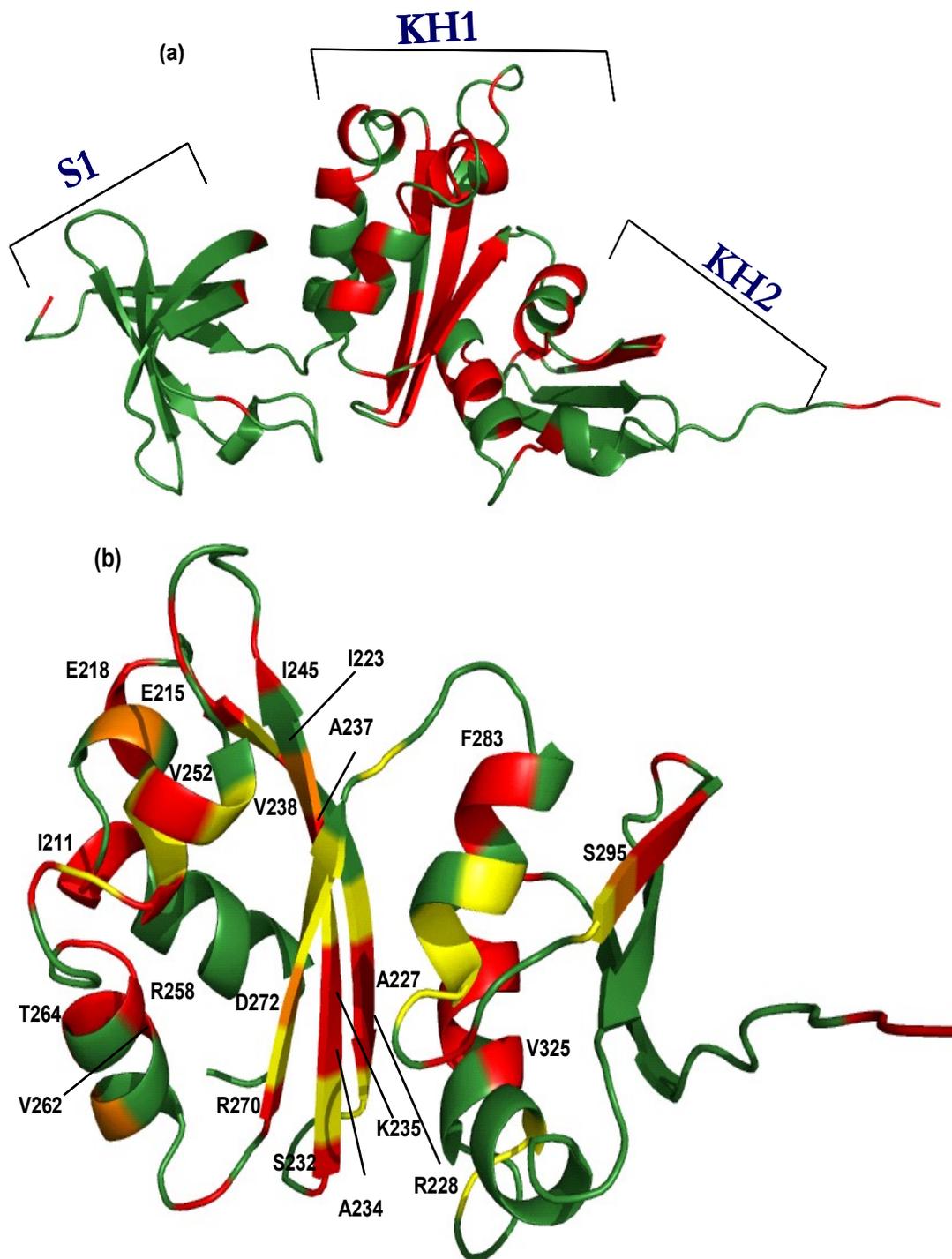


Figure 4.3 (a) Ribbon representation of SKK domain, where the binding region on SKK upon interacting with λ *nutL* RNA is shown in red color. (b) Mapping of SKK residues showing significant chemical shifts when λ *nutL* RNA binds. The S1 domain is unaffected and are not displayed for clarity. Residues with $\Delta\delta$ in between 0.4 and 0.1 ppm are colored red, and residues whose $\Delta\delta > 0.1$ are colored orange, while those which disappear upon interacting are colored in yellow. Some significant perturbed residues are numbered in black. The shift mapping indicates KH1 site with significant shifts.

4.5 Regulation of NusA-RNA interaction by autoinhibition

Reported results [Mah et al., 2000] show that the α subunit of *E. coli* RNA polymerase and the λ N protein binds to the carboxy-terminal regions of NusA, suggesting that α -CTD and λ N may act in similar ways to control the binding of NusA to RNA, and that interaction of α -CTD with NusA promotes the association of NusA with RNA. It has also been suggested that an interaction of NusA with α -CTD in a transcription complex would allow NusA to bind the nascent transcript and stimulate pausing and termination by RNA polymerase.

Recently it was shown that the 79 carboxy-terminal residues of NusA (NusA ar2) hinders RNA interaction in isolation but is sequestered in transcription complexes by the C-terminal domain (CTD) of RNAP subunit α (α -CTD) [Mah et al., 2000]. So, with the reported results, we wanted to evaluate the potential autoinhibition effect of NusA ar2, on RNA binding domains of NusA.

As a primary step, the interaction between SKK domain and NusA ar2 have been studied. To observe the autoinhibition effect, NusA construct lacking the N-terminal domain and the two acidic-repeat domains was used, since these regions are not directly involved in RNA binding. Addition of unlabeled NusA ar2 to ^2H , ^{15}N -labeled SKK domain led to significant chemical shift changes in the ^1H , ^{15}N -TROSY spectra for the residues located in the α -helix and β -sheet of KH1 domain (3.8.1). The significant changes are seen for the residues A234, G249, V252, Q260, A261, S263, T264, R270 and I271. Most of the resonances change lies in the α 4-helix of KH1 domain. Apart from the α 4-helix, the significant changes observed for the residue A234 lies in β 2-sheet, G249 in α 3-helix, V252 in α 1-helix and R270/I271 in β 1-sheet.

All the resonances change which were observed are in the fast-exchange regime on the NMR time scale and thus could be traced easily during titration. Relatively small number of shifted peaks indicates a comparatively small binding surface, which is consistent with the weak interaction (Fig 3.35). The various resonances change which were observed with SKK + NusA ar2 was also seen in the SKK + λ *nutL* RNA titration. But compared with the long stretch of residues which were affected on SKK domain upon interacting with λ *nutL* RNA, the changes seen in SKK domain upon interacting with NusA ar2 is relatively small. Though the binding surface of NusA ar2 on SKK domain is small, it could still cause an interference of RNA binding by NusA. It has been already suggested that one or more of the RNA binding domains

of NusA may be occluded by the second HhH motif or other determinants within the 79 carboxy-terminal amino acids of NusA [Mah et al., 2000]. So, based on our results, the significant resonance changes are seen only in KH1 domain, indicating that this domain could be occluded by NusA ar2 and thereby inhibiting RNA binding by NusA.

By plotting the chemical shift changes as a function of ligand concentration and fitting the equation 2.4 in a two state model yielded a $K_D \sim 150\text{-}250 \mu\text{M}$, which corresponds to a weak affinity interaction (Fig 3.36). The stability of protein-protein complex or protein-RNA complex is strongly influenced by entropic contributions. During the study of autoinhibition effect of NusA ar2, entropy changes (ΔS) between the two states of opening a complex (SKK + λ *nutL* RNA) and forming a new complex (SKK + NusA ar2) can be different. Hence, the calculated K_D values for SKK + NusA ar2 complex may not be exact.

By affinity chromatography experiments [Mah et al., 2000], it has been illustrated that there is a direct interaction of α -CTD and NusA, and thereby stimulating RNA binding by NusA. To explore the effect of α -CTD on NusA, interaction of the complex (SKK + NusA ar2) with α -CTD subunit of RNA polymerase have been studied. The expected output from this experiment should be, that on titrating the complex with α -CTD subunit of RNA polymerase, the NusA ar2 has to be displaced from its bound state, thereby facilitating the RNA binding by NusA.

Hence, to observe this output, a series of $^1\text{H}, ^{15}\text{N}$ -TROSY experiments on the complex (SKK + NusA ar2) were measured by gradually adding increased molar ratio of unlabeled α -CTD subunit of RNA polymerase. From the overlay of $^1\text{H}, ^{15}\text{N}$ -TROSY spectra, we could observe clearly the displacement of NusA ar2 from the SKK domain and thereby allowing SKK domain to interact with RNA.

To summarize the autoinhibition effect of NusA ar2, our experimental results by NMR shows that NusA ar2 has an effect on SKK domain by affecting those resonances which are also involved in the RNA binding, directly indicating that NusA ar2 inhibits RNA binding by NusA. This autoinhibition effect of NusA ar2 is relieved when α -CTD subunit of RNA polymerase is available in the system, there by NusA ar2 primarily focuses on α -CTD subunit of RNA polymerase by binding to it. In this manner, NusA ar2 releases its autoinhibitory effect on SKK domain and facilitates RNA binding.

5 Summary

In phage λ , antitermination is initiated by the λ -encoded N protein which recruits a number of host proteins called Nus factors. Several of these host proteins which are essential for effective transcription termination and antitermination have been identified, these includes NusA, NusB, NusE and NusG. The subject of this work is mainly focused on characterization of interactions between various Nus host factors in the termination and antitermination system by NMR spectroscopy.

Like N protein, Nun also requires additional host factors for efficient termination. It has been reported already that NusA interacts with C-terminal region of Nun and also that Nun binding to NusA requires NusA ar1 region. On the basis of these results, ^1H , ^{15}N -HSQC spectra were recorded to monitor the interaction between HK022-Nun and NusA ar1. NMR titration experiments between Nun and NusA clearly showed a lack of chemical shift perturbations. Therefore, it can be concluded that there is no intermolecular interaction between Nun and NusA ar1. Up to date, no information about the interaction between Nun and NusG as well as Nun and NusB are known. Titration experiments between Nun and both Nus factors, have also revealed no direct interaction. Altogether, it can be deduced that there might be no direct interaction between Nun and NusA ar1, and NusG, and NusB.

The NusA transcription elongation protein, which binds *nut* site RNA, contains sequences corresponding to the S1 and KH classes of identified RNA binding domains. To gain comprehensive insights into binding surface on SKK domain upon *nut* RNA binding, backbone resonances of SKK domain was assigned using sequential C^α , C^β and CO chemical shift information derived from an array of TROSY based triple-resonance experiments. With virtually complete backbone assignment (80.5 %) of the SKK domain it was possible to characterize the interaction between SKK domain and λ *nutL* RNA by NMR titration experiments. Significant chemical shift changes observed on SKK domain upon addition of unlabeled λ *nutL* RNA, had reflected a direct interaction. Mapping of chemical shift perturbations on SKK domain revealed that the RNA binding interface is mainly located in the KH domains. The results implied a sequence-specific RNA binding.

In the free state, NusA cannot bind to RNA. Once α -CTD of RNA polymerase is bound to NusA the RNA binding inhibition is released. A direct interaction between α -CTD and NusA ar2 have been reported and therefore NusA ar2 could be a prime candidate for inhibiting the RNA binding of NusA. To further evaluate the autoinhibition effect of NusA ar2 on SKK domain, titrations between NusA ar2 and SKK domain have been performed. Upon NusA ar2 binding, notable chemical shift changes were observed in the KH1 region of SKK domain. The residues which were affected on binding to NusA ar2 were also affected during the SKK and λ *nutL* titration experiments. The results are in good agreement with the proposed idea that NusA ar2 possibly occludes the RNA binding domains of NusA.

To investigate the effect of α -CTD on RNA binding by NusA, titration of the complex containing SKK domain and NusA ar2 by gradually adding an increased molar ratio of α -CTD have been carried out. On addition of α -CTD, it was clearly observed that α -CTD displaced NusA ar2 from the complex suggesting that the inhibition of RNA binding by NusA ar2 could be released by α -CTD.

6 Zusammenfassung

Im Phagen λ wird die Antitermination durch das λ -codierte N Protein initiiert, das eine Vielzahl von Wirtsproteinen, die sogenannten Nus Faktoren, rekrutiert. Bisher sind nur einige von diesen Wirtsproteinen, die für eine effiziente Termination und Antitermination der Transkription wichtig sind, identifiziert. Dazu gehören NusA, NusB, NusE und NusG. Gegenstand dieser Arbeit war die Charakterisierung von möglichen Wechselwirkungen zwischen den verschiedenen Nus Wirtsfaktoren im Terminations- und Antiterminationssystem mit Hilfe von NMR Spektroskopie.

Ebenso wie das N Protein benötigt auch Nun zusätzliche Wirtsfaktoren für eine funktionierende Termination. Bisher wurde nur gefunden, dass NusA mit der C-terminalen Region von Nun interagiert, wobei für die Bindung von Nun an NusA die NusA ar1 Region benötigt wird. Darauf aufbauend wurde mit Hilfe von $^1\text{H},^{15}\text{N}$ -HSQC Spektren die Wechselwirkung zwischen HK022-Nun und NusA ar1 untersucht. Dabei konnten keinerlei Veränderungen der chemischen Verschiebung während der Titration beobachtet werden. Folglich findet keine Interaktion zwischen NusA ar1 und Nun statt. Über die Wechselwirkung zwischen Nun und NusG oder NusB gibt es bis jetzt noch keine Informationen. Die in dieser Arbeit durchgeführten Titrationsstudien mit Nun und diesen beiden Nus Faktoren zeigten ebenso keine direkten Interaktionen. Basierend auf den durchgeführten Experimente kann gefolgert werden, dass keine direkten Wechselwirkungen zwischen Nun und NusA ar1, NusG oder NusB vorhanden sind.

Der Transkriptions-Elongationsfaktor NusA, der die *nut* RNA bindet, enthält Bereiche, die zu der Klasse der S1 und KH homologen Domänen gehören und als RNA Bindungsdomänen identifiziert wurden. Um einen detaillierten Einblick in die Bindungsfläche der SKK Domänen bei der Bindung an die *nut* RNA zu bekommen, erfolgte eine sequenzspezifische Zuordnung der Amidresonanzen des Proteinrückgrats. Die fast vollständige Zuordnung (80,5%) ermöglichte nun die Untersuchung der Wechselwirkung zwischen der SKK Domäne und der *nut* RNA mit Hilfe von NMR Spektroskopie.

Bei der Titration der SKK Domäne mit der unmarkierten λ *nutL* RNA konnten deutliche Veränderungen der chemischen Verschiebung beobachtet werden, die auf eine direkte Interaktion schließen lassen. Eine Visualisierung der Veränderungen der chemischen Verschiebung auf der Oberfläche von NusA SKK zeigt, dass die RNA Bindungsfläche hauptsächlich im Bereich der beiden KH Domänen zu finden ist. Dies deutet somit auf eine sequenzspezifische RNA Bindung hin.

NusA kann im freien Zustand keine RNA binden. Erst durch die Bindung an die α -CTD der RNA Polymerase wird diese Selbstblockade aufgehoben. Da Experimente auf eine Interaktion zwischen der α -CTD und NusA ar2 hindeuteten, könnte möglicherweise NusA ar2 die RNA Bindungsstelle blockieren. Um diesen autoinhibitorischen Effekt zu untersuchen, wurde NusA ar2 zu der SKK Domäne titriert. Dabei konnte für diejenigen Aminosäuren eine Veränderung der chemischen Verschiebung beobachtet werden, die auch an der Bindung der RNA beteiligt sind, so dass hierdurch die Hypothese der Selbstblockade durch NusA ar2 bestätigt werden konnte. Eine Titration des Komplexes aus NusA ar2 und der SKK Domäne mit α -CTD zeigte, dass durch die Zugabe von α -CTD NusA ar2 von der RNA Bindungsstelle verdrängt wird und bestätigt damit ebenso die autoinhibitorische Rolle von NusA ar2.

7 Abbreviations

ϵ	molar extinction coefficient
1D	one dimensional
2D	two dimensional
3D	three dimensional
aa	amino acid
A_{280}	absorption at 280 nm
APS	ammonium peroxy disulfate
ARM	arginine rich motif
ar1	acidic repeat 1
ar2	acidic repeat 2
ATP	adenosine-5'-triphosphate
bp	base pair
CSA	chemical shift anisotropy
CTD	carboxy terminal domain
CV	column volume
Da	dalton
DSS	2,2-dimethyl-2-silapentane-5-sulfonic acid
DD	dipole-dipole
DNA	deoxyribonucleic acid
DNase I	deoxyribonuclease I
DTT	dithiothreitol
EMSA	electrophoretic mobility-shift assay
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	ethylenediaminetetraacetic acid
FPLC	fast protein liquid chromatography
Fig	figure
h	hour
HK022	HongKong 022
HPLC	high performance liquid chromatography
HSQC	heteronuclear single quantum coherence
INEPT	Insensitive Nuclei Enhancement by Polarization Transfer
IPTG	isopropyl- β -D-thiogalactopyranoside
kDa	kilo Dalton
K_D	dissociation constant
L	liter
LB	Luria Bertani medium
mRNA	messenger RNA
μ	micro (10^{-6})
μm	micromolar ($\mu\text{mol/L}$)
m	milli (10^{-3})
mAU	milli-Absorption Unit

mL	milliliter
mM	millimolar (mmol/L)
M9	minimal medium
min	minute
M	molar (mol/L)
MWCO	molecular weight cutoff
nm	nanometer
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser effect spectroscopy
nt	nucleotide
NTD	amino terminal domain
NTPs	nucleoside triphosphates
NS	number of scans
Nus	N-utilization substance
nut	N-utilization
OD ₆₀₀	optical density at 600 nm
PAGE	polyacrylamide gel electrophoresis
PEG	polyethylene glycol
PCR	polymerase chain reaction
PDB	protein data bank
PMSF	phenylmethylsulfonylfluoride
ppm	parts per million
put	polymerase utilization
RNase	ribonuclease
RNA	ribonucleic acid
rut	rho-utilization
rpm	rotations per minute
RT	room temperature
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
sec	second
SKK	S1+KH1+KH2
SW	spectral width
SF0	spectrometer frequency used
SAR	Structure Activity Relationships
TEMED	N-,N-,N'-,N'-Tetramethylethyldiamine
TEV	tobacco etch virus
TD	total number of data points
TFA	trifluoroacetic acid
TPPI	time proportional phase incrementation
Tris	tris(hydroxymethyl)aminomethane
TROSY	Transverse Relaxation Optimized Spectroscopy
tsp	transcription stop points
TS2	trace element solution 2
U	unit
v/v	volume by volume
w/v	weight by volume

8 References

- Agnieszka S.P., Barbara S., Antosiewicz A.H., Wegrzyn G and Thomas M.S. (2003)**
Genetic analysis of bacteriophage lambda-N dependent antitermination suggests a possible role for the RNA polymerase alpha subunit in facilitating specific functions of NusA and NusE.
Arch Microbiol. **180**, 161-168
- Ahmad Z and Huang K.P. (1981)**
Dephosphorylation of rabbit skeletal muscle glycogen synthase (phosphorylated by cyclic AMP-independent synthase kinase 1) by phosphatases.
J Biol Chem. **256**, 757-760
- Archer S.J., Ikura M., Torchia D.A and Bax A. (1991)**
An alternative 3D NMR technique for correlating backbone ¹⁵N with side chain H β resonances in larger proteins.
J Magn Reson. **95**, 636-641
- Arnvig K.B., Pennell S., Gopal B., Colston M.J. (2004)**
A high-affinity interaction between NusA and the rrn nut site in Mycobacterium tuberculosis.
Proc Natl Acad Sci U S A. **101**, 8325-8330
- Barik S., Ghosh B., Whalen W., Lazinski D and Das A. (1987)**
An antitermination protein engages the elongating transcription apparatus at a promoter-proximal recognition site.
Cell. **50**, 885-899
- Barkhuijsen H., De Beer W., Bovee M.M.J and van Ormondt D. (1985)**
Retrieval of frequencies, amplitudes, damping factors, and phases from time domain signals using a linear least squares procedure.
J Magn Reson. **61**, 465-481
- Bax A and Ikura M. (1991)**
An efficient 3D NMR technique for correlating the proton and ¹⁵N backbone amide resonances with the α -carbon of the preceding residue in uniformly ¹⁵N/¹³C-enriched proteins.
J Biomol NMR. **1**, 99-104
- Berg K.L., Squires C and Squires C.L. (1989)**
Ribosomal RNA operon antitermination. Function of leader and spacer region boxB-boxA sequences and their conservation in diverse micro-organisms.
J Mol Biol. **209**, 345-358

Bodenhausen G and David J.R. (1980)

Natural abundance nitrogen-15NMR by enhanced heteronuclear spectroscopy.

Chem Phys Lett. **69**, 185-189

Burgess R.R., Erickson B., Gentry D., Griskov M and Hager D. (1987)

RNA polymerase and the regulation of transcription.

New York: Elsevier, pp. 3-15

Burova E., Hung S.C., Chen J., Court D.L., Zhou J.G., Mogilnitskiy G., Gottesman M.E. (1999)

Escherichia coli nusG mutations that block transcription termination by coliphage HK022 Nun protein.

Mol Microbiol. **31**, 1783–1793

Campbell A. (1994)

Comparative molecular biology of lambdoid phages.

Annu Rev Microbiol. **48**, 193-222

Cavanagh J., Fairbrother W.J., Palmer A.G and Skelton N.J. (1996)

Protein NMR Spectroscopy: Principles and Practice.

Academic Press, Inc., San Diego

Christopher A.L., Jonathan M.M and Jeffrey W.P. (2004)

Theory and applications of NMR based screening in pharmaceutical research.

Chem Rev. **104**, 3641-3675

Clubb R.T., Thanabal V and Wagner G. (1992a)

A constant-time 3-dimensional triple-resonance pulse scheme to correlate intra residue H-1(N), N-15, and C-13(') chemical shifts in N-15-C-13-labeled proteins.

J Magn Reson. **92**, 213-217

Clubb R.T., Thanabal V and Wagner G. (1992b)

A new 3D HN(CA)HA experiment for obtaining fingerprint HN-Halpa peaks in 15N- and 13C-labeled proteins.

J Biomol NMR. **2**, 203-210

Court D.L., Oppenheim A.B., and Adhya S.L. (2007)

A new look at bacteriophage lambda genetic networks.

J Bacteriol. **189**, 298-304

Das A. (1992)

How the phage lambda N gene product suppresses transcription termination: communication of RNA polymerase with regulatory proteins mediated by signals in nascent RNA.

J Bacteriol. **174**, 6711–6716

Das A. (1993)

Control of transcription termination by RNA-binding proteins.

Annu Rev Biochem. **62**, 893-930

Das A., Pal M., Mena J.G., Whalen W., Wolska K., Crossley R., Rees W., Hippel P.Y., Costantino N., Court D., Mazzulla M., Altieri A.S., Byrd A., Chatopadhyay S., Devito, B and Ghosh B. (1996)

Components of multiprotein-RNA complex that controls transcription elongation in *E. coli* phage lambda.

Methods Enzymol. **274**, 374-402

DeVito J and Das A. (1994)

Control of transcription processivity in phage λ Nus factors strengthen the termination-resistant state of RNA polymerase induced by N antiterminator.

Proc Natl Acad Sci U S A. **91**, 8660-8664

Dhillon E.K., Dhillon T.S., Lam Y.Y and Tsang A.H. (1980)

Temperate coliphages: classification and correlation with habitats.

Appl Environ Microbiol. **39**, 1046-1053

Dhillon T.S., Dhillon E.K and Lai A.N. (1981)

Genetic recombination between phage HK022, lambda, and phi 80.

Virology. **109**, 198-200

Dodd I.B., Shearwin K.E and Egan B.J. (2005)

Revisited gene regulation in bacteriophage λ .

Curr Opin Genet Dev. **15**, 145-152

Dombroski A.J., Walter W.A and Gross C.A. (1993)

Amino-terminal amino acids modulate sigma-factor DNA-binding activity.

Genes Dev. **7**, 2446-2455

Eisenmann A., Schwarz S., Prash S., Schweimer K and Rösch P. (2005)

The *E. coli* NusA carboxy-terminal domains are structurally similar and show specific RNAP and lambda N interaction.

Protein Sci. **14**, 2018-2029

Engelke J and Ruterjans H. (1995)

Sequential protein backbone assignments using an improved 3D-HN(CA)CO pulse scheme.

J Magn Reson. Series B. **109**, 318-322.

Ernst R.R., Bodenhausen B and Wokaun A. (1992)

Principles of Nuclear magnetic resonances in one or two dimensions.

Oxford University Press

Faber C., Schärpf M., Becker T., Sticht H and Rösch P. (2001)

The structure of the coliphage HK022 Nun protein-lambda-phage boxB RNA complex. Implications for the mechanism of transcription termination.

J Biol Chem. **276**, 32064-32070

Feeney J., Batchelor J.G., Albrand J.P and Roberts G.C.K. (1979)

The effects of intermediate exchange processes on the estimation of equilibrium constants by NMR.

J Magn Reson. **33**, 519-529

Fernandez C and Wider G. (2003)

TROSY in NMR studies of the structure and function of large biological macromolecules.
Curr Opin Struct Biol. **13**, 570-580

Fesik S.W and Zuiderweg E.R.P. (1988)

Heteronuclear Three-Dimensional NMR Spectroscopy. A Strategy for the Simplification of Homonuclear Two-Dimensional NMR Spectra.
J Magn Reson. **78**, 588-593

Fielding L. (2007)

NMR methods for the determination of protein-ligand dissociation constants.
Progress in Nuclear Magnetic Resonance Spectroscopy. **51**, 219-242

Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A., Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G and Kelley J.M. (1995).

The minimal gene complement of *Mycoplasma genitalium*.
Science. **270**, 397-403

Freifelder David. (2001)

University of California, San Diego. Molecular biology. Second edition. Jones and Bartlett Publishers, Inc., U.S.A.

Friedman D.I., Baumann M.F., Baron L.S. (1976)

Cooperative effects of bacterial mutations affecting λ N gene expression: Isolation and characterization of a *nusB* mutant.
Virology. **73**, 119-127

Friedman D.I., Olson E.R., Johnson L.L., Alessi D and Craven M.G. (1990)

Transcription-dependent competition for a host factor: The function and optimal sequence of the phage λ boxA transcription antitermination signal.
Genes Dev. **4**, 2210-2222

Friedman D.I and Court D.L. (1995)

Transcription antitermination: the lambda paradigm updated.
Mol Microbiol. **18**, 191-200

Friedman D.I and Court D.L. (2001)

Bacteriophage lambda: alive and well and still doing its thing.
Curr Opin Microbiol. **4**, 201-207

Friedrich M.S. (1995)

A Model-free algorithm for the removal of baseline artifacts.
J Biomol NMR. **5**, 147-153

Garber M.E., Wei P., Kewal Ramani V.N., Mayall T.P., Herrmann C.H., et al. (1998)

The interaction between HIV-1 Tat and human cyclin T1 requires zinc and a critical cysteine residue that is not conserved in the murine CycT1 protein.

Genes Dev. **12**, 3512–27

Gill S.C and Hippel P.H. (1989)

Calculation of Protein extinction coefficients from amino acid sequence data.

Anal Biochem. **182**, 319–326

Gopal B., Haire L.F., Gamblin S.J., Dodson E.J., Lane A.N., Papavinasasundaram K.G., Colston M.J and Dodson G. (2001)

Crystal structure of the transcription elongation/anti-termination factor NusA from *Mycobacterium tuberculosis* at 1.7 Å resolution.

J Mol Biol. **314**, 1087–1095

Gottesman M.E and Weisberg R.A. (2004)

Little lambda, Who Made Thee?.

Microbiol Mol Biol Rev. **68**, 796-813

Gottesman M.E. (1999)

Bacteriophage λ: The untold story.

J Mol Biol. **293**, 177-180

Grabski A., Mehler M and Drott D. (2003)

Unattended high-density cell growth and induction of protein expression with the overnight express TM autoinduction system.

InNovations. **17**, 3-6

Greenblatt J and Li J. (1981a)

Interaction of the sigma factor and the nusA gene protein of *E. coli* with RNA polymerase in the initiation-termination cycle of transcription.

Cell. **24**, 421-428

Greenblatt J., Mah T.F., Legault P., Mogridge J., Li J and Kay L.E. (1998)

Structure and mechanism in transcriptional antitermination by the bacteriophage lambda N protein.

Cold Spring Harb Symp Quant Biol. **63**, 327-336

Greenblatt J., McLimont M and Hanly S. (1981)

Termination of transcription by nusA gene protein of *Escherichia coli*.

Nature. **292**, 215-220

Greenblatt J., Nodwell J.R and Mason S.W. (1993)

Transcriptional antitermination.

Nature. **364**, 401–406

Greive S.J and von Hippel P.H. (2005)

Thinking quantitatively about transcriptional regulation.

Nat Rev Mol Cell Biol. **6**, 221-232

Gross C., Lonetto M and Losick R. (1992)

Bacterial sigma factors: In transcriptional regulation.

Cold Spring Harbor. Cold Spring Harbor Laboratory Press. 129-176

Grzesiek S and Bax A. (1992b)

Improved 3D triple resonance NMR techniques applied to a 31 kDa protein.

J Magn Reson. **96**, 432-440

Grzesiek S and Bax A. (1993a)

Amino acid type determination in the sequential assignment procedure of uniformly ¹³C/¹⁵N-enriched proteins.

J Biomol NMR. **3**, 185-204

Grzesiek S and Bax A. (1993b)

The importance of not saturating water in protein NMR-application to sensitivity enhancement and NOE measurements.

J Am Chem Soc. **115**, 12593-12594

Gusarov I and Nudler E. (1999)

The mechanism of intrinsic transcription termination.

Mol Cell. **3**, 495-504

Hajduk P.J., Dinges J., Miknis G.F., Merlock M., Middleton T., Kempf D.J., Egan D.A., Walter K.A., Robins T.S., Shuker S.B., Holzman T.F and Fesik S.W. (1997)

NMR-based discovery of lead inhibitors that block DNA binding of the human papillomavirus E2 protein.

J Med Chem. **40**, 3144-3150.

Hall D.A., Vander Kooi C.W., Stasik C.N., Stevens S.Y., Zuiderweg E.R and Matthews R.G.(2001)

Mapping the interactions between flavodoxin reductase and cobalamin- dependent methionine synthase.

Proc Natl Acad Sci U S A. **98**, 9521-9526

Henkin T.M and Yanofsky C.Y. (2002)

Regulation by transcription attenuation in bacteria: how RNA provides instructions for transcription termination/antitermination decisions.

Bioessays. **24**, 700-707

Herbrüggen S.T and Sorensen OW. (2000)

Clean TROSY: compensation for relaxation-induced artifacts.

J Magn Reson. **144**, 123-128

Hershey A.D and Dove W. (1971)

Introduction to lambda.

Cold spring harbor, New York, Cold Spring Harbor Laboratory Press. 3-11

Hershey A.D. (1967)

Annu. Report of the Director of the Genetics Research Unit. Washington, DC.

Horvath C.G., Preiss B.A and Lipsky S.R. (1967)

Fast liquid chromatography: an investigation of operating parameters and the separation of nucleotides on pellicular ion exchangers.

Anal Chem. **39**, 1422–1428

Horwitz R.J., Li J and Greenblatt J. (1987)

An elongation control particle containing the N gene transcriptional antitermination protein of bacteriophage lambda.

Cell. **51**, 631-641

Hung S.C and Gottesman M.E. (1995)

Phage HK022 Nun protein arrests transcription on phage λ DNA in vitro and competes with the phage λ N antitermination protein.

J Mol Biol. **247**, 428–442

Huth J.R., Bewley C.A., Jackson B.M., Hinnebusch A.G., Clore G.M and Gronenborn A.M. (1997)

Design of an expression system for detecting folded protein domains and mapping macromolecular interactions by NMR.

Protein Sci. **6**, 2359-2364.

Igarashi K and Ishihama A. (1991)

Bipartite functional map of the *E. coli* RNA polymerase alpha subunit involvement of the C-terminal region in transcription activation by cAMP-CRP.

Cell. **65**, 1015-1022

Ikura M., Bax A., Marius C.G and Angela M.G. (1990)

Detection of Nuclear Overhauser effects between degenerate amide proton resonances by heteronuclear three-dimensional nuclear magnetic resonance spectroscopy.

J Am Chem Soc. **112**, 9020-9022

Ikura M., Kay L.E and Bax A. (1990)

A novel approach for sequential assignment of ^1H , ^{13}C , and ^{15}N spectra of proteins: heteronuclear triple resonance three-dimensional NMR spectroscopy. Application to calmodulin.

Biochemistry. **29**, 4659-4667

Ingle J.D.G and Crouch S.R. (1988)

Spectrochemical Analysis, Prentice Hall, New Jersey.

Ishihama A. (1988)

Promoter selectivity of prokaryotic RNA polymerases.

Trends Genet. **4**, 282-286

Ishihama A. (1992)

Role of the RNA polymerase alpha subunit in transcription activation.

Mol Microbiol. **6**, 3283-3288

Jason R.C and John B.M (2006)

Bacteriophages and biotechnology: Vaccines, gene therapy and antibacterials.

Trends Biotechnol. **24**, 212-218

Jeon Y.H., Negishi T., Shirakawa M., Yamazaki T., Fujita N., Ishihama A and Kyogoku Y. (1995)

Solution structure of the activator contact domain of the RNA polymerase alpha subunit.
Science. **270**, 1495-1497

Jeremy N.S.E. (1995)

Biomolecular NMR spectroscopy. Oxford University Press.

Johnson B.A & Blevins R.A. (1994)

NMRView: A computer program for the visualization and analysis of NMR data.
J Biomol NMR. **4**, 603-614

Kaiser A.D. (1957)

Mutations in a temperate bacteriophage affecting its ability to lysogenize *Escherichia coli*.
Virology. **3**, 42-61

Kay L.E., Keifer P and Saarinen T. (1992)

Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity.
J Am Chem Soc. **114**, 10663-10665

Kim H.C., Robert S.W and Gottesman M.E (2006)

Role of *e. coli* NusA in phage HK022 Nun-mediated transcription termination.
J Mol Biol. **359**, 10-21

Kojima C and Kainosho M. (2000)

Simple suppression of spurious peaks in TROSY experiments.
J Magn Reson. **143**, 417-422

Kramer M.F and Ceon D.M. (2001)

Enzymatic amplification of DNA by PCR: Standard procedures and optimization.
Curr Protoc Mol Biol. Chapter **15**, Unit 15.1

Laemmli U.K. (1970)

Cleavage of structural proteins during the assembly of the head of bacteriophage T4.
Nature. **227**, 680-685

Lederberg E.M. (1951)

Lysogenicity in *E. coli* K-12.
Genetics. **36**, 560

Lederberg E.M and Lederberg J. (1953)

Genetic studies of lysogenicity on *Escherichia coli*.
Genetics. **38**, 51-64

Li J., Mason S.W and Greenblatt J. (1993)

Elongation factor NusG interacts with termination factor λ to regulate termination and

antitermination of transcription.

Genes Dev. **7**, 161–172

Liu K and Hanna M.M. (1995)

NusA contacts nascent RNA in *Escherichia coli* transcription complexes.

J Mol Biol. **247**, 547-558

Live D.H., Davis D.G., Agosta W.C and Cowburn D. (1984)

Long-range hydrogen bond mediated effects in peptides 15N NMR-study of gramicidin-S in water and organic solvents.

J Am Chem Soc. **106**, 1939-1941

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

Marion D., Ikura M., Tschudin R and Bax A. (1989)

Rapid recording of 2D NMR spectra without phase cycling. Applications to the study of hydrogen exchange in proteins.

J Magn Reson. **85**, 393-399

Marley J., Lu M and Bracken C. (2001)

A method for efficient isotopic labeling of recombinant proteins.

J Biomol NMR. **20**, 71–75

Meyer O and Schlegel H.G. (1983)

Biology of aerobic carbon monoxide-oxidizing bacteria.

Annu Rev Microbiol. **37**, 227–310

Mogridge J., Legault P., Li J., van Oene M.D., Kay L.E and Greenblatt J. (1998)

Independent ligand-induced folding of the RNA-binding domain and two functionally distinct antitermination regions in the phage λ 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 λ N protein.

Genes Dev. **9**, 2831-2845

Mori S., Abeygunawardana C., Johnson M.O and Vanzijl P.C.M. (1995)

Improved sensitivity of HSQC spectra of exchanging protons at short interscan delays using a new fast HSQC (fHSQC) detection scheme that avoids water saturation.

J Magn Reson. B. **108**, 94-98

Morton C.J., Pugh D.J., Brown E.L., Kahmann J.D., Renzoni D.A and Campbell ID. (1996)

Solution structure and peptide binding of the SH3 domain from human Fyn.
Structure. **4**, 705-714

Müller Luciano. (1979)

Sensitivity enhanced detection of weak nuclei using heteronuclear multiple quantum coherence.

J Am Chem Soc. **101**, 4481-4484

Muskett F.W., Frenkiel T.A., Feeney J., Freedman R.B., Carr M.D and Williamson RA. (1998)

High resolution structure of the N-terminal domain of tissue inhibitor of metalloproteinases-2 and characterization of its interaction site with matrix metalloproteinase-3.

J Biol Chem. **273**, 21736-21743

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

Nudler E. (1999)

Transcription elongation: Structural basis and mechanisms.

J Mol Biol. **288**, 1-12

Nudler E and Gottesman M.E. (2002)

Transcription termination and anti-termination in *E. coli*.

Genes Cells. **7**, 755-768

Nudler E and Gusarov I. (2001)

Control of intrinsic transcription termination by N and NusA: The basic mechanism.

Cell. **107**, 437-449

Oberto J., Weisberg R.A and Gottesman M.E. (1989)

Structure and function of the *nun* gene and the immunity region of the lambdaoid phage HK022.

J Mol Biol. **207**, 675-693

Oppenheim A.B., Kobiler O., Stavans J., Court D.L and Adhya S. (2005)

Switches in bacteriophage lambda development.

Annu Rev Genet. **39**, 409-429

Pasman Z and Peter H. Von Hippel. (2000)

Regulation of Rho-dependent transcription termination by NusG is specific to the *Escherichia coli* elongation complex.

Biochemistry. **39**, 5573-5585

Patterson T.A., Zhang Z., Baker T., Johnson L.L., Friedman D.I and Court D.L. (1994)

Bacteriophage lambda N-dependent transcription antitermination. Competition for an RNA

site may regulate antitermination.

J Mol Biol. **236**, 217-228

Pellacchia M., Montgomery D.L., Stevens S.Y., Vander Kooi C.W., Feng H.P., Gierasch L.M and Zuiderweg E.R. (2000)

Structural insights in to substrate binding by the molecular chaperone DnaK.

Nat Struct Biol. **7**, 298-303

Pervushin K. (2000)

Impact of transverse relaxation optimized spectroscopy (TROSY) on NMR as a technique in structural biology.

Q Rev Biophys. **33**, 161-197

Pervushin K., Riek R., Wider G and Wüthrich K. (1997)

Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution.

Proc Natl Acad Sci U S A. **94**, 12366-12371

Piotto M., Saudek V and Sklenar V. (1992)

Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions.

J Biomol NMR. **2**, 661-665.

Prasch S., Schwarz S., Eisenmann A., Wöhrl B.M., Schweimer K and Rösch P. (2006)

Interaction of the intrinsically unstructured phage λ N protein with *Escherichia coli* NusA.

Biochemistry. **45**, 4542-4549

Prasch S., Jurk M.A.W., Washburn R.S., Gottesman M.E., Wöhrl B.M and Rösch P. (2008)

E. coli NusA recognizes phage λ *nut* spacer RNA sequences.

In revision.

Press W.H., Teukolsky S.A., Vetterling W.T and Flannery B.P. (1992)

Numerical recipes in C 2nd ed. New York: Cambridge University Press

Ptashne M and Gann A. (2002)

Genes and signals. Cold spring harbor, New York, Cold Spring Harbor Laboratory Press

Ptashne M. (1992)

A genetic switch: Phage lambda and higher organisms.

Cambridge, MA: Blackwell Sci. 2nd edition.

Rajagopal P., Waygood E.B., Reizer J., Saier M.H and Klevit R.E. (1997)

Demonstration of protein-protein interaction specificity by NMR chemical shift mapping.

Protein Sci. **6**, 2624-2627

Rees W.A., Weitzel S.E., Yager T.D., Das A and von Hippel P.H. (1996)

Bacteriophage lambda N protein alone can induce transcription antitermination *in vitro*.

Proc Natl Acad Sci U S A. **93**, 342-346

Rhodium V.A and Stephen Busby J.W. (1998)

Positive activation of gene expression.

Curr Opin Microbiol. **1**, 152-159

Riek R., Pervushin K and Wüthrich K. (2000)

TROSY and CRINEPT: NMR with large molecular and supramolecular structures in solution.

Trends Biochem Sci. **25**, 462-468

Robert C.T., Hassan K.S., Shanteri S., David J.A., Craig A.B., John L.M and Brian G.F. (2005)

Auto-induction medium for the production of [U-¹⁵N]- and [U-¹³C, U-¹⁵N]-labelled proteins for NMR screening and structure determination.

Protein Expr Purif. **40**, 268-278

Robledo R.A and Gottesman M.E. (1991)

Escherichia coli mutations that block transcription termination by phage HK022 Nun protein.

J Mol Biol. **220**, 613-619

Roger W Hendrix. (2003)

Bacteriophage genomics.

Curr Opin Microbiol. **6**, 506-511

Rosenberg M., Court D., Shimatake H., Brady C and Wulff D.L. (1978)

The relationship between function and DNA sequence in an intercistronic regulatory region in phage lambda.

Nature. **272**, 414-423

Sambrook J., Fritsch E.F and Maniatis M. (1989)

Molecular cloning: a laboratory manual.

Sattler M., Schleucher J and Griesinger C. (1999)

Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients.

Progress in Nuclear Magnetic Resonance Spectroscopy. **34**, 93-158

Schagger H and Von Jagow G. (1987)

Tricine-sodium dodecyl-sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa.

Anal Biochem. **166**, 368-379

Scharpf M., Sticht H., Schweimer K., Boehm M., Hoffmann S and Rosch P. (2000)

Antitermination in bacteriophage λ . The structure of the N36 peptide-boxB RNA complex.

Eur J Biochem. **267**, 2397-2408

Schleucher J., Sattler M and Griesinger C. (1993)

Coherence selection by gradients without signal attenuation: application to the three-

dimensional HNCOC experiment.

Angew Chem. **105**, 1518-1521

Schleucher J., Schwendinger M.G., Sattler M., Schmidt P., Schedletzky O., Glaser S.J., Sorensen O.W and Griesinger C. (1994)

A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients.

J Biomol NMR. **4**, 301-306

Schweimer K. (2000)

Mehrdimensionale NMR Spektroskopie zur Bestimmung der Strukturen des Birken pollenallergens Bet v1, des Guillardia theta Rubredoxins und des [2Fe-2S] Ferredoxins aus Halobakterium salinarium.

Doctoral Dissertation, University of Bayreuth, Bayreuth, Germany.

Sekhar T and Gerhard W. (1996)

An optimized 3D NOESY-HSQC.

J Magn Reson. B. **112**, 200-205

Shuker S.B., Hajduk P.J., Meadows R.P and Fesik S.W. (1996)

Discovering high-affinity ligands for proteins: SAR by NMR.

Science. **274**, 1531-1534

Sklenar V., Piotto M., Leppik R and Saudek V. (1993)

Gradient-tailored water suppression for 1H-15N HSQC experiments optimized to retain full sensitivity.

J Magn Reson. **102**, 241-245

States D.J., Haberkorn R.A and Ruben D.J. (1982)

A Two-Dimensional nuclear overhauser experiment with pure absorption phase in four quadrants.

J Magn Reson. **48**, 286-292

Steven S.Y., Sanker S., Kent C and Zuiderweg E.R. (2001)

Delineation of the allosteric mechanism of a cytidylyl transferase exhibiting negative cooperativity.

Nat Struct Biol. **8**, 947-952

Studier F.W. (2004)

Auto-induction for protein production in inducible T7 expression systems.

Annual Meeting of the American Crystallographic Association, Chicago, IL, July 17–22

Studier F.W. (2005)

Protein production by auto-induction in high density shaking cultures.

Protein Expr Purif. **41**, 207-234

Sugar I.P and Neumann E (1984)

Stochastic model for electric field-induced membrane pores- Electroporation.

Biophys Chem. **19**, 211–25

Sullivan S.L and Gottesman M.E. (1992)

Requirement for *E. coli* NusG protein in factor-dependent transcription termination.
Cell. **68**, 989–994

Szalewska-Palasz A., Strzelczyk B., Herman-Antosiewicz A., Wegrzyn G and Thomas M.S (2003)

Genetic analysis of bacteriophage lambda-N dependent antitermination suggests a possible role for the RNA polymerase alpha subunit in facilitating specific functions of NusA and NusE.

Arch Microbiol. **180**, 161-168

Thomas M.R. (1994)

Simple, effective clean-up of DNA ligation reaction prior to electro-transformation of *E. coli*.
Biotechniques. **16**, 988–990.

Uptain S.M., Kane C.M and Chamberlin M.J. (1997)

Basic mechanisms of transcript elongation and its regulation.

Annu Rev Biochem. **66**, 117-172

Van Nuland N.A., Kroon G.J., Dijkstra K., Wolters G.K., Scheek R.M and Robillard G.T. (1993)

The NMR determination of the IIA (mt1) binding site on Hpr of the *Escherichia coli* phosphoenol pyruvate-dependent phosphotransferase system.

FEBS Lett. **315**, 11-15

Vassilyev D.G., Vassilyeva M.N., Perederina A., Tahirov T.H and Artsimovitch I. (2007)

Structural basis for transcription elongation by bacterial RNA polymerase.

Nature. **448**, 157-165

Voet Donald, Voet Judith and Charlottes w-Pratt.

Fundamentals of Biochemistry. 2nd edition

von Hippel P.H. (1998)

An integrated model of the transcription complex in elongation, termination and editing.

Science. **281**, 660-665

von Hippel P.H., Rees W.A., Rippe K and Wilson K.S. (1996)

Specificity mechanisms in the control of transcription.

Biophys Chem. **59**, 231-246

Von Hippel PH., Pasman Z. (2002)

Reaction pathways in transcript elongation.

Biophys Chem. **101-102**, 401-423

Vuister G.W and Bax A. (1992)

Resolution enhancement and spectral editing of uniformly ¹³C enriched proteins by homonuclear broadband ¹³C decoupling.

J Magn Reson. **98**, 428-435

- Washburn R.S., Court D.L and Gottesman M.E. (2006)**
Role of Rnase III binding site in transcription termination an lambda *nutL* by HK022 Nun protein.
J Bacteriol. **188**, 6824-6831
- Washburn R.S., Wang Y and Gottesman M.E. (2003)**
Role of *E. coli* transcription-repair coupling factor Mfd in Nun-mediated transcription termination.
J Mol Biol. **329**, 655–662
- Watnick R.S and Gottesman M.E. (1998)**
Escherichia coli NusA is required for efficient RNA binding by phage HK022 nun protein.
Proc Natl Acad Sci U S A. **95**, 1546–1551
- Watnick R.S., Herring S.C., Palmer A.G and Gottesman M.E. (2000)**
The carboxyl terminus of phage HK022 Nun includes a novel zinc-binding motif and a tryptophan required for transcription termination.
Genes Dev. **14**, 731–739
- Weisberg R.A and Gottesman M.E. (1999)**
Processive antitermination.
J Bacteriol. **181**, 359-367
- Weisberg R.A., Gottesman M.E., Hendrix R.W and Little JW. (1999)**
Family values in the age of genomics:Comparative analyses of temperate bacteriophage HK022.
Annu Rev Genet. **33**, 565-602
- Whalen W., Ghosh B and Das A. (1988)**
NusA protein is necessary and sufficient *in vitro* for phage λ N gene product to suppress a rho-independent terminator placed downstream of *nutL*.
Proc Natl Acad Sci U S A. **85**, 2494-2498
- Wider G and Wüthrich K. (1999)**
NMR spectroscopy of large molecules and multimolecular assemblies in solution.
Curr Opin Struct Biol. **9**, 594-601
- William R. C and Robert M. K.C. (1994)**
Two-Dimensional NMR Spectroscopy. Applications for chemists and Biochemists. Second Edition
- Williamson R.A., Carr M.D., Frenkiel T.A., Feeney J and Freedman R.B. (1997)**
Mapping the binding site for matrix metalloproteinase on the N-terminal domain of the tissue inhibitor of metalloproteinases-2 by NMR chemical shift perturbation.
Biochemistry. **36**, 13882-13889
- Wilson C.M. (1983)**
Staining of proteins on gels: Comparisons of dyes and procedures.
Methods Enzymol. **91**, 236-246

Wisdom G.B. (1997)

Molecular weight determinations using polyacrylamide gel electrophoresis with tris-tricine buffers.

Methods Mol Biol. **73**, 97 -100

Wittekind M and Mueller L. (1993)

HNCACB, a High-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha and beta carbon resonances in proteins.

J Magn Reson. **101**, 201-205

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

Wüthrich K., Salzmann M., Pervushin K., Wider G., Senn H. (1998)

TROSY in triple-resonance experiments: New perspectives for sequential NMR assignment of large proteins.

Proc Natl Acad Sci U S A. **95**, 13585-13590

Zheng C and Friedman D.I. (1994)

Reduced Rho-dependent transcription termination permits NusA-independent growth of *Escherichia coli*.

Proc Natl Acad Sci U S A. **91**, 7543-7547

Zhou P., Lugovskoy A.A and Wagner G. (2001)

A solubility-enhancement tag (SET) for NMR studies of poorly behaving proteins.

J Biomol NMR. **20**, 11-4

Zhou Y., Mah T.F., Greenblatt J and Friedman D.I. (2002)

Evidence that the KH RNA-binding domains influence the action of the *E. coli* NusA protein.

J Mol Biol. **318**, 1175-1188

Zuiderweg E.R., Hamers L.F., Rollema H.S., de Bruin S.H and Hilbers C.W. (1981)

³¹P NMR study of the kinetics of binding of myo-inositol hexakisphosphate to human hemoglobin-observation of fast-exchange kinetics in high affinity systems.

Eur J Biochem. **118**, 95-104.

Zuiderweg E.R. (2002)

Mapping protein-protein interactions in solution by NMR spectroscopy.

Biochemistry. **41**, 1-7

9 Appendix

9.1 Script for K_D fitting

```

clear;
global Data;
fqz_1H = 700.13;
fqz_15N = 70.5;
%fqz_1H = 1.0;
%fqz_15N = 1.0;
filename_SHIFT = sprintf('./text.data')
fp = fopen(filename_SHIFT)
shiftdata = fscanf(fp,'%g %g',[2 inf])
fclose(fp)
filename_CONC = sprintf('./ratio.dat')
fp = fopen(filename_CONC)
conc_data = fscanf(fp,'%g %g',[2 inf])
fclose(fp)
shiftdata = shiftdata'
conc_data = conc_data'
ratio = conc_data(:,1)
conc = conc_data(:,2)
conc = conc .* 1000.0
proton = shiftdata(:,1)
proton = proton - proton(1);
proton = proton * fqz_1H;
nitrogen = shiftdata(:,2)
nitrogen = nitrogen(1) - nitrogen
nitrogen = nitrogen * fqz_15N;
norm_shift = sqrt(proton .* proton + nitrogen .* nitrogen)
% plot(ratio, norm_shift, 'ro')
Data = zeros(length(conc), 3);
Data(:,1) = conc
Data(:,2) = ratio
Data(:,3) = norm_shift
a = Data(:,2)
b = Data(:,3)
plot(a, b, 'ro')
hold on
global Plothandle
Plothandle = plot(a, b, 'EraseMode','xor');
k = [10 1];
trace = 0; % If this is nonzero, intermediate steps in the solution are displayed.
tol = .00001; % This is the termination tolerance.
k = fminsearch('kd_fit_func',k,[trace tol]);
k

```

9.2 Nucleotide sequence of HK022 Nun

atgctgatgggtgaaaaaaaaaccatttatgtgaaccgcatagcggccagaaccgcaaagtg
M L M V K K T I Y V N P D S G Q N R K V
agcgatcgcggcctgaccagccgcatcgccgcccgcattgcgcgctgggaaaaacgcatt
S D R G L T S R D R R R I A R W E K R I
gcgatcgcgctgaaaaacggcgtgacccccgggctttaacgcgattgatgatggcccggaa
A Y A L K N G V T P G F N A I D D G P E
tataaaattaacgaagatccgatggataaagtggataaagcgcctggcgacccccgtttccg
Y K I N E D P M D K V D K A L A T P F P
cgcgatgtggaaaaaattgaagatgaaaaatatgaagatgtgatgcatcgcgctgggtgaac
R D V E K I E D E K Y E D V M H R V V N
catgcgcatcagcgcaaccgcaacaaaaaatggagc
H A H Q R N P N K K W S

Physical/Chemical parameters of HK022 Nun

Number of amino acids	112
Molecular weight	13107.8
Theoretical pI	9.52
Extinction coefficient	16960
Instability index	41.05

9.3 Nucleotide sequence of Nun CTD

ggcgcgatgggctgacccccgggctttaacgcgattgatgatggcccggaaatataaaatt
G A M G V T P G F N A I D D G P E Y K I
aacgaagatccgatggataaagtggataaagcgcctggcgacccccgtttccgcgcatgtg
N E D P M D K V D K A L A T P F P R D V
gaaaaaattgaagatgaaaaatatgaagatgtgatgcatcgcgctgggtgaaccatgcgcat
E K I E D E K Y E D V M H R V V N H A H
cagcgcaaccgcaacaaaaaatggagc
Q R N P N K K W S

Physical/Chemical parameters of Nun CTD

Number of amino acids	69
Molecular weight	7892.7
Theoretical pI	5.21
Extinction coefficient	8480
Instability index	30.50

9.4 Nucleotide sequence of NusA (1-495)

atgaacaaagaaattctggcgggtggtggaagcgggtgagcaacgaaaaagcgctgccgcgc
M N K E I L A V V E A V S N E K A L P R
 gaaaaaatttttgaagcgctggaagcgcgctggcgaccgcgaccaaaaaaaaaatatgaa
E K I F E A L E S A L A T A T K K K Y E
 caggaaattgatgtgcgcgctgcagattgatcgcaaaagcggcgattttgatacctttcgc
Q E I D V R V Q I D R K S G D F D T F R
 cgctggctggtggtggaatgaagtgaaccagccgaccaaaagaattaccctggaagcggcg
R W L V V D E V T Q P T K E I T L E A A
 cgctatgaagatgaaagcctgaacctgggcgattatgtggaagatcagattgaaagcgtg
R Y E D E S L N L G D Y V E D Q I E S V
 acctttgatcgcattaccaccagaccgcaaacaggtgattgtgcagaaagtgcgcgaa
T F D R I T T Q T A K Q V I V Q K V R E
 gcggaacgcgcgatggtggtggaatcagtttcgcaaacatgaaggcgaaattattaccggc
A E R A M V V D Q F R E H E G E I I T G
 gtggtgaaaaaagtgaaccgcaataacattagcctggatctgggcaacaacgcggaagcg
V V K K V N R D N I S L D L G N N A E A
 gtgattctgcgcgaagatatgctgccgcgcgaaaactttcgcccgggcgatcgcgtgcgc
V I L R E D M L P R E N F R P G D R V R
 ggctgctgtatagcgtgcgcccgaagcgcgcggcgcgcagctgtttgtgaccgcgagc
G V L Y S V R P E A R G A Q L F V T R S
 aaaccggaaatgctgattgaactgtttcgcattgaagtgccggaaattggcggaagaagtg
K P E M L I E L F R I E V P E I G E E V
 attgaaattaaagcggcggcgcgcgatccgggagccgcgcgaaaattgacggtgaaaacc
I E I K A A A R D P G S R A K I A V K T
 aacgataaacgcattgatccggtgggcgctgctggtggcatgcgcggcgcgcgctgcag
N D K R I D P V G A C V G M R G A R V Q
 gcggtgagcaccgaactgggcggcgaacgcattgatattgtgctgtgggatgataaccgc
A V S T E L G G E R I D I V L W D D N P
 gcgcagtttgtgattaacgcgatggcgccggcggatgtggcgagcattgtggtggatgaa
A Q F V I N A M A P A D V A S I V V D E
 gataaacataccatggatattgcggtggaagcgggcaacctggcgagggcattggccgc
D K H T M D I A V E A G N L A Q A I G R
 aacggccagaacgtgcgcctggcgagccagctgagcggctgggaactgaacgtgatgacc
N G Q N V R L A S Q L S G W E L N V M T
 gtggatgatctgcaggcgaaacatcaggcgggaagcgcgatgcggcgattgatacctttacc
V D D L Q A K H Q A E A H A A I D T F T
 aaatatctggatattgatgaagattttgcgaccgtgctggtggaagaaggcttttagcacc
K Y L D I D E D F A T V L V E E G F S T
 ctggaagaactggcgtatgtgccgatgaaagaactgctggaattgaaggcctggatgaa
L E E L A Y V P M K E L L E I E G L D E
 ccgaccgtggaagcgcctgcgcgaacgcgcgaaaaacgcgctggcgaccattgcgagggcg
P T V E A L R E R A K N A L A T I A Q A
 caggaagaagcctgggcgataacaaaccggcggatgatctgctgaacctggaaggcgtg
Q E E S L G D N K P A D D L L N L E G V
 gatcgcgatctggcgtttaactggcggcgcgcggcgtgtgcaccctggaagatctggcg
D R D L A F K L A A R G V C T L E D L A

gaacagggcattgatgatctggcggatattgaaggcctgaccgatgaaaaagcgggcgcg
E Q G I D D L A D I E G L T D E K A G A
 ctgattatggcggcgcgcaacatttgcctgggttggcgatgaagcg
L I M A A R N I C W F G D E A

Physical/Chemical parameters of NusA (1-495)

Number of amino acids	495
Molecular weight	54870.9
Theoretical pI	4.53
Extinction coefficient	31065
Instability index	35.15

9.5 Nucleotide sequence of NusA ar1

atgaccgtggatgatctgcaggcgaaacatcaggcgggaagcgcgatgcggcgattgatacc
M T V D D L Q A K H Q A E A H A A I D T
 ttaccaaatatctggatattgatgaagattttgcgaccgtgctggtggaagaaggcttt
F T K Y L D I D E D F A T V L V E E G F
 agcaccctggaagaactggcgtatgtgccgatgaaagaactgctggaaattgaaggcctg
S T L E E L A Y V P M K E L L E I E G L
 gatgaaccgaccgtggaagcgcctgcgcggaacgcgcgaaaaacgcgctggcgaccattgcg
D E P T V E A L R E R A K N A L A T I A
 caggcgcaggaagaaagcctgggc
Q A Q E E S L G

Physical/Chemical parameters of NusA ar1

Number of amino acids	88
Molecular weight	9751.8
Theoretical pI	4.11
Extinction coefficient	2980
Instability index	40.00

9.6 Nucleotide sequence of NusA ar2

agcctgggcgataacaaaccggcggatgatctgctgaacctggaaggcgtggatcgcgat
S L G D N K P A D D L L N L E G V D R D
 ctggcggtttaaactggcggcgcgcgggcgtgtgcacctggaagatctggcggaacagggc
L A F K L A A R G V C T L E D L A E Q G
 attgatgatctggcggatattgaaggcctgaccgatgaaaaagcgggcgcgattatggcg
I D D L A D I E G L T D E K A G A I M A
 gcgcgcaacatttgctggtttggcgatgaagcg
A R N I C W F G D E A

Physical/Chemical parameters of NusA ar2

Number of amino acids	72
Molecular weight	7663.5
Theoretical pI	3.95
Extinction coefficient	5625
Instability index	31.28

9.7 Nucleotide sequence of SKK domain

atgcataaccatcatcatcatgaaggcgaaattattaccggcgtggtgaaaaagtgaac
M H N H H H H E G E I I T G V V K K V N
 cgcgataacattagcctggatctgggcaacaacgcggaagcggtgattctgcgcgaagat
R D N I S L D L G N N A E A V I L R E D
 atgctgccgcgcgaaaactttcgccccggcgatcgcgctgcgcggcgtgctgtatagcgtg
M L P R E N F R P G D R V R G V L Y S V
 cgcccggaagcgcgcggcgcgcagctgtttgtagcccgagcaaaccggaaatgctgatt
R P E A R G A Q L F V T R S K P E M L I
 gaactgtttcgcattgaagtgcggaaattggcgaagaagtgattgaaattaaagcggcg
E L F R I E V P E I G E E V I E I K A A
 gcgcgcgatccgggcagccgcgcgaaaattgcggtgaaaaccaacgataaacgcattgat
A R D P G S R A K I A V K T N D K R I D
 ccgggtgggcgcgctgctggtggcatgcgcggcgcgcgctgcaggcggtgagcaccgaactg
P V G A C V G M R G A R V Q A V S T E L
 ggcggcgaacgcattgatattgtgctgtgggatgataaccggcgcagtttgtgattaac
G G E R I D I V L W D D N P A Q F V I N
 gcgatggcgccggcggatgtggcgagcattgtggtggatgaagataaacataccatggat
A M A P A D V A S I V V D E D K H T M D
 attgcggtggaagcgggcaacctggcgcaggcgattggccgcaacggccagaacgtgcgc
I A V E A G N L A Q A I G R N G Q N V R
 ctggcgagccagctgagcggctgggaactgaacgtgatgaccgtggatgatctgcaggcg
L A S Q L S G W E L N V M T V D D L Q A
 aaacat
K H

Physical/Chemical parameters of SKK domain

Number of amino acids	222
Molecular weight	24416.7
Theoretical pI	5.54
Extinction coefficient	12490
Instability index	36.19

9.8 Nucleotide sequence of NusG

atgagcgaagcgccgaaaaaacgctggtatgtggtgcaggcgtttagcggctttgaaggc
M S E A P K K R W Y V V Q A F S G F E G
 cgcggtggcgaccagcctgcgcgaacatattaaactgcataacatggaagatctgtttggc
R V A T S L R E H I K L H N M E D L F G
 gaagtgatggtgccgaccgaagaagtggtggaaattcgcgggcgccagcgccgcaaaagc
E V M V P T E E V V E I R G G Q R R K S
 gaacgcaaattttttccgggctatgtgctggtgcagatggtgatgaacgatgcgagctgg
E R K F F P G Y V L V Q M V M N D A S W
 catctggtgcgcgagcgtgccgcgctgatgggctttattggcggcaccagcgatcgcccg
H L V R S V P R V M G F I G G T S D R P
 gcgccgattagcgataaagaagtggatgcgattatgaaccgcctgcagcaggtgggcgat
A P I S D K E V D A I M N R L Q Q V G D
 aaaccgccccgaaaacctggttgaaccgggcgaaatggtgcgcgatgaacgatggcccg
K P R P K T L F E P G E M V R V N D G P
 tttgcggtatttaacggcgtggtggaagaagtggattatgaaaaagccgcctgaaagtg
F A D F N G V V E E V D Y E K S R L K V
 agcgtgagcatttttgccgcgcgacccccgggtggaactggatttttagccaggtggaaaaa
S V S I F G R A T P V E L D F S Q V E K
 gcg
A

Physical/Chemical parameters of NusG

Number of amino acids	181
Molecular weight	20531.5
Theoretical pI	6.34
Extinction coefficient	15470
Instability index	43.37

9.9 Nucleotide sequence of NusB

ggcgcgatggaaccggcgggcgcgcccgccgcgcgcgcggaatgcgcggtgcagggcgctgtat
G A M E P A A R R R A R E C A V Q A L Y
 agctggcagctgagccagaacgatattgcggatgtggaatatcagtttctggcggaacag
S W Q L S Q N D I A D V E Y Q F L A E Q
 gatgtgaaagatgtggatgtgctgtatcttctcggaactgctggcgggcgctggcgaccaac
D V K D V D V L Y F R E L L A G V A T N
 accgcgtatctggatggcctgatgaaaccgtatctgagccgctgctggaagaactgggc
T A Y L D G L M K P Y L S R L L E E L G
 caggtggaaaaagcgggtgctgcgctgctgctgactgagcaaacgcagcgatgtg
Q V E K A V L R I A L Y E L S K R S D V
 ccgtataaagtggcgattaacgaagcgattgactggcgaaaagcttggcgcggaagat
P Y K V A I N E A I E L A K S F G A E D
 agccataaatttgtgaacggcgctgctggataaagcggcgccggtgattcgcccgaacaaa
S H K F V N G V L D K A A P V I R P N K
 aaa
K

Physical/Chemical parameters of NusB

Number of amino acids	141
Molecular weight	15818.1
Theoretical pI	5.51
Extinction coefficient	15930
Instability index	43.99

9.10 Backbone resonance assignment for SKK domain

AA	HN	N	CA	CA(i-1)	CB	CB(i-1)	CO	CO(i-1)
* M	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
* H	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
* N	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
* H	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
* H	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
* H	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
133 H	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
134 E	8.515	121.037	57.800	54.254	29.000	30.550	177.640	174.043
135 G	9.373	112.948	44.700	57.860	-100.000	29.026	173.360	177.635
136 E	7.615	118.152	54.671	44.769	30.943	-100.000	176.389	173.360
137 I	8.733	121.442	60.983	54.933	36.899	30.964	175.806	176.344
138 I	9.108	124.526	57.821	60.646	41.100	37.020	174.658	175.816
139 T	8.065	115.440	60.869	57.772	69.539	41.214	173.824	174.653
140 G	8.843	111.420	43.639	61.020	-100.000	69.653	170.732	173.871
141 V	8.027	118.498	59.700	43.400	33.500	-100.000	177.000	170.700
142 V	9.247	127.266	64.346	59.643	30.435	33.345	176.379	177.028
143 K	9.317	130.657	56.556	64.238	32.975	30.500	175.444	176.361
144 K	7.420	117.473	55.180	56.428	35.514	33.089	173.535	175.429
145 V	8.844	124.441	61.894	55.247	31.793	35.383	174.767	173.582
146 N	8.319	124.859	51.400	61.771	40.100	32.073	174.578	174.837
147 R	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
148 D	8.449	124.072	53.600	60.500	41.000	38.335	175.856	175.700
149 N	7.475	113.927	53.000	53.607	40.512	41.400	172.270	175.939
150 I	8.487	115.673	59.529	53.050	40.137	40.087	174.635	172.275
151 S	8.812	118.244	56.281	59.433	64.968	40.199	174.147	174.671
152 L	9.307	123.370	52.408	56.133	45.000	65.083	175.259	174.221
153 D	9.521	121.121	52.706	52.290	41.608	45.000	177.051	175.264
154 L	8.574	126.734	54.152	52.793	41.100	41.722	177.542	177.313
155 G	8.610	109.680	43.654	53.897	-100.000	41.112	174.100	177.641
156 N	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
157 N	9.019	110.873	54.257	53.978	37.100	37.300	173.827	174.904
158 A	7.723	122.695	51.299	54.418	19.770	37.037	176.272	173.933
159 E	8.463	122.504	54.192	51.151	32.975	19.885	174.576	176.285
160 A	8.533	122.854	49.319	54.091	24.014	32.979	175.047	174.622

198 T	9.358	118.076	59.911	59.496	69.539	33.551	172.016	175.302
199 R	9.167	122.761	55.928	59.940	31.550	-100.000	-100.000	172.034
200 S	8.345	116.431	59.172	55.954	-100.000	-100.000	174.800	177.070
201 K	7.123	123.127	55.950	59.042	29.928	64.067	177.099	174.798
202 P	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
203 E	10.425	116.673	60.422	66.171	28.550	23.400	-100.000	176.839
204 M	7.200	115.708	57.643	60.267	31.451	28.400	175.861	178.123
205 L	6.518	117.300	57.627	57.627	42.623	31.221	177.300	175.961
206 I	7.734	116.851	65.514	57.508	37.545	42.588	177.651	177.300
207 E	8.254	115.591	57.858	65.445	27.896	37.640	179.905	177.693
208 L	7.983	119.020	57.313	57.858	41.608	28.011	180.029	179.937
209 F	7.958	118.158	63.952	57.200	38.053	41.722	177.379	180.123
210 R	8.178	118.246	59.343	63.932	28.912	37.994	177.000	177.405
211 I	7.230	113.416	63.045	59.327	38.053	29.026	177.862	177.000
212 E	7.618	118.450	56.698	63.162	30.300	38.085	176.380	177.900
213 V	8.086	117.866	58.000	56.700	30.900	30.360	179.976	176.396
214 P	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
215 E	9.966	115.387	60.768	64.534	29.089	30.894	179.522	178.401
216 I	8.020	119.800	63.800	60.650	36.300	29.000	180.700	179.500
217 G	8.158	112.201	47.100	63.777	-100.000	36.356	175.851	180.701
218 E	7.944	116.366	56.001	47.100	29.600	-100.000	175.798	175.890
219 E	7.960	112.515	57.611	55.980	25.357	29.534	175.470	175.796
220 V	8.086	119.560	64.200	57.580	32.104	25.500	175.600	175.500
221 I	7.170	115.056	58.296	64.377	39.576	32.253	173.703	175.670
222 E	8.746	122.270	53.288	58.200	32.675	39.700	175.566	173.745
223 I	8.953	123.116	60.694	53.540	36.529	33.089	176.339	175.537
224 K	8.653	126.924	54.194	60.609	31.959	36.644	176.212	176.330
225 A	7.523	117.319	51.761	54.106	21.802	32.073	174.577	176.099
226 A	8.713	121.133	51.400	51.737	21.935	22.148	174.589	174.595
227 A	8.949	122.192	51.456	51.406	22.518	21.917	175.399	174.638
228 R	9.114	117.631	55.719	51.263	35.006	22.932	174.475	175.373
229 D	9.393	123.855	51.188	55.667	39.576	35.120	174.459	174.543
230 P	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
231 G	10.651	114.734	44.600	-100.000	-100.000	-100.000	-100.000	177.489
232 S	8.438	112.515	58.600	44.769	63.300	-100.000	174.890	173.300
233 R	8.628	122.907	54.980	58.600	34.498	63.500	171.208	174.882
234 A	9.358	125.779	50.286	54.838	24.341	34.105	176.299	171.242

235 K	9.050	115.709	54.618	50.179	38.053	24.456	174.504	176.323
236 I	8.632	116.958	58.739	54.525	41.100	38.167	172.250	174.559
237 A	8.787	126.082	48.656	58.648	22.518	41.214	177.170	172.288
238 V	8.928	114.204	58.422	48.523	35.514	22.524	173.837	177.185
239 K	8.463	119.763	53.926	58.390	36.700	35.628	174.782	173.800
240 T	9.050	116.984	58.302	53.900	70.554	37.152	172.429	174.754
241 N	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
242 D	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
243 K	8.525	118.113	57.000	57.700	29.700	31.300	176.496	176.989
244 R	8.034	119.835	58.200	56.900	32.100	29.800	174.600	176.300
245 I	8.281	126.485	61.570	58.228	38.000	32.100	176.050	174.684
246 D	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
247 P	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
248 V	7.330	120.178	65.974	64.000	30.600	30.550	178.200	177.240
249 G	7.859	128.565	46.991	65.940	-100.000	30.700	172.000	178.280
250 A	7.904	121.980	54.145	46.911	17.240	-100.000	178.468	172.000
251 C	7.074	113.131	61.933	54.188	-100.000	-100.000	174.000	178.472
252 V	7.980	116.400	65.200	61.800	-100.000	-100.000	180.000	174.100
253 G	7.981	127.337	42.623	65.314	-100.000	-100.000	-100.000	180.034
254 M	7.654	122.137	56.264	-100.000	29.031	-100.000	-100.000	175.303
255 R	7.913	121.345	54.922	56.200	41.100	29.000	178.554	178.029
256 G	8.413	131.504	45.114	54.907	-100.000	41.210	177.472	178.291
257 A	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
258 R	7.984	116.392	58.874	55.300	28.400	19.224	177.300	181.108
259 V	7.529	118.263	64.776	58.900	30.435	28.600	177.188	177.300
260 Q	8.259	119.648	58.214	64.646	27.896	30.500	178.518	177.199
261 A	7.451	123.083	54.961	58.200	17.232	27.900	181.103	178.560
262 V	7.270	119.988	66.333	54.837	30.435	17.594	176.856	181.112
263 S	8.475	112.235	61.779	66.291	63.445	30.600	178.278	176.913
264 T	8.668	115.195	66.355	61.722	68.523	63.394	177.161	178.200
265 E	7.236	122.178	57.942	66.251	28.404	68.280	176.499	177.153
266 L	7.247	117.695	52.146	57.899	39.000	28.011	176.482	176.497
267 G	7.278	127.304	46.700	52.380	-100.000	39.183	-100.000	176.472
268 G	8.601	129.330	43.969	46.713	-100.000	-100.000	-100.000	175.549
269 E	6.774	123.374	58.215	44.059	31.451	-100.000	175.399	174.311
270 R	7.718	130.227	55.720	58.087	29.420	31.565	175.400	175.431
271 I	8.633	125.526	60.280	55.650	39.100	29.534	174.848	175.400

272 D	9.061	128.122	52.780	60.100	41.000	39.180	174.400	175.100
273 I	8.594	124.671	60.100	52.899	36.022	41.214	174.759	174.487
274 V	9.216	124.872	58.900	60.095	34.000	36.000	174.462	174.750
275 L	8.920	126.250	54.390	58.900	41.600	34.100	177.200	174.488
276 W	8.802	124.757	57.581	54.399	-100.000	41.800	174.690	177.200
277 D	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
278 D	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
279 N	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
280 P	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
281 A	7.376	117.792	54.584	64.747	17.232	31.058	178.299	177.436
282 Q	6.939	114.761	56.564	54.704	27.677	17.094	177.759	178.337
283 F	8.363	118.648	58.381	56.370	-100.000	27.700	176.892	177.811
284 V	8.020	118.530	67.200	58.100	30.000	-100.000	177.500	176.900
285 I	7.711	119.950	66.703	67.142	36.529	-100.000	179.186	177.572
286 N	8.272	120.329	55.090	66.523	36.022	36.494	178.834	179.240
287 A	8.743	122.360	53.400	54.900	18.000	36.300	177.100	178.800
288 M	7.778	114.992	53.216	53.589	29.928	17.854	177.085	177.085
289 A	7.071	124.046	51.891	53.054	16.724	30.042	175.419	176.937
290 P	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
291 A	8.752	127.993	52.200	63.277	18.500	33.051	175.777	176.399
292 D	8.271	119.362	52.730	52.260	40.700	18.700	176.000	175.700
293 V	8.518	123.750	62.567	52.780	31.959	40.706	175.304	176.000
294 A	8.943	131.270	53.531	62.455	18.755	32.097	178.061	175.205
295 S	7.678	110.076	57.368	53.451	64.460	18.870	171.096	178.104
296 I	8.190	120.550	59.100	57.293	41.400	64.700	174.000	171.000
297 V	8.826	127.407	61.177	59.338	33.773	41.700	175.134	174.056
298 V	8.829	128.846	62.000	61.049	32.975	33.597	175.142	175.142
299 D	9.051	127.949	52.255	62.100	40.800	32.728	176.900	175.122
300 E	9.180	124.034	59.030	52.226	29.420	40.706	179.194	176.957
301 D	8.428	119.773	56.900	58.900	40.300	29.200	178.000	179.200
302 K	7.703	115.901	55.216	56.982	31.959	40.400	176.556	178.016
303 H	8.242	123.409	56.100	55.200	27.008	32.069	173.524	176.593
304 T	8.136	109.829	59.476	56.245	72.078	26.995	173.175	173.572
305 M	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
306 D	8.684	123.320	52.700	54.632	41.200	34.612	174.500	174.008
307 I	9.148	124.666	59.500	52.733	39.500	41.000	173.400	174.300
308 A	8.796	127.889	49.067	59.597	20.787	39.691	177.020	173.481

309 V	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
310 E	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
311 A	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
312 G	7.895	128.948	45.739	56.771	-100.000	-100.000	-100.000	176.548
313 N	8.133	117.137	52.272	45.785	40.084	-100.000	175.544	174.474
314 L	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
315 A	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
316 Q	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
317 A	7.999	120.735	54.400	57.100	18.500	28.300	178.479	176.271
318 I	8.154	113.893	63.673	54.436	38.053	18.521	179.213	178.552
319 G	7.783	129.848	43.969	63.449	-100.000	38.428	179.242	179.248
320 R	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
321 N	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
322 G	8.317	128.710	42.038	-100.000	-100.000	-100.000	-100.000	173.800
323 Q	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
324 N	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
325 V	7.258	116.494	64.525	57.426	31.000	38.200	-100.000	176.729
326 R	8.180	122.674	55.200	64.524	40.800	31.256	179.497	177.340
327 L	8.117	117.610	57.200	55.200	41.600	-100.000	-100.000	179.500
328 A	8.345	121.045	55.245	57.210	17.740	41.600	180.494	179.320
329 S	8.691	118.128	60.600	-100.000	63.064	-100.000	-100.000	180.446
330 Q	8.476	121.739	58.278	60.500	28.500	-100.000	178.940	177.310
331 L	8.569	118.116	56.777	58.110	42.623	28.518	177.024	178.896
332 S	7.682	129.900	59.257	56.720	-100.000	-100.000	-100.000	177.027
333 G	8.614	113.376	45.700	59.201	-100.000	65.835	174.150	176.248
334 W	7.944	121.027	56.929	45.760	29.600	-100.000	175.260	174.200
335 E	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
336 L	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
337 N	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
338 V	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
339 M	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
340 T	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000	-100.000
341 V	7.831	118.531	64.825	60.442	-100.000	70.900	-100.000	176.630
342 D	8.044	119.121	56.000	64.800	40.240	-100.000	-100.000	176.158
343 D	7.832	120.382	55.550	56.000	40.200	40.200	177.945	177.803
344 L	7.879	121.118	56.800	55.300	41.000	40.400	178.500	177.681
345 Q	7.992	117.839	56.300	56.800	28.000	41.000	176.600	178.602

346 A	7.675	121.336	52.300	56.400	18.247	28.011	177.792	176.601
347 K	7.762	119.608	56.349	52.249	32.467	18.302	175.626	177.808
348 H	7.719	124.828	56.999	56.202	30.435	32.581	179.724	175.646

AA – Amino acid; CA – C^α chemical shift; CB – C^β chemical shift; -100.000 represents the unassigned chemical shifts; (*) indicates the amino acid residues in the tag region; the numbering starts from 133-348 of NusA representing S1+KH1+KH2 domain.

10 Acknowledgements

At this juncture, I would like to extend my gratitude to all the people who have contributed to my work in one way or the other during the course of my doctoral study.

First and foremost I feel immense pleasure in availing this opportunity for expressing my deepest sense of sincere gratitude and regards to my supervisor Prof. Dr. Paul Rösch for his support and constant encouragement, earnest guidance, constructive suggestions, and excellent research facilities extended during my entire doctoral research work. I am greatly indebted to him for having given me an opportunity to work in a very interesting topic and especially in a very healthy and friendly atmosphere.

My sincere thanks to Dr. Kristian Schweimer, who introduced me to the arena of biomolecular NMR spectroscopy. He had guided and supported me during all these years. He was always being very patient in teaching even very minute things which has definitely helped me a lot to learn the nuances of research work. Without this opportunity, I would not have got a chance to assess my research abilities.

I would like to acknowledge Prof. Dr. Birgitta Wöhrl for her timely help, scientific discussions and giving valuable suggestions in the practical aspects of molecular biology.

I would especially like to acknowledge the contribution of Dr. Stefan Prasch, who has been a good mate over this period sharing his invaluable suggestions and explaining various practical approaches with patience. I appreciate his efforts in correcting my thesis meticulously. Many thanks to him, for sharing his thoughts with me and suggesting the way forward at several points in my research.

I express my deep gratitude to Rainer Hofmann who had not only made computers running for us but provided lots of information, helpful tips and timely support. I enjoyed his company during many lunch times with him. I also thank him for his cheerful face and extremely helping nature.

I extend my appreciation and thanks to Dr. Irena Matecko for her support, caring, and providing me with lots of scientific information.

I owe a huge debt of thanks to all my wonderful colleagues who made my stay in the lab to be memorable forever. Thanks to Sabine Wenzel for a nice company and making me stress free during the work in the lab by talking “Flavor of love” stories. Björn Burmann and Marcel Jurk, not only being a good pals, but helped me a lot in scientific discussions with their good comments. Hanna Berkner, thanks a lot for cheering me up many times during the toughest period of my work.

Now, its my turn to thank the rest of present and ex members, Dr. Stephan Schwarzinger, Maximilian Hartl, Dr. Christian Mangels, Claudia knake, Philipp Weiglmeier, Britta Zimmermann, Ramona Heissmann, Ulrike Persau, Rao Jampani, Ulrich Scheckenhofer, Nadine Herz, and Katrin Weiss for a very nice working atmosphere. Many special thanks and love to Andrea Hager for not only helping me in ordering the chemicals but providing lots of love and care personally.

My lots of thanks to Gudrun Wagner, Angela Rösler, and Violaine Zigan for all the administration helps. Special thanks to Violaine Zigan who has always being endowed with lot of patience in helping me out with lot of administration related things.

I want to take this moment to thank my wonderful friend Punnagai (smiley) for spending some splendid time together and having lots of fun with cocktails. She really deserves a special gratitude for being one of my best buddy ever.

My special thanks to my husband Jay for his love, ongoing support, encouragement, and with who exchange of ideas, scientific and otherwise have enriched me a lot.

I am indebted to thank all my brothers for their love and helping me to come through the tough times. In particular, I wish to thank my younger brother Bobby for his affection, inspiration, care and all time support.

Last, but yet the most significant in my life is to thank my dad who stood behind me for everything, providing his incredible support, care, love through out of my life, cheering me up, and making me feel special. Without him I would not have reached to this point of success in my life.

11 Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet habe.

Ferner erkläre ich, dass ich nicht anderweitig mit oder ohne Erfolg versucht habe, eine Dissertation einzureichen oder mich der Doktorprüfung zu unterziehen.

Bayreuth, den

Pagadala Santhanam Sujatha

